Chaotic Neural Networks and Their Applications

Yuyao He
College of Marine Engineering
Northwestern Polytechnical University
Xi'an, Shaanxi 710072, P.R.China

Lipo Wang
Shool of Electrical and Electronic Engineering
Nanyang Technological University
Block S2, Nanyang Avenue, Singapore 639798

Abstract – Many difficult combinatorial optimization problems arising from science and technology are often difficult to solve exactly. Hence a great number of approximate algorithms for solving combinatorial optimization problems have been developed [10], [15]. Hopfield and Tank applied the continuous-time, continuous-output Hopfield neural network (CTCO-HNN) to TSP, thereby initiating a new approach to optimization problems. But Hopfield neural network is often trapped in local minima because of its gradient descent property. A number of modifications have been done on Hopfield neural network for escaping from local minima. As so far, incorporating chaos into the Hopfield neural network has been proved to be successful approach to improve the convergent property of the HNNs. 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 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 has more powerful ability to obtain global minima than any other chaotic neural network model when applied to difficult combinatorial optimization problems.

Keywords: Neural Networks, Transient Chaos, Simulated Annealing, Combinatorial Optimization Problems, TSP

I. INTRODUCTION

Since Hopfield and Tank first applied their CTCO-HNN [8] to solving the TSP, the HNNs [6], [7] have been shown to provide powerful tools for a wide variety of combinatorial optimization problems. However, using the CTCO-HNN to solve TSP suffers from several shortcomings. First, the network can often be trapped at a local minimum on the complex energy terrain. Second, a CTCO-HNN may converge to an infeasible solution. Third, sometimes, a CTCO-HNN does not converge at all within prescribed iteration [13].

Chaotic neural networks have recently received attention due to the potential capability for information processing [4], [12]. 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 minima when applied to difficult combinatorial optimization problems. Based on chaotic property of biological neurons, Aihara *et al* [1] proposed a chaotic neural network (CNN) which includes relative refractoriness in the model. Yamada *et al* [14] solved TSP using CNN model. They found that the relative refractoriness in CNN improves optimization by escaping from local minima. Based on the CNN model, Chen and Aihara proposed chaotic simulated annealing (CSA) to illustrate the features and effectiveness of a transiently chaotic neural network (TCNN) in solving optimization problem [2]. 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 [10]. 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 Δt, where the dynamics are chaotic, and then gradually reducing the time step Δt, the system be assured to converge and to minimize the CTCO-HNN energy function. Since in the limit of $\Delta t \rightarrow 0$, the system approaches the CTCO-HNN, which is stable and minimizes the CTCO-HNN energy function. This approach eliminates the need for difficult choice of any other system parameter.

Hayakawa *et al* [5] proposed another approach for TSP by adding chaotic noise to the discretized CTCO-HNN. 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 their paper.

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, we propose a chaotic neural network which is initially chaotic but eventually 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 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.

II. 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.

A Chaotic Simulated Annealing with Decaying Self-Coupling

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

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

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

$$z_i(t+1) = (1-\beta)z_i(t) \quad (3)$$

where $(i = 1,2,\cdots,n)$

$x_i(t)$ — 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,

α — positive scaling parameter for neuronal inputs,

k — damping factor of nerve membrane, $0 \le k \le 1$,

$z_i(t)$ — self-feedback connection weight (refractory strength) ≥ 0,

β — damping factor of $z_i(t)$, $0 < \beta < 1$,

I_0 — positive parameter,

ε — 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 β.

B Chaotic Simulated Annealing with Decreasing Time-step

Considering the Euler approximation of the CTCO-HNN,

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

Wang & Smith [11] 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 Δt 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. For asynchronous updating of eq. (4), together with a symmetric w, $\tau > 0$, and an increasing activation function with a maximum slope of μ_{max}, the network stabilizes if [$w_{ii} \le 0$ and $\Delta t \le \tau$].

