Solving Optimization Problems Using Transiently Chaotic Neural Networks: Choices of Network Parameters

Fuyu Tian*, Lipo Wang* and Yuyao He**

*School of Electrical and Electronic Engineering
Nanyang Technological University
Block S2, Nanyang Avenue, Singapore 639798
Email: elpwang@ntu.edu.sg
URL: http://www.ntu.edu.sg/home/elpwang

**P.O.Box 32
College of Marine Engineering
Northwestern Polytechnical University
Xi'an Shaanxi 710072
People's Republic of China

***Abstract* - Chen and Aiahara recently proposed a transiently chaotic neural network (TCNN). They reported that the TCNN has rich dynamics and higher ability of searching for globally optimal or near-optimal solutions; however, only one set of network parameters was published and it was not clear how the network performance depends on the choice of parameters. In this paper we re-examine the TCNN's solutions to the traveling salesman problem (TSP) with a wide range of network parameters. From our simulations and analysis, we propose a guideline on choosing these parameters. We show that the performance depends on the parameters sensitively.**

1. INTRODUCTION

Since Hopfield and Tank's seminal work on solving the travelling salesman problem (TSP) with a Hopfield neural network (HNN)[1], the HNN's have been recognized as powerful tools for solving combinatorial optimization problems. Although these neural networks can guarantee convergence to a stable equilibrium point due to their gradient descent dynamics, the main drawback is that the networks can often be trapped in local minima, thereby leading to poor solutions to optimization problems[2]. Various simulated annealing techniques have been suggested to overcome this drawback.

Chen and Aihara [5] recently proposed a transient chaotic neural network (TCNN) by modifying a chaotic neural network which Aihara et al proposed earlier[4]. Since the optimization process of the TCNN is deterministically chaotic rather than stochastic, the TCNN is also called as chaotic simulated annealing (CSA), in contrast to the conventional stochastic simulated annealing (SSA)[3]. The mechanics of CSA uses slow damping of negative self-feedback to produce successive bifurcations so that the neurodynamics eventually converges to from strange attractors to a stable equilibrium point.

The TCNN is defined below:

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

$$y_i(t+1) = ky_i(t) + \alpha\Big(\sum_{j=1, j\neq i}^{n} w_{ij}x_j(t) + I_i\Big) - z_i(t)(x_i(t) - I_0) \tag{2}$$

$$z_i(t+1) = (1-\beta)z_i(t) \qquad (i = 1,\dots,n) \tag{3}$$

where

x_i = output of neuron i,

y_i = internal state of neuron j,

$w_{ij} = w_{ji}$; $w_{ii} = 0$; $\sum_{j=1, j\neq i}^{n} w_{ij}x_j + I_i = -\partial E / \partial x_i$ connection weight from neuron j to neuron i,

I_i = input bias of neuron i,

k = damping factor of nerve membrane ($0 \le k \le 1$),

α = positive scaling parameter for inputs,

$z_i(t)$ = self-feedback connection weight or refractory strength $(z(t) \ge 0)$,

β = damping factor of the time dependent ($0 \le \beta \le 1$),

I_0 = positive parameter,

ε = steepness parameter of the output function,(ε >0)

E = energy function,

usually

$E = W_1 E_{constraint} + W_2 E_{optimization}$,

W_1 and W_2 are the coupling parameters corresponding to the constraints and the cost function of tour length, respectively.

There are several network parameters that need to be properly chosen in Chen-Aihara approach, i.e., the relative weighting factors W_1 and W_2 for the constraint and the optimization solution, respectively, the initial self-feedback $z(0)$, the decay rate for the self-feedback β, the relative strength of the energy terms α, and the damping factor of nerve membrane k. In their paper [5], Chen and Aihara examined the influence of α, β on the systems by applying CSA to 4-city and 10-city TSP. For the 48-city TSP, they showed the result on only one set of network parameters. No details were given about how to choose these parameters or how sensitive the performance is with respect to different choices of parameters. However, this is very important for solving combinatorial problems with TCNN. Since we show in this paper that the performance of the TCNN sensitively depends on the choice of system parameters. We propose a guideline on how the parameters should be chosen.

