We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

6

Solving TSP by Transiently Chaotic Neural Networks

Shyan-Shiou Chen[1] and Chih-Wen Shih[2]
[1]*Department of Mathematics, National Taiwan Normal University, Taipei,*
[2]*Department of Applied Mathematics, National Chiao Tung University, Hsinchu, Taiwan, R.O.C.*

1. Introduction

Inspired by the information processing of human neural systems, the artificial neural networks (ANNs) have been developed and applied to solve problems in various disciplines with varying degrees of success. For example, ANNs have been applied to memory storage, pattern recognition, categorization, error correction, decision making, and machine learning in object oriented machine. Various computational schemes and algorithms have been devised for solving the travelling salesman problem which is a difficult NP-hard combinatorial optimization problem. The use of ANN as a computational machine to solve combinatorial optimization problems, including TSP, dates back to 1985 by Hopfield and Tank (1985). Although the achievement of such an application broadens the capacity of ANNs, there remain several insufficiencies to be improved for such a computational task, cf. (Smith, 1999). They include that the computations can easily get trapped at local minimum of the objective function, feasibility of computational outputs, and suitable choice of parameters. Improvements of feasibility and solution quality for the scheme have been reported subsequently. Among them, there is a success in adding the chaotic ingredient into the network to enhance the global searching ability. Chaotic behavior is an inside essence of stochastic processes in nonlinear deterministic system. Recently, chaotic neural networks have been paid much attention to, and contribute toward solving TSP. Chaotic phenomena arise from nonlinear system, and the discrete-time analogue of Hopfield's model can admit such a dynamics. Notably, the discrete-time neural network models can also be implemented into analogue circuits, cf. (Hänggi et al., 1999 ; Harrer & Nossek, 1992).

The chapter aims at introducing recent progress in discrete-time neural network models, in particular, the transiently chaotic neural network (TCNN) and the advantage of adopting piecewise linear activation function. We shall demonstrate the use of TCNN in solving the TSP and compare the results with other neural networks. The chaotic ingredients improve the shortcoming of the previous ODE models in which the outputs strongly depend on the initial conditions and are easily trapped at the local minimum of objective function. There are transiently chaotic phase and convergent phase for the TCNN. The parameters for convergent phase are confirmed by the nonautonomous discrete-time LaSalle's invariant principle, whereas the ones for chaotic phase are derived by applying the Marotto's theorem. The Marotto's theorem which generalizes the Li-York's theorem on chaos from one-dimension to multi-dimension has found its best application in the discrete-time neural network model considered herein.

Source: Travelling Salesman Problem, Book edited by: Federico Greco, ISBN 978-953-7619-10-7, pp. 202, September 2008, I-Tech, Vienna, Austria

In Section 2, we will introduce the setting of solving the TSP by the neural networks, including the description of objective functions to be minimized at optimal routes, and the original work by Hopfield and Tank. In Section 3, we review the LaSalle's invariant principle, the Marotto's theorem, and introduce the discrete-time analogue of Hopfield's network. The recent progress in the transiently chaotic neural network is summarized in Section 4. We arrange some numerical simulations in applying the TCNN with piecewise linear activation function in Section 5. Finally, the chapter is concluded with some discussions.

2. Solving TSP via Hopfield neural network

Suppose there are N cities indexed by i= 1, 2,…, N and the distance between city i and city k is d_{ik}. The optimal solution to the TSP consists of an ordered list of the N cities. The list expresses the order of the cities visited and indicates the path with shortest total tour length. Let us describe how to map the TSP into the computational networks. For each city, its final location in the ordered list is to be specified by the asymptotic output states of a set of N neurons. For example, for a 10-city problem, if city i is the seventh city visited in an optimal solution, then it is represented by the corresponding outputs of 10 neurons:

$$0\,0\,0\,0\,0\,0\,1\,0\,0\,0.$$

Accordingly, N^2 neurons will be needed in the computational network for a N-city TSP. We thus arrange the outputs of these neurons into a $N\times N$ matrix. In such a representation, an ideal output matrix with only one entry equal to one in each row and in each column, and other entries all zero, will then correspond to a feasible solution of the TSP.
Thereafter, the TSP with N cities can be formulated as the following optimization problem:

