Chaotic Simulated Annealing with Decaying Chaotic Noise

Yuyao He [1] Lipo Wang [2] and Demin Xu [1]

[1]College of Marine Engineering
Northwestern Polytechnical University
Xi'an, Shaanxi 710072, P.R.China
Email: heyyao@nwpu.edu.cn

[2]Shool of Electrical and Electronic Engineering
Nanyang Technological University
Block S2, Nanyang Avenue, Singapore 639798
Email: elpwang@ntu.edu.sg

Abstract

By adding chaotic noise to each neuron of the discrete-time, continuous-output Hopfield neural network (DTCO-HNN) and gradually reducing the noise, we propose a neural network that is initially chaotic but eventually convergent, and thus has richer and more flexible dynamics compared to the HNN. We apply the proposed network to the Traveling Salesman Problem (TSP) and find that results are highly satisfactory. That is, the transient chaos enables the network to escape from local energy minima and to find global minima in 100% of the simulations for 4-city and 10-city TSP, as well as near-optimal solution in every run for 48-city TSP.

Index Terms: Neural Networks, Transiently Chaotic Noise, Combinatorial Optimization Problem, TSP.

1. Introduction

Combinatorial optimization problems arising from science and technology can be described as a search for the best solutions. Because a number of interesting combinatorial optimization problems is difficult to find exact solutions due to their NP hardness, many researchers have been paying much attention to approximate or heuristic algorithms[1], [2].

Since Hopfield and Tank applied a continuous-time, continuous-output Hopfield Neural Network (CTCO-HNN) [3] to solving the traveling salesman problem, the HNNs [4], [5] have been recognized as powerful tools for combinatorial optimization problems. However, using HNNs to solve the TSP suffers from several shortcomings. First, the network can often be trapped at a local minimum on the complex energy terrain. Second, a HNN may not converge to a valid tour. Third, a HNN sometimes does not converge at all within a prescribed number of iteration

[6].

Chaotic neural networks have recently received attention due to the potential capability for information processing [7], [8]. In a number of methods proposed to prevent the HNN from being trapped at local minima, chaotic neural networks have been shown to be powerful approaches. Based on chaotic properties of biological neurons[9] Chen and Aihara recently proposed "chaotic simulated annealing" to illustrate the features and effectiveness of a transiently chaotic neural network (TCNN) in solving optimization problems [10]. By adding a negative self-coupling to a Hopfield-like neural network and gradually removing it, they used the transient chaos for searching and self-organizing. Compared to other neural network methods [11], [12], the Chen--Aihara approach significantly increased the probability of finding near-optimal solutions. However, the Chen--Aihara method requires that a number of network parameters must be subtly adjusted so as to guarantee the convergence of the TCNN and its minimization of the energy upon the removal of the transient chaos [13].

Wang and Smith presented an alternate approach to chaotic simulated annealing [14]. 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 is guaranteed to converge and to minimize the CTCO-HNN energy function. Since in the limit of $\Delta t \to 0$, the network approaches the CTCO-HNN. This approach eliminates the need for difficult choices of any other system parameter.

Hayakawa et al [15] proposed another approach for TSP by adding chaotic noise to the discretized CTCO-HNN. Since the system with chaotic noise does not converge, a statistical average over a period of time is used to determine the state of a neuron. Because this technique has hill-climbing features, the rate of reaching global optimization is lower.

In this paper we add chaotic noise to each neuron of DTCO-HNN and gradually reduce the noise. As the chaotic noise approaches zero, the network becomes the DTCO-HNN, thereby stabilizing and minimizing the energy. We apply our method to Travelling Salesman Problem and find that results are remarkably good: the network is very likely to find global optimal solutions or near-optimal solutions after transiently chaotic searching.

2. Chaotic Simulated Annealing with Decaying Chaotic Noise

The chaotic neural network that we propose in this paper is defined as follows:

$$u_i(t+1) = \alpha(\sum_j w_{ij} v_j + I_i) + \gamma \eta_i(t) \qquad (1)$$

$$v_i(t) = f(u_i(t)) = \frac{1}{1+e^{-u_i(t)/\varepsilon}} \qquad (2)$$

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

$v_i(t)-$ output of neuron i,

$u_i(t)-$ 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,

$\gamma-$ positive scaling parameter for the chaotic noise,

$\alpha-$ positive scaling parameter for neuronal inputs,

$\eta_i(t)-$ chaotic noise for neuron i,

$\varepsilon-$ gain of the output function, $\varepsilon > 0$.

The chaotic noise can be generated from the logistic map:

$$\eta_i(t) = z_i(t) - h \qquad (3)$$

$$z_i(t+1) = a(t) z_i(t)(1 - z_i(t)) \qquad (4)$$

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

$$a(t+1) = (1-\beta) a(t) + \beta \cdot a_0 \qquad (5)$$

and

$$h = z^* = 1 - \frac{1}{a_0} \qquad (6)$$

where $1 \le a_0 \le 2.9$. $0 < \beta < 1$ is a decay rate.