2. PERFORMANCE EVALUATION ON THE TSP

A classical combinatorial optimization problem is the travelling salesman problem (TSP), which is NP-hard. It is to seek the shortest route through n cities, visiting each once and only once, and returning to the starting point. Since Hopfield and Tank applied their neural networks to TSP, TSP has been intensively studied in the field of artificial neurocomputing.

Here the solution of TSP with n cities is represented by a network with $n \times n$ neurons. The neuron output $x_{ij} = 1$ represents visiting city i in visiting order j. The computational energy function to be minimized consists of two parts:

$$E = \frac{W_1}{2}\left\{\sum_{i=1}^{n}\left(\sum_{j=1}^{n} x_{ij} - 1\right)^2 + \sum_{j=1}^{n}\left(\sum_{i=1}^{n} x_{ij} - 1\right)^2\right\} + \frac{W_2}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\left(x_{kj+1} + x_{kj-1}\right)x_{ij} d_{ik} \qquad (4)$$

where $x_{i0} = x_{in}, \quad x_{in+1} = x_{i1}$

d_{ij} is the distance between city i and city j.

The first part, with a coefficient W_1, is minimized (assumes zero) for any valid tour, i. e, when the constraints are satisfied. The second part, with a coefficient W_2, is the tour length. Hence a global minimized of E represents a shortest valid tour.

From equation (1)-(4), the network dynamics of TCNN for TSP can be obtained as follows:

$$x_{ij}(t) = \frac{1}{1+e^{-y_{ij}(t)/\varepsilon}}$$

$$y_{ij}(t) = K + F + E$$

where $K = ky_{ij}$, $F = -Z(t)(x_{ij}(t) - I_0)$

$$E = \alpha\{-W_1(\sum_{l \neq j}^{n} x_{il}(t) + \sum_{k \neq i}^{n} x_{kj}(t)) - W_2(\sum_{k \neq i}^{n} d_{ik} x_{kj+1}(t) + \sum_{k \neq i}^{n} d_{ik} x_{kj-1}(t)) + W_1\}$$

$$(i, j = 1, \ldots, n)$$

$$z(t+1) = (1-\beta)z(t)$$

In our experiment, the updating scheme is cyclic and asynchronous. It means that all neurons are cyclically updated in a fixed order. When all the neurons are updated once, we call it one iteration. Once the state of a neuron is updated, the new state information is immediately available to the other neurons in the network (asynchronous).

The initial neuron inputs are generated randomly between [-1,1]. After each update, the new energy is compared to the old. If the energy does not change by more than a threshold 10^{-10}, the system is assumed to have reached a stable state.

To help us determine the network parameters, we analyze relative magnitudes of the three parts in the input to a neuron in the TCNN model. For convenience, we call them the membrane term (decay from the previous membrane state, represented by K), the energy term (tour validity and length, represented by E), and the self-feedback term (cause of chaos, represented by F).

Hasegawa has investigated the relationship between solving abilities and decay parameter k. So we fix the value of k to 0.90 which can give high solving ability of TSP [15]. In their experiments on TSP, $k = 0.90$ was also used by Chen & Aihara in their experiments on CSA.

For each set of network parameters chosen, we did 100 simulations, each with a different random initial condition.

2.1 EXPERIMENTS ON TSP WITH 4 CITIES AND 10 CITIES

In this subsection we use the Hopfield-Tank original data[2] on TSP with 4 and 10 cities.

Chen and Aihara's experiments show that α represents the influence of the energy function on the neurodynamics. If α is too large, the influence of the energy function becomes too strong to generate the transient chaos. On the other hand, the energy function cannot be sufficiently reflected in the neurodynamics if α is too small. The parameter β can be considered as damping speed parameter of the negative self-feedback strength, which controls the annealing schedule. To produce longer chaotic dynamics, β should be set smaller. But too small β will make the convergence time too long. For 4 and 10 cities, β =0.001 is sufficiently small.