$$\text{Minimize} \qquad E(\mathbf{y}) = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{N} d_{ij}\, y_{ik}\,(y_{j(k-1)} + y_{j(k+1)}), \tag{1}$$

where matrix $\mathbf{y}=[y_{ij}]$ is constrained by

$$\sum_{i=1}^{N} y_{ij} = 1 \text{ and } \sum_{j=1}^{N} y_{ij} = 1, \tag{2}$$

for all i,j =1,…,N, and $y_{i0}=y_{iN}$ and $y_{i1}=y_{i(N+1)}$. The variables $y_{ij}\in[0,1]$, i,j =1,…,N, can also be regarded as the probability for the i th city to be visited the j th time. If every y_{ij} is either 0 or 1, then the constraint Eq. (2) means that every city must be visited only once. Under such a circumstances, the optimal solution of the objective function $E(\mathbf{y})$ equals the shortest distance of the traveling route. Notably, any shift of an optimal solution also yields an optimal solution (with the same shortest tour length). Thus, the optimal solution is not unique.
The main idea of using neural networks to solve TSP is to search the global minimum of the objective function which involves the data of TSP, through evolutions of the states of the network. Hopfield & Tank (1985) considered the following objective function

$$E(\mathbf{y}) = \frac{\gamma_1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k\neq j}^{N} y_{ij}y_{ik} + \frac{\gamma_2}{2}\sum_{j=1}^{N}\sum_{i=1}^{N}\sum_{k\neq i}^{N} y_{ij}y_{kj} + \frac{\gamma_3}{2}(\sum_{i=1}^{N}\sum_{j=1}^{N} y_{ij} - N)^2 \tag{3}$$

$$+\frac{\gamma_4}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{N} d_{ij}y_{ik}(y_{j(k-1)} + y_{j(k+1)}). \tag{4}$$

Note that the minimum, i.e., zero, of (3) is attained at a feasible solution. They aimed at using the Hopfield network to locate the global minimum of this objective function. The Hopfield network is a continuous-time ODE system which consists of a fully interconnected system of computational elements or neurons arranged in, say, lattice L:

$$C_{\mathbf{i}}\frac{dx_{\mathbf{i}}}{dt} = -\frac{x_{\mathbf{i}}}{R_{\mathbf{i}}} + \sum_{\mathbf{j}\in L} w_{\mathbf{ij}}y_{\mathbf{j}} + I_{\mathbf{i}}, \mathbf{i}\in L, \tag{5}$$

$$y_{\mathbf{i}} = g_{\mathbf{i}}(x_{\mathbf{i}}). \tag{6}$$

The synapse (or weight) between two neurons is defined by w_{ij}, which may be positive or negative depending on whether the action of neurons is in an excitatory or inhibitory manner; $x_{\mathbf{i}}$ is the internal state of neuron **i,** and $y_{\mathbf{i}}$ with $0 \le y_{\mathbf{i}} \le 1$ is the external (output) state of neuron **i.** The parameter $C_{\mathbf{i}}$ (resp. $R_{\mathbf{i}}$) is the input capacitance of the cell membrane (resp. the transmembrane resistance) of neuron **i**. The activation function $g_{\mathbf{i}}$ is a monotonically increasing function and thus has an inverse. Typical $g_{\mathbf{i}}$ is given by

$$g_{\mathbf{i}}(\xi) = \frac{1}{2}(1+\tanh(\xi/\varepsilon)),$$

where ε is a parameter controlling the slope of the activation function. The gradient descent dynamics of the neural network provides a decreasing property of the objective function for the TSP. For convenience of expression and derivation, we consider (5) on the one-dimensional array {1,2,…,n}. There exists a Lyapunov function (energy function for the network)

$$V = -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}y_iy_j - \sum_{i=1}^{n} I_iy_i + \sum_{i=1}^{n}\frac{1}{R_i}\int_0^{y_i} g_i^{-1}(\xi)d\xi. \tag{7}$$