C Discretized Continuous Hopfield Network with Chaotic Noise

Hayakawa [5] investigated adding 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 $\eta_i(t)$ whose amplitude is normalized to be unity, 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 σ_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 solution feasibility at each of iteration a discretised state variable is introduced,

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

Since this system with chaotic noise is no longer convergent, a statistical average over a period of time should be employed to determine the state of a neuron. Parameters to be adjusted are the logistic map parameter a in (7), and the amplification factor A in (6).

D Simulated Annealing with Decaying Chaotic Noise

We propose a chaotic neural network model that is described as

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

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

where $\eta_i(t)$ — chaotic noise for input to $y_i(t)$, ($0 < \eta_i(t) < 1$).

γ —positive scaling parameter for chaotic noise,

Other parameters have the same meaning as in (1)-(2).

The chaotic noise can be generated from the logistic map:

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

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

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

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

and

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

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

III. APPLICATION OF THE ABOVE METHODS TO THE TRAVELLING SALESMAN PROBLEM

We adopt the formulation of Hopfield and Tank [8] 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 to visit city i in visiting order k. A computational energy function which is to minimize the total tour length while simultaneously satisfying all constraints takes the following form:

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

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 [8].

A Chaotic Simulated Annealing with Decaying Self-Coupling

In Chen and Aihara's original paper [2], they applied transiently chaotic neural network to 10-city TSP. The parameters were set as follows:

$$k = 0.9 \qquad \alpha = 0.015 \qquad \varepsilon = 1/250$$

$$I_0 = 0.65 \qquad z(0) = 0.8$$

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.

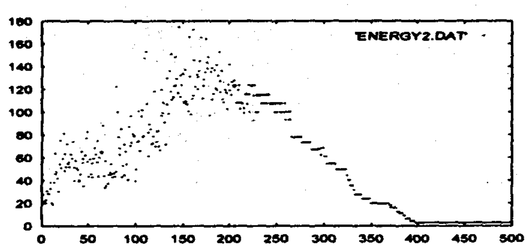

Fig. 1 Time evolution of the discrete energy function for 10 city TSP with Chen and Aihara's Method.

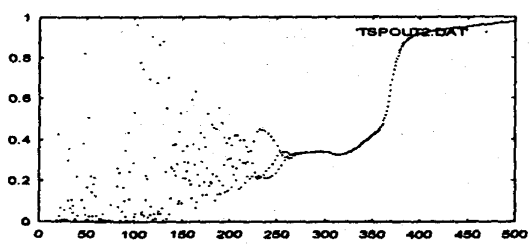

Fig. 2 Time evolution of the neuron output x23 or 10 city TSP with Chen and Aihara's Method.

Table 1. Results of 5000 Different Initial Conditions for Each Value of β on 10 city TSP with Chen and Aihara's Method.

β	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

B 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 α. The initial time step $\Delta t(0)$ and the annealing factor β should be carefully adjusted. The parameters we choose are as follows:

$$\Delta t(0) = 2 \qquad \tau = 0.5 \qquad \alpha = 0.015$$

We only adjust the parameter β. 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.

Table 2. Results of 100 Different Initial Conditions for Each Value of β on 10 city TSP with Wang and Smith's Method.

β	0.010	0.005	0.001
NG	65	81	92
RG	65%	81%	92%
NI	130	225	965

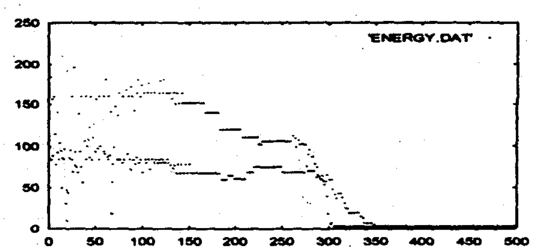

Fig. 3 Time evolution of the discrete energy function for 10 city TSP with Wang and Smith's Method.