Now we consider different W_2 with W_1=1. Other parameters are set as follows:

$$k = 0.9; \varepsilon = 1/250; I_0 = 0.65; z(0) = 0.08; \alpha = 0.015; \beta = 0.001$$

In both 4-city and 10-city cases, we find that the solving abilities are the highest when the three terms, i.e., the membrane term K, the energy term E, and the self-feedback term F, are comparable in magnitude as shown in Figure 1, with the parameters chosen.

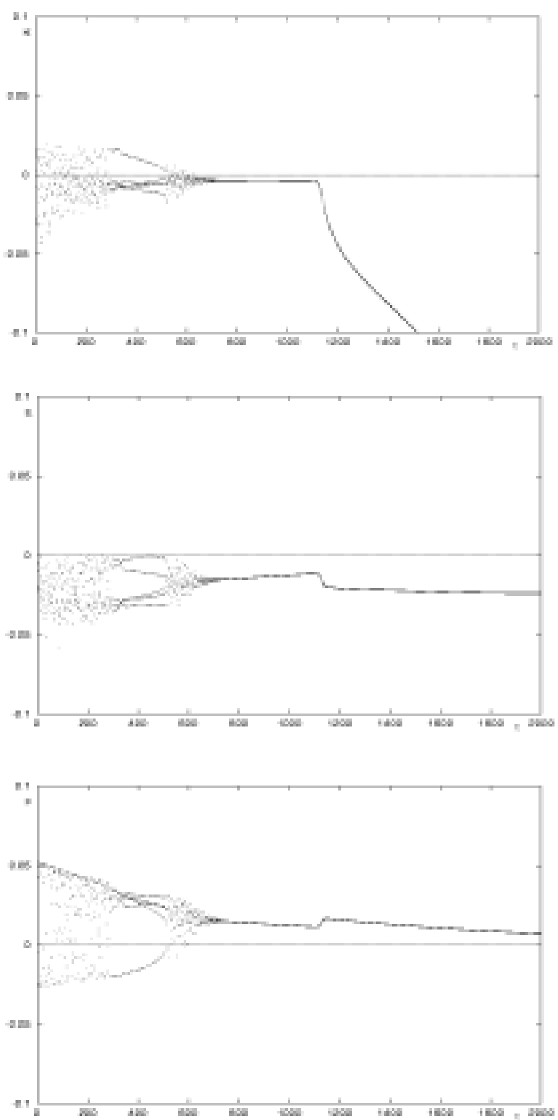


Figure 2 Term K, E, F against t

2.2 EXPERIMENTS ON TSP WITH 48 USA CITIES

In this subsection we use the data of 48 USA cities[6], which do not fit into a unit square as in the cases of 4 and 10 cities. We thus normalize the data by dividing all distances between the cities by the longest distance 2662. After the network converges to a solution, the distances are then converted back before the tour-length is calculated.

The parameters are as follows:

$$k = 0.9; \varepsilon = 1/250; I_0 = 0.5; z(0) = 0.10;$$
$$\alpha = 0.015; \beta = 5.0 \times 10^{-5}$$

We evaluate the performance of TCNN on 48 USA city problem for different choices of W_2 with $W_1 = 1$.

When $1.7 \le W_2 \le 2.0$, 100% of the solutions converge to a near-optimal route (tour length 10805, as compared to the shortest tour with length 10628). The performance of the TCNN decreases as W_2 deviates from this range. For example, when W_2=1.0, 100% converge to tours with length=10992. When W_2=0.4, only 55% converge to tours with length = 10992 and 45% converge to tours with length longer than 10992.

The three terms, i.e., the membrane term K, the energy term E, and the self-feedback term F, are comparable in magnitude in the first condition. Figure 2 shows the Energy term. The chaos helps the searching in a little different way. The transient chaos is quite distinct and contributes to the searching processes.