The time derivative of V along a solution $\mathbf{x}(t)$ is computed as

$$\frac{dV}{dt} = -\sum_{i=1}^{n}\frac{dy_i}{dt}(\sum_{j=1}^{n} w_{ij}y_j - \frac{x_i}{R_i} + I_i) = -\sum_{i=1}^{n}\frac{dy_i}{dt}C_i\frac{dx_i}{dt}$$
$$= -C_i\sum_{i=1}^{n}[\frac{dg_i(x_i)}{dx_i}](\frac{dx_i}{dt})^2.$$

Due to the increasing property of the activation function g_i, we obtain

$$\frac{dV}{dt} \le 0, \text{ and } \frac{dV}{dt} = 0 \text{ if } \frac{dx_i}{dt} = 0. \tag{8}$$

3. Discrete-time dynamical systems

Biological neurons are much more complicated than the simple threshold elements in ANNs. Chaotic behaviors have actually been observed experimentally in biological neurons, as pointed out in Aihara et al. (1990) and the references therein.

Discrete-time dynamical systems have attracted much attention in recent years, thanks to its capacity of applications and underlying sophisticated mathematical theory. Indeed, not only that discrete-time counterparts of classical theorems for continuous-time systems have been developed successfully, but also the chaotic behaviors for discrete-time systems can be characterized lucidly.

Due to the shortcomings that solutions get trapped at local minimum of objective function, and dependence of performance upon choosing initial conditions in continuous-time systems, researchers have attempted to introduce the chaotic ingredient into the networks. The stage was thus set for the development of discrete-time neural networks, cf. (Aihara et al., 1990; Chen & Aihara, 1995; Nozawa, 1992; Yamada & Aihara, 1997). In particular, Nozawa (1992) showed that the Euler approximation of the continuous-time Hopfield neural network with a negative self-feedback connection possesses chaotic dynamics, and has a much better searching ability in solving the TSP than the original continuous-time Hopfield neural network.

Notably, although it has been reported in (Bersini,1998; Bersini & Senser, 2002) that there exist chaotic behaviours in continuous-time Hopfield-type neural networks, it is still unknown whether the same concept or technique can be applied to the TSP problem.

We shall introduce the discrete-time Hopfield neural network in Subsection 3.1; then review the LaSalle's invariant principle for convergent dynamics and the Marotto's theorem for chaos, for discrete-time dynamical systems in Subsections 3.2, 3.3, respectively.

3.1 Discrete-time Hopfield neural networks

Discrete-time neural network model of Hopfield type can be described by the following equations: for $i = 1,\ldots,n$,

$$x_i(t+1) = \mu x_i(t) + \sum_{j=1}^{n} w_{ij} y_j(t) + I_i. \tag{9}$$

Here, x_i is the internal state of neuron i; y_i is the output of neuron i; μ is the damping factor; w_{ii} is the self-feedback connection weight; w_{ij} is the connection weight from neuron j to neuron i; I_i is the input bias. The parameter μ (resp. ω_{ij}, I_i) in Eq. (9) can be compared to $1-\frac{\Delta t}{C_i R_i}$ (resp. $\frac{\omega_{ij}\Delta t}{C_i}$, $\frac{I_i \Delta t}{C_i}$) in terms of the parameters in Eq. (5), where Δt is the discretization time step. System (9) is the Euler approximation of Eq. (5). There also exists a Lyapunov function for the discrete-time system (9):

$$V(\mathbf{y}) = -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} y_i y_j - \sum_{i=1}^{n} I_i y_i - (\mu - 1)\sum_{i=1}^{n} \int_0^{y_i} g_i^{-1}(\xi) d\xi, \tag{10}$$

where again $y_i = g_i(x_i)$. It has been shown in (Chen & Aihara, 1997; Chen & Shih, 2002) that, under some conditions, the energy function (10) is decreasing along the solution of the system:

$$V(\mathbf{y}(t+1)) - V(\mathbf{y}(t)) \le 0, \text{ for all } t \in \boldsymbol{N}.$$

Notably, the Lyapunov function (10) for the discrete-time network and the one (7) for the continuous-time network are quite similar. The existence of Lyapunov function basically guarantees the convergence of evolutions for the system to certain steady states, by the LaSalle's invariant principle. The transiently chaotic neural network is developed from this discrete-time network with transient chaos imbedded in. Before introducing the theory for the TCNN, let us review the LaSalle's invariant principle and the Marotto's theorem.