With $\eta_i(t) = 0$, the network described by eqs. $(1) \sim (4)$ reduces to the DTCO-HNN. We will show that the transient chaos induced by the decaying chaotic noise is very useful in preventing the network from being trapped in local minima. In our model, the chaotic noise is assigned to each neuron, and the noise generators for the neurons operate with the same rule but different initial conditions.

3. Application to the Travelling Salesman Problem

To verify and illustrate the features and effectiveness of the proposed approach of simulated annealing with chaos, we apply it to TSP in this section. In the Travelling Salesman Problem, the salesman is to visit all n cities once and only once, returning to his starting point after travelling the minimum total distance.

We adopt the formulation of Hopfield and Tank [3] for TSP, namely, a solution of TSP with n cities is represented by the output of an $n \times n$ network, with $v_{ik} = 1$ signifying that the salesman visits city i in order k. A computational energy function used to minimize the total tour length while simultaneously assuring that all constraints are satisfied takes the following form:

$$E = \frac{A}{2}\{\sum_{i=1}^{n}(\sum_{k=1}^{n} v_{ik} - 1)^2 + \sum_{k=1}^{n}(\sum_{i=1}^{n} v_{ik} - 1)^2\} + \frac{B}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}(v_{jk+1} + v_{jk-1})v_{ik}d_{ij} \qquad (7)$$

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

This energy can be cast into a general energy expression

$$E = -\frac{1}{2}\sum_{ik}\sum_{jl} w_{ikjl} v_{ik} v_{jl} - \sum_{ik} I_{ik} v_{ik} \qquad (8)$$

if we choose the weight between neuron ik and neuron jl as

$$w_{ikjl} = -A[\delta_{ij}(1-\delta_{kl}) + \delta_{kl}(1-\delta_{ij})] - Bd_{ij}(\delta_{lk+1} + \delta_{lk-1}) \quad (9)$$

and the input bias to neuron ik as

$$I_{ik} = A \qquad (10)$$

where $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ otherwise.

The chaotic discrete dynamics of the neural network for TSP can be obtained from eqs. (1) ～ (5):

$$u_{ik}(t+1) = \alpha\{-A(\sum_{l\neq k}^{n} v_{il}(t) + \sum_{j\neq i}^{n} v_{jk}(t)) - B\sum_{j\neq i}^{n} d_{ij}(v_{jk+1}(t) + v_{jk-1}(t)) + A\} + \gamma\eta_{ik}(t) \qquad (11)$$

$$v_{ik}(t) = f(u_{ik}(t)) = \frac{1}{1+e^{-u_{ik}(t)/\varepsilon}} \qquad (12)$$

$$\eta_{ik}(t) = z_{ik}(t) - h \qquad (13)$$

$$z_{ik}(t+1) = a(t)z_{ik}(t)(1 - z_{ik}(t)) \qquad (14)$$

$$a(t+1) = (1-\beta)a(t) + \beta \cdot a_0 \qquad (15)$$

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

Although the dynamics of equation (11) is discrete, the output v_{ik} from equation (12) takes a continuous value between 0 and 1. The corresponding energy function of equation (7) is also continuous. Since a solution to TSP requires the states of the neurons to be either 0 or 1, we introduce a discrete output as follows

$$v_{ik}^{d} = \begin{cases} 1 & \text{iff} \quad v_{ik} > \sum_{i}^{n}\sum_{k}^{n} v_{ik}(t)/n \times n \\ 0 & \text{otherwise} \end{cases} \qquad (16)$$

In our simulation, we calculate simultaneously the continuous energy function E^{C} using eq. (7) and the discrete energy function E^{d} by replacing v_{ik} with v_{ik}^{d} in eq. (7).

In the following studies, we use the same initial value of $a(0)$ and the same decaying rule for $a(t)$ in eq. (14) for all neurons for simplicity. The parameters in the equations (7), (11) ～ (15) are set as follows:

$$\mathrm{A = B = 1.0} \quad \alpha = 0.015 \quad \varepsilon = 0.004 \quad \mathrm{a(0) = 3.9} \quad \mathrm{a_0 = 2.5} \quad h = 1 - \frac{1}{a_0} = 0.6 \qquad (17)$$

We vary β and γ to investigate the dynamics of the proposed chaotic neural network model. The focus of our simulations is on the optimization performance (the rate of global optimization) and efficiency (the average number of iterations) with various values of parameter β and γ. We use asynchronous cyclic updating of the neural network model. One of iteration means one cyclic updating of all neuron states and $a(t)$ is decreased by one step after each of iteration.