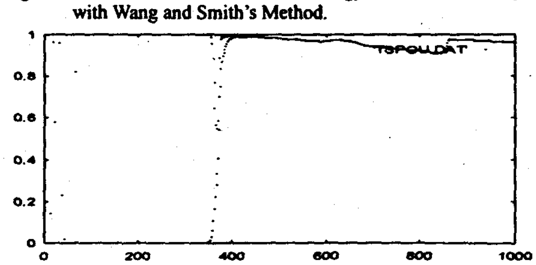

Fig. 4 Time evolution of the neuron output x23 or 10 city TSP with Wang and Smith's Method.

C 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. Iterations less than 1000 correspond to the transient period. 2000 is chosen to be the maximum number of iterations. In order to determine the solution feasibility the visiting frequency to the optimal solutions is tested:

$$P_{OS} = \frac{\textit{number of steps staying at the optimal solution}}{\textit{total steps}} \quad (16)$$

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 in 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 were found to have good optimization ability on TSP.

Table 3. Results of 100 Different Initial Conditions for Each Value of 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

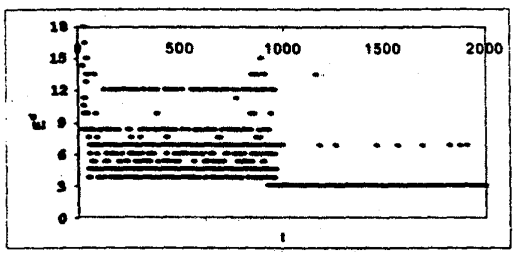

Fig. 5 Time evolution of the discrete energy function for 10 city TSP with Hayakawa et al's Method.

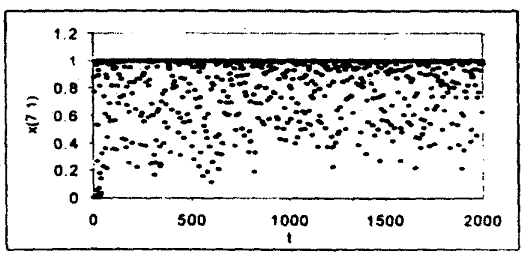

Fig. 6 Time evolution of the neuron output x71 for 10 city TSP with Hayakawa et al's Method.

D Simulated Annealing with Decaying Chaotic Noise

We analyze 10-city TSP using our approach described in section II. D. The parameters are set as follows:

$$\alpha = 0.015 \qquad \gamma = 0.1 \qquad \varepsilon = 1/250 \qquad a(0) = 3.9$$

$$a_0 = 2.5 \qquad h = 1 - 1/a_0 = 0.6$$

We vary only β to investigate the dynamics of our chaotic neural network model. The results obtained are summarized in Table 4. Fig. 7 and Fig. 8 show up 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 of β on 10 city TSP with He and Wang's Method.

β	0.1	0.015	0.003	0.001
NG	5000	5000	5000	5000
RG	100%	100%	100%	100%
NI	19	82	350	1022

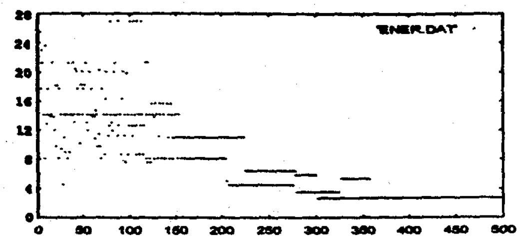

Fig. 7 Time evolution of the discrete energy function for 10 city TSP with He and Wang's Method.

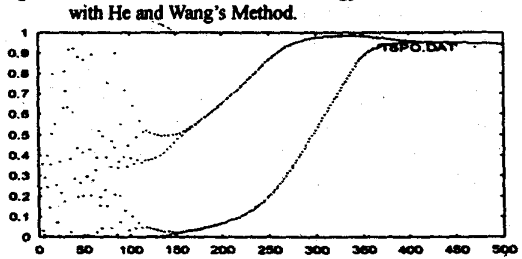

Fig. 8 Time evolution of the neuron output x23 for 10 city TSP with He and Wang's Method.

IV. DISCUSSION

In this section, we illustrate the properties and optimization ability of chaos in neural networks by analyzing the results obtained by solving the 10-city problem.