3.2 LaSalle's invariant principle

Long-time asymptotic behaviors of solutions for a dynamical system, such as neural network, are always important concerns. In 1960, LaSalle discovered the relation between Lyapunov function and Birkhoff limit set. Extended from the Lyapunov direct method, a uniform concept was developed to describe the asymptotic behaviors in terms of limit set. By the invariant property of limit set, a basic theory for the stability of motion of dynamical systems was derived. In this section, we review the invariant principle for both autonomous and non-autonomous discrete-time dynamical systems, cf. (LaSalle, 1976). First, we consider an autonomous difference equation

$$\mathbf{x}(t+1) = \mathbf{F}(\mathbf{x}(t)), \mathbf{x} \in R^n, t \in \boldsymbol{N}, \tag{11}$$

where $\mathbf{F}: R^n \to R^n$ is continuous. We assume that $\mathbf{x}^*$ is a fixed point (i.e. $\mathbf{F}(\mathbf{x}^*) = \mathbf{x}^*$). Suppose there exists a continuous, positive definite, and radially unbounded function $V: \overline{G} \to R, G \subset R^n$ with

$$\Delta V(\mathbf{x}) \le 0, \forall \mathbf{x} \in G$$

where $\Delta V(\mathbf{x}) = V(\mathbf{F}(\mathbf{x})) - V(\mathbf{x})$, then every solution to Eq. (11) converges to the largest invariant set *M* contained in $\{\mathbf{x} \in G \mid \Delta V(\mathbf{x}) = 0\}$. If the set *M* only contains the equilibrium **x*** , then **x*** is asymptotically stable. The function *V* satisfying $\Delta V(\mathbf{x}) \le 0$ for all $\mathbf{x} \in G$ is said to be a *Lyapunov function* on set *G*.

Now we consider a discrete-time non-autonomous system. Let $\boldsymbol{N}$ be the set of positive integers. For a given continuous function $\mathbf{F}: \boldsymbol{N} \times R^n \longrightarrow R^n$, we consider the non-autonomous dynamical system

$$\mathbf{x}(t+1) = \mathbf{F}(t, \mathbf{x}(t)). \tag{12}$$

A sequence of points $\{\mathbf{x}(t) \mid t = 1, 2, \cdots\}$ in R^n is a solution of (12) if $\mathbf{x}(t+1) = \mathbf{F}(t, \mathbf{x}(t))$, for all $t \in \boldsymbol{N}$. Let $O_{\mathbf{x}} = \{\mathbf{x}(t) \mid t \in \boldsymbol{N}, \mathbf{x}(1) = \mathbf{x}\}$ be the orbit of **x**. We say that **p** is a ω-limit point of $O_{\mathbf{x}}$ if there exists a sequence of positive integers $\{t_k\}$ with $t_k \to \infty$ as $k \to \infty$, such

that $\mathbf{p} = \lim_{k\to\infty} \mathbf{x}(t_k)$. Denote by $\omega(\mathbf{x})$ the set of all ω-limit points of $O_{\mathbf{x}}$. Let N_i represent the set of all positive integers larger than n_i, for some positive integer n_i. Let $\Omega \subseteq R^n$ and $\overline{\Omega}$ be its closure. For a function $V : \boldsymbol{N}_0 \times \Omega \longrightarrow R$, we define $\dot{V}(t,\mathbf{x}) = V(t+1,\mathbf{F}(t,\mathbf{x})) - V(t,\mathbf{x})$ so that if $\{\mathbf{x}(t)\}$ is a solution of Eq. (12), then $\dot{V}(t,\mathbf{x}(t)) = V(t+1,\mathbf{x}(t+1)) - V(t,\mathbf{x}(t))$. V is said to be a *Lyapunov function* for (12) if