We have carried out over 2000 simulations for the 48-city

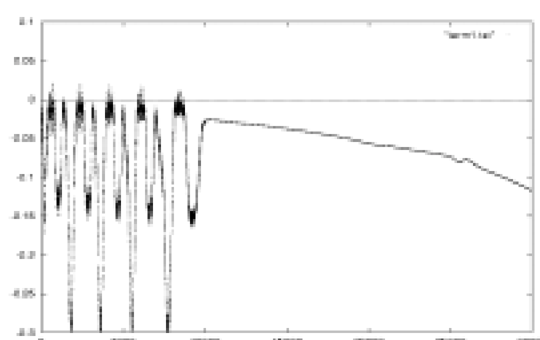
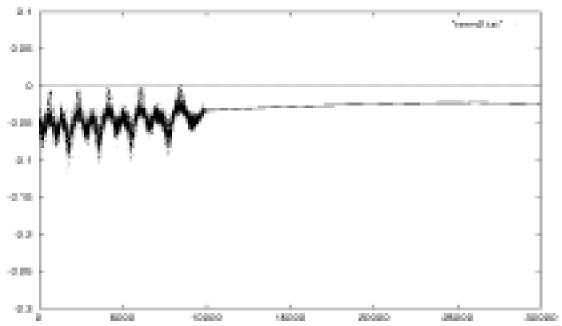
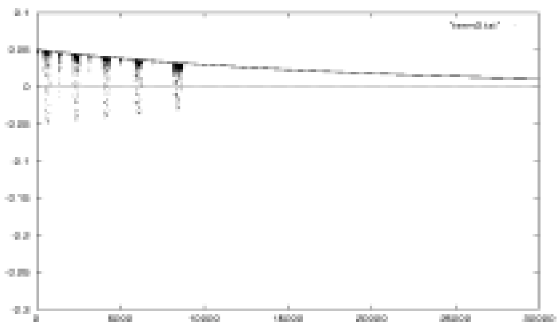


Figure 2 (a)



(b)



(c)

Figure 2 Energy term K, E, F against t

TSP with random initial conditions and with different choices of W_2 in the range of 0.08 and 887. The network never found the shortest tour with length 10628.

The simulations also show that if W_2 is too small, for example, W_2<0.3, the solutions are tours longer than 10992 (local minimum). In these cases, the tour-length term is too weak compared to the other terms.

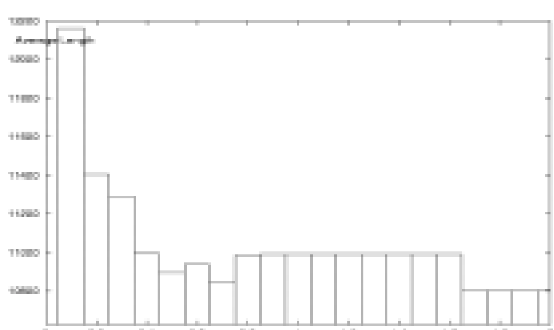


Figure 3. Average tour lengths of 100 runs against W_2

When W_2 is greater than 2.0, invalid tours begin to appear, when W_2 is greater than 2.3, e.g., W_2=887, which corresponds to the case W_2=1/3 [5] but no distance normalization, the system can not find any feasible tour. In these cases, the tour-length term is too large and dominates the input of the system. In comparison, the constraint term and the self-feedback term are ineffective, thereby resulting in invalid tours and absence of chaos.

Figure 3 demonstrates the average tour lengths of 100 runs with various parameter W_2, The initial inputs y_{ij} for each neuron are randomly generated between [-1,1]. Table 1 summarizes the shortest lengths with various W_2.

Table1. Shortest tour lengths of 48-city TSP

W_2	0.1	0.2	0.3	0.4	0.5	0.6	0.7-1.6	1.7-2.0
Shortest	12164	11209	10943	10857	10791	10739	10992	10805

3. CONCLUSIONS

We have carried out extensive simulations for solving the TSP problem using the TCNN of Chen and Aihara, Our results show that when the number of cities is small, for example, 4 or 10, the transient chaos helps the network very easily find an global optimum, and the solutions are less sensitive to W_1, W_2. But when the number of cities becomes larger, the TCNN's global searching ability becomes more sensitive to the choice of parameters. According to our analysis and simulations, we suggest that all parameters should be chosen such that all the three terms in the neuronal input, i.e., the membrane term K, the energy term (constraint and tour length) E, and the self-feedback term F, are comparable.