We first analyze 4-city TSP with data originally used by Hopfield and Tank [3] with 5000 randomly generated initial conditions of $u_{ik} \in [1,-1]$ and $z_{ik} \in (0,1)$ for $\beta = 0.003$ and $\gamma = 0.1$.

Fig.1 shows the time evolutions of the continuous energy function E^c and the discrete energy function E^d. The shortest route or the global energy minimum is 2-3-4-1-2 (tour length=1.341768). In addition, we find that other representations of the shortest routes with the same energy value, such as 3-4-1-2-3 and 1-2-3-4-1 can be obtained, depending on the initial conditions of u_{ik}.

Fig.2 shows the time evolutions of the neuron output $v_{23}(t)$ and $v_{24}(t)$ which illustrate chaotic behavior during the first 120 iterations. They go through a reversed period-doubling bifurcation after iteration 120. As chaotic noise input $\eta_{ik}(t)$ becomes smaller, the neural network finally converges to a fixed point corresponding to a global minimum. In our simulations of 4-city TSP with 5000 different initial conditions of u_{ik} for each of $\beta = 0.05, 0.015, 0.008, 0.005, 0.003, 0.001$ and $\gamma = 0.1$, the rates of reaching global optimization are all 100%.

We now analyze 10-city TSP with data in [3]. The system parameters are set as in eq. (17) besides β and γ. The results obtained are summarized in Table 1 and 2 (GM= Rate of reaching the global minimum, NI= Number of iteration required for the network to converge). In Table 1, the decaying rate β varies from a larger value of 0.1 to a smaller value of 0.001 and $\gamma = 0.1$. The number of iterations required for the network to converge (for the discrete energy function) increases as β decreases. 5000 simulations are run for each β and the network finds the global minimum (tour length=2.690670) in all simulations. In Table 2, we fix $\beta = 0.015$ and change γ from 0.1 to 1.0. We find the rate of reaching the global minimum is still 100 percent for different γ. As γ gets larger, the system becomes more chaotic and needs more iteration to converge to the global minimum.

Finally, we apply the proposed approach to ATT48-city TSP [16]. The parameters are set as follows:

$\alpha = 0.015$ $\varepsilon = 0.004$ $a(0) = 3.9$ $a_0 = 2.0$ $h = 0.5$ $\gamma = 1.5$ $A = 1.0$ $B = 1.445$

In our simulation, we use Pseudo-Euclidean Distance. The edge weight type ATT corresponds to a special "Pseudo-Euclidean" distance function [16].

The decaying rate β in ATT48 city TSP has two values, i.e. when $3.0 \le a(t) \le 3.9$, we choose $\beta_1 = \beta = 0.5 \times 10^{-4}$; when $2.0 \le a(t) \le 3.0$, we adopt $\beta_2 = (2 \sim 10)\beta_1$ to speed up the convergence of the system. In the following simulation we employ $\beta_2 = 8\beta_1 = 4.0 \times 10^{-4}$. The results with 200 different initial conditions are summarized in Table 3, where the number of iteration is 20236. The simulation results show that 22% of the solutions converge to the best tour length (=10628). While 73% solutions converge to near-optimal routes.

In the end of this Section we list our results and Chen & Aihara's results in Table 4 for comparison in Section 4.

4. CONCLUSIONS and DISCUSSIONS

We have proposed an alternative approach to simulated annealing using decaying chaotic noise for solving combinatorial optimization problems. The dynamics of our model is composed of two combined dynamics. Chaotic dynamics at early stage is used to search for finding a valley involving globally optimal solutions. At later stage, the network stabilizes. Compared with other models mentioned in section I in terms of optimization performance, computational efficiency and ease of choosing parameters, our new model has many advantages:

1. The rate of global optimization of our approach is higher than that of other approaches [10], [15].

2. The parameters like γ and β can be easily chosen because the performance is insensitive to their choices (Table 1 and 2), although the number of required iterations increases as chaos parameter γ increases or as the decaying rate β decreases.

3. The computational cost is smaller. If the same decaying rate β in our method is chosen as that in Chen & Aihara's method, the computational cost of our method is a little smaller than that of Chen & Aihara's method. But

because the decaying rate β in our method can be chosen much larger than that in other method (Table 4), we can say the computational cost of our method is much smaller.

Compared with other methods, our method has a high feasibility, good robustness and not problem-specific. Like Chen and Aihara's method, Our method has a transiently chaotic simulated annealing property. γ stands for the strength of chaotic noise, and parameter β can be considered as a damping speed parameter of chaotic noise, which controls the simulated annealing speed.