i. each $V(t,\cdot)$ is continuous, and

ii. for each $\mathbf{p} \in \overline{\Omega}$, there exists a neighborhood U of $\mathbf{p}$ such that $V(t,\mathbf{x})$ is bounded below for $\mathbf{x} \in U \cap \Omega$ and $t \in \boldsymbol{N}_1$, $n_1 \geq n_0$, and

iii. there exists a non-degenerate continuous function $Q_0 : \overline{\Omega} \to R$ such that $\dot{V}(t,\mathbf{x}) \leq -Q_0(\mathbf{x}) \leq 0$ for all $\mathbf{x} \in \Omega$ and for all $t \in \boldsymbol{N}_2$, $n_2 \geq n_1$,

or

iii.' there exist a non-degenerate continuous function $Q_0 : \overline{\Omega} \to R$ and an equi-continuous family of functions $Q(t,\cdot) : \overline{\Omega} \to R$ such that $\lim_{t\to\infty} |Q(t,\mathbf{x}) - Q_0(\mathbf{x})| = 0$ for all $\mathbf{x} \in \Omega$ and $\dot{V}(t,\mathbf{x}) \leq -Q(t,\mathbf{x}) \leq 0$ for all $(t,\mathbf{x}) \in N_2 \times \Omega$, $n_2 \geq n_1$.

If there exists such a Lyapunov function *V*, then the LaSalle's invariant principle states that the ω-limit set of any point **x** lies in S_0, i.e. $\omega(\mathbf{x}) \subset S_0$, where

$$S_0 = \{\mathbf{x} \in \overline{\Omega} : Q_0(\mathbf{x}) = 0\}. \tag{13}$$

3.3 Marotto's theorem on chaos

Originally, a chaotic phenomenon was numerically found in the research of Lorenz on weather prediction in 1963. Later, the mathematical definition of chaos was initiated by Li & Yorke (1975) for one-dimensional continuous maps. A criterion of existence of chaos has been termed as "period three implies chaos" therein. More precisely, let $f : I \to I$ be a continuous map of the compact interval *I* of the real line *R* into itself; if *f* has a periodic point of period three, then *f* exhibits chaotic behavior. Three years later, the above result was generalized by Marotto (1978). He proposed the definition of "snapback repeller" and proved that "snapback repellers imply chaos" for multi-dimensional maps. The definition of snapback repeller was further clarified in (Marotto, 2005).

The theorem has provided the best analytic argument of chaos for multi-dimensional maps. The detailed description of chaos in the sense of Marotto is as follows. Let us define a system as $\mathbf{x}_{k+1} = F(\mathbf{x}_k)$ where $\mathbf{x}_k \in R^n$, and *F* is C^1 or piecewise C^1 with non-smooth points at suitable locations. A fixed point $\overline{\mathbf{x}}$ is said to be a *snapback repeller* (see Fig. 1) of *F* if all eigenvalues of $DF(\overline{\mathbf{x}})$ exceeding one in magnitude, and there exists a point $\mathbf{x}_0 \neq \overline{\mathbf{x}}$ in a repelling neighborhood of $\overline{\mathbf{x}}$, such that $F^m(\mathbf{x}_0) = \overline{\mathbf{x}}$ for some $m \in \mathbf{N}$, and