With Chen & Aihara's method of CSA with decaying self-coupling 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 states have a little fluctuation and then tend 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. The above characteristic transition, together with the proof of the existence of chaos in this CSA scheme by Chen & Aihara [3], strongly suggest a chaotic role in the process. As shown in Table 1 the rate of global optimization is 98.8% for $\beta = 0.015$, 100% of the rate of global optimization 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 & 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. This illustrates the fact that decaying the time-step (an inherent quantity to the network) is fundamentally different from decaying the self-coupling term (an externally introduced quantity). In general, this method requires fewer parameters to be adjusted, but the difficulty arises when choosing an effective annealing rate β.

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 at the beginning where 0 is often visited. The final state of this neuron is 1 according to (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. The ability of a network with chaotic noise having a persistent attraction towards an optimal state was also reported. Two values of the logistic parameter a found to have high optimization ability are used, which yield a maximum of 7/10 feasibility as shown in Table 3. This is comparatively low against the other methods discussed so far.

We can see from Table 4 that β can be varied 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). Where among 5000 cases, 100 percent are to global minima (tour length=2.690670). When β 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 β is that the number of iterations is different.

V. CONCLUSIONS

We have introduced four methods of incorporating chaos into the Hopfield network for combinatorial optimization. Two such methods, namely Chen & Aihara's CSA with a decaying self-coupling term and Wang & 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 first three methods, our method is to add chaotic noise to each neuron of DTCO-HNN and then gradually decrease it. The network finally converges to optimal solution. The main feature of our method is the fact that the method is more efficient, the rate of optimization is higher than any other method and the parameter can be easily chosen.

VI. REFERENCES

[1] Aihara K, Takabe T, Toyoda M. Chaotic Neural Networks. Physics Letter A, 1990, 144 (6,7) :333-340

[2] Chen L, Aihara K. Chaotic Simulated Annealing by a Neural Network Model with Transient Chaos. Neural Networks, 1995, 8(6): 915-930.

[3] Chen L, Aihara K. Chaos and Asymptotical Stability in Discrete-time Neural Networks. Physica D, 1997, Vol. 104: 286-325.

[4] Freeman W J. Tutorial on Neurobiology: from Single Neurons to Brain Chaos. International Journal of Bifurcation and Chaos, 1992, 2 (3):451-482.

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

[6] Hopfield J J. Neural Networks and Physical Systems with Emergent Collective Computational Abilities. Proc. Nat. Acad. Sci. Vol. 79, 1982:2554-2558.

[7] Hopfield J J. Neurons with Graded Response have Collective Computational Properties like those of Two-state Neurons. Proc. Nat. Acad. Sci., Vol. 81, 1984:3088-3092.

[8] Hopfield J J, Tank D W. "Neural" Computation of Decisions in Optimization Problems. Biological Cybernetics, *1985, Vol* 52:141-152.

[9] Kirkpatrick S, Gelatt C D, Vecchi M P. Optimisation by Simulated Annealing. *Science*, 1983, Vol. 220:671-680.

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

[11] Wang L, Smith K. On Chaotic Simulated Annealing. IEEE Transactions on NN, 1998, 9(4):716-718.

[12] Wang L. Effects of Noise in Training Patterns on the Memory Capacity of the Fully Connected Binary Hopfield Neural Network: Mean-Field Theory and Simulations. *IEEE Transactions on Neural Networks*, 1998, 9 (4):697-704.

[13] Wilson G V, Pawley G S. On the stability of the traveling salesman problem algorithm of Hopfield and Tank. Biol Cybern., 1988, Vol. 58:63-70.

[14] Yamada T, Aihara K, Kotani M. Chaotic Neural Networks and The Travelling Salesman Problem. *Proceedings of 1993 International Joint Conference on Neural Networks*, 1993: 1549-1552.

[15] Zanakis S H, Evans J R, Vazacopoulos A A. Heuristic Methods and Applications: a Categorized Survey. Europ. J. Operat. Res., 1989, Vol. 43:88-110.