Why is the proposed method so effective? In our model, all noise generators operate with the same initial value of $a(0)$ and the same decaying rule of $a(t)$ for all neurons. That means chaotic noise signals have the same distribution and is mutually correlated. Further more, the chaotic noise within each neuron has short time correlation, as opposed to white noise with no correlation. It has been suggested that correlation in chaotic noise may play an important role for the network system to find the best solution [15]. However, the effect of the distribution and correlation of the noise signal on the performance of the neural network system will be the subject of future research.

The proposed method offers a robust and efficient approach to global optimization, as shown in our experiments on TSP. Applying the approach to other difficult combinatorial optimization problems will also be studied in the forthcoming papers.

References

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

[2] 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.

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

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

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

[6] 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.

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

[8] Wang L. Oscillatory and Chaotic Dynamics in Neural Networks Under Varying Operating Conditions. IEEE Transactions on Neural Networks, 1996, 7(6):1382-1388.

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

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

[11] Hasegawa M, Ikeguchi T, Matozaki T. Solving Combinatorial Optimization Problems by Nonlinear Neural Dynamics. IEEE International Conference on Neural Networks Proceedings, Perth, Vol. 6, 1995:3140-3145.

[12] Nozawa H. A Neural Network Model as a Globally Coupled Map and Applications based on Chaos. Chaos, 1992, 2(3):377-386.

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

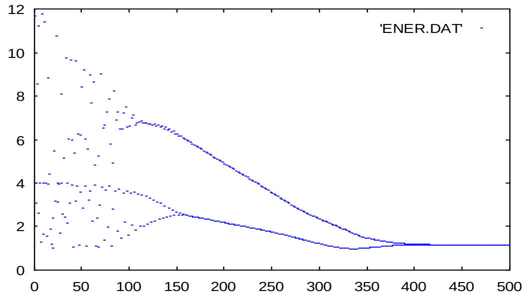
[14] Wang L, Smith K. On Chaotic Simulated Annealing. IEEE Transactions on NN, 1998, 9(4):716-718.

[15] 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.

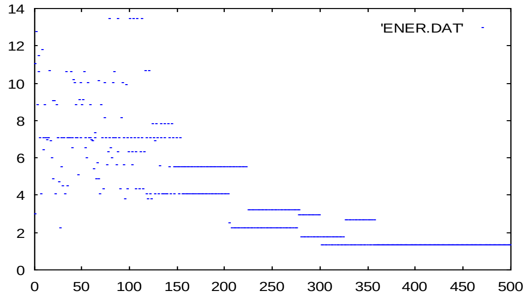
[16] Reinelt G. TSPLIB-A Traveling Salesman Problem Library. ORSA J. On Computing, 1991, 3(4):376-384.

List of Figures

List of Tables

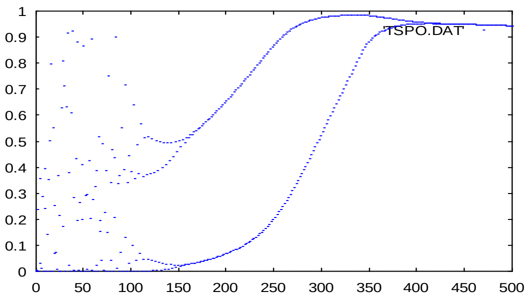


(a)

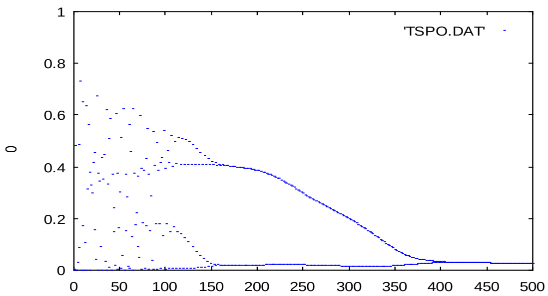


(b)

Fig.1



(a)



(b)

Fig.2

Yuyao He, Lipo Wang

Table 1.

	β					
	0.1	0.05	0.015	0.008	0.003	0.001
GM	100%	100%	100%	100%	100%	100%
NI	19	33	80	138	350	1022

Table 2.

	γ					
	1.0	0.8	0.6	0.45	0.3	0.1
GM	100%	100%	100%	100%	100%	100%
NI	175	159	139	119	93	80

Table 3.

Route Length	Number (%)
10628	22%
10835	63%
10926	10%
11298	5%

Table 4.

TSP	β	He & Wang's Method		Chen & Aihara's method	
		GM	NI	GM	NI
10 city TSP	0.1	100%	19	--	--
	0.05	100%	33	--	--
	0.015	100%	80	98.9%	81
	0.010	100%	112	99.4%	119
	0.005	100%	208	99.9%	234
	0.003	100%	350	100%	398
ATT48 City TSP	$\begin{cases} 0.5\times10^{-4} \\ 4.0\times10^{-4} \end{cases}$ 0.5×10^{-4}	22%	20236	5%	25632

Yuyao He, Lipo Wang