$\det(DF^j(\mathbf{x}_0)) \neq 0$, for all $1 \leq j \leq m$. If F has a snapback repeller, then the dynamical system defined by F is chaotic in the following sense: (i) There exists a positive integer m_0 such that for each integer $p \geq m_0$, F has p-periodic points. (ii) There exists a scrambled set, that is, an uncountable set L containing no periodic points such that the following pertains: (a) $F(L) \subset L$; (b) for every $\mathbf{y} \in L$ and any periodic point $\mathbf{x}$ of F,

$$\limsup_{k\to\infty} \| F^k(\mathbf{y}) - F^k(\mathbf{x}) \| > 0;$$

(c) for every $\mathbf{x}, \mathbf{y} \in L$ with $\mathbf{x} \neq \mathbf{y}$,

$$\limsup_{k\to\infty} \| F^k(\mathbf{y}) - F^k(\mathbf{x}) \| > 0;$$

(iii) There exists an uncountable subset L_0 of L such that for every $\mathbf{x}, \mathbf{y} \in L_0$,

$$\liminf_{k\to\infty} \| F^k(\mathbf{y}) - F^k(\mathbf{x}) \| = 0.$$

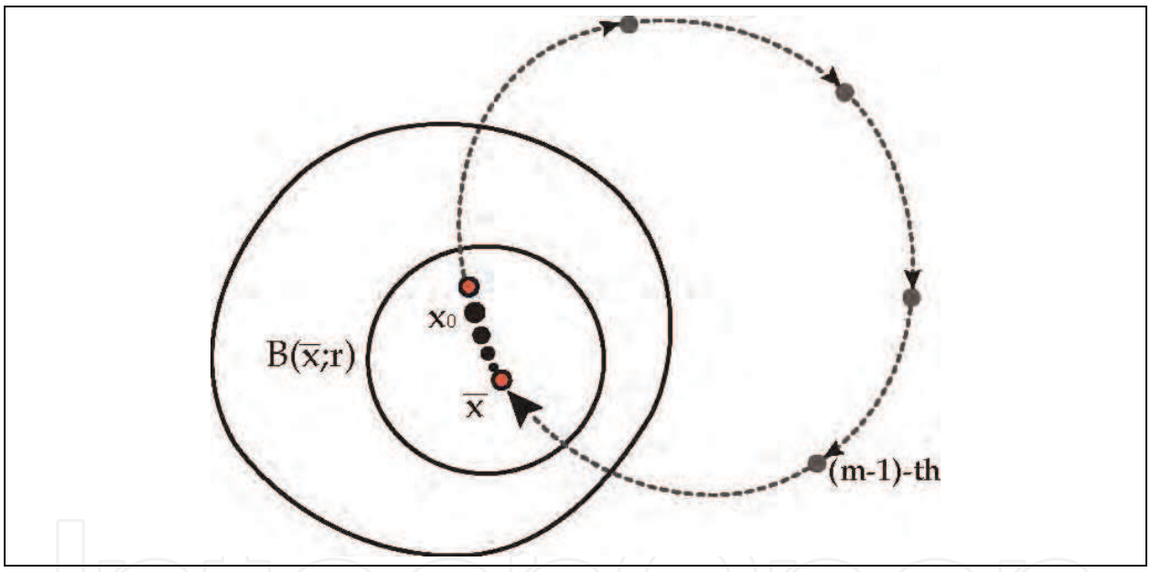

Fig. 1. Diagram of a snapback repeller. The point $\overline{X}$ is a snapback repeller. The point X_0 is a snapback point such that $F^m(X_0) = \overline{X}$ for some integer m. Note that the value of F at the (m-1)-th point is the snapback repeller $\overline{X}$.

Notably, (ii) (b) describes that any point in the scrambled set L does not converge to any periodic point of F under the iteration of F. It bears a sense that there only exist unstable periodic points in the system. (ii)(c) shows that there only exist unstable points in the scrambled set L. In other words, points in the scrambled set do not attract each other. (iii) describes that distances between the iterations of any pair of points in an uncountable subset of L approach zero. Although it seems that there exists no rule for the dynamical behavior, the behavior is controlled by the underlying deterministic system. It is not similar to the concept of randomness of a stochastic system. The chaotic behavior is very random but ordered.

Let us illustrate the existence of period-three points and a snap-back repeller in the sense of Li-Yorke and Marotto respectively, for the logistic map, $f_\mu(x) = \mu x(1-x), x \in [0,1]$, as an example. The period-three points and a snap-back repeller are presented in Fig. (2). There exists chaos in the sense of Li-Yorke (resp. Marotto) for the logistic map f_μ with $\mu = 4$.

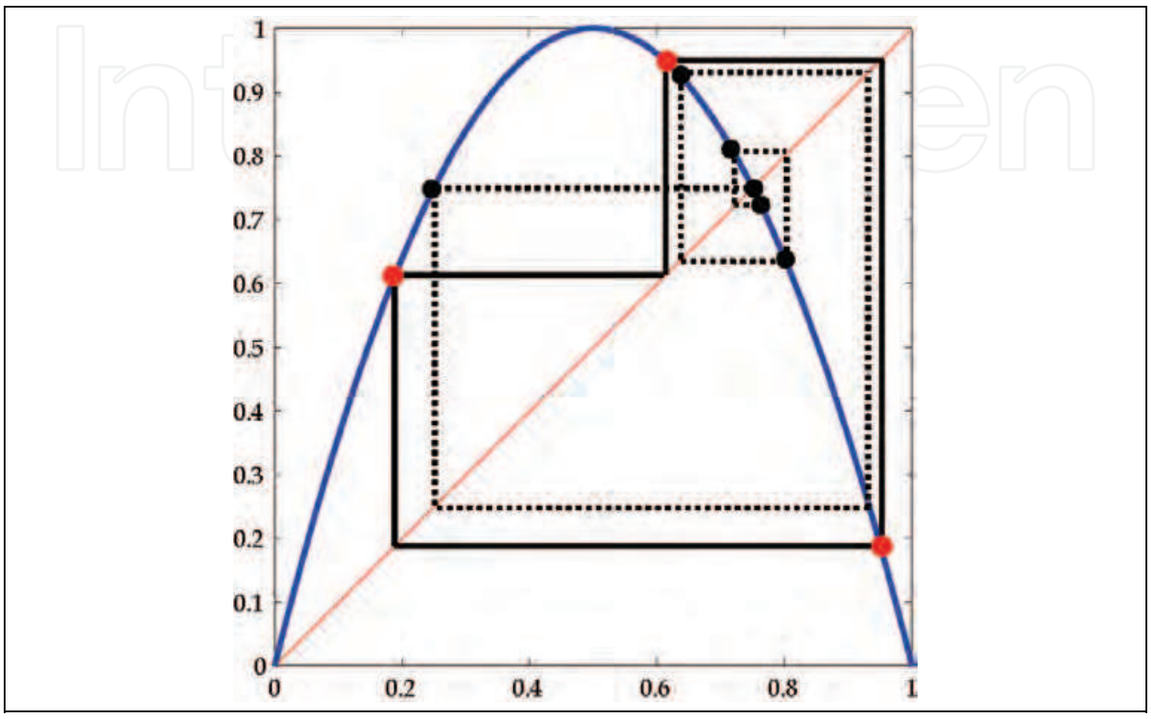

Fig. 2. The blue line is the graph of logistic map with $\mu = 4$. Black line illustrates the period 3 trajectory. The dotted line depicts a homoclinic orbit with snap-back points. This logistic map possesses Li-Yorke's and Marotto's chaos.

4. Transiently chaotic neural networks

The transiently chaotic neural network (TCNN) is equipped with a chaotic phase which prevails in the first stage of computation to enhance global searching and reduce the effect of variations from choosing initial values. This procedure can be realized by a suitable choice of parameters which typically starts from sufficiently large negative self-feedback connection weights. The process is then cooled down, as the self-feedback connection weights increases, while maintaining decreasing property of the energy (objective) function, and finally settles at a state with lower value of objective function. The characteristic of dynamical phenomena from chaotic phase to convergent phase is called "chaotic simulated annealing".

The TCNN, inherited from the Hopfield type network, was first proposed by Chen & Aihara (1995, 1997, 1999). Later, Chen & Shih (2002) performed a systematic analysis on the chaotic behaviors of the TCNN. The existence of chaos is