Our simulations show that although the TCNN performs much better than the original Hopfield-Tank approach, one still needs to choose network parameters carefully. Invalid or poor solutions will result from parameters selected arbitrarily. Even after our guideline above is followed, fine-tuning of the parameters is necessary for satisfactory performance. Our future work will include applications of an alternative approach to chaotic simulated annealing proposed recently by Wang and Smith [14], which requires a smaller number of parameters to be selected. We are also in the process of developing a neural network approach to effectively solving optimization problems *without the need to select parameters.*

REFERENCES

[1] J. J. Hopfield and D. W. Tank, “Neural computation of decisions in optimization problems,” *Biological Cybernetics*, vol. 52, pp. 141-152, 1985.

[2] G.V. Wilson and G. S. Pawley, “On the stability of the traveling salesman problem algorithm of Hopfield and Tank,” *Biological Cybernetics*, vol. 58, pp. 63, 1988.

[3] E. Aarts and J. Korst, *Simulated Annealing and Boltzmann Machines*, Chichester: John Wiley, 1989.

[4] K. Aihara, T. Takabe and M. Toyoda, “Chaotic neural networks,” *Physics Letters A*, vol. 144, no. 6, 7, pp. 333-340, 1990.

[5] L. Chen and K. Aihara, “Chaotic simulated annealing by a neural network model with transient chaos,” *Neural Networks*, vol. 8, no. 6, pp. 915-930, 1995.

[6] G. Reinelt, TSPLIB--A traveling salesman library. *ORSA Journal on Computing*, vol.3, no. 4, pp. 376-384, 1991.

[7] G.W. Davis, “Sensitivity analysis in neural net solutions,” IEEE Transactions on Systems, Man, And Cybernetics, vol. 19, no. 5, pp. 1078-1082, 1989.

[8] L. Wang and J. Ross, “Osillations and chaos in neural networks: an exactly solvable model.” *Proc. National Academy of Sciences* (USA) vol. 87, pp. 9467-9471, 1990.

[9] L. Wang and D.L. Alkon, (eds.), *Artificial neural networks: Oscillations, chaos, and sequence processing*. (IEEE Computer Society Press), Los Alamitos, California, 1993.

[10] L. Wang, “Oscillatory and chaotic dynamics in neural networks under varying operating conditions,” *IEEE Transactions on Neural Networks*, vol.7, no. 6, pp. 1382-1388, 1996.

[11] T. Kwok, K. Smith and L. Wang, “Solving combinatorial optimization problems by chaotic neural networks,” C. dagli et al. (eds.) *Intelligent Engineering Systems through Artificial Neural Networks*, vol. 8, pp. 317-322, 1998.

[12] T. Kwok, K. Smith and L. Wang, “Incorporating chaos into the Hopfield neural network for combinatorial optimization,” in *Proc. 1998 World Multiconference on Systemics, Cybernetics and Informatics*, N. Callaos, O. Omolayole, and L. Wang, (eds.) vol.1, pp. 646-651, 1998.

[13] L. Wang and K. Smith, “Chaos in the discretized analog Hopfield neural network and potential applications to optimization,” *Proc. International Joint Conference on Neural Networks*, vol.2, pp. 1679-1684, 1998.

[14] L. Wang and K. Smith, “On chaotic simulated annealing,” *IEEE Transactions on Neural Networks.*, vol. 9, no. 4, pp. 716-718, 1998.

[15] M. Hasegawa, T.Ikeguchi and T. Matozaki, “Solving combinatorial optimization problems by nonlinear neural dynamics,” *Proc. International Conference on Neural Networks,* pp. 3140-3145,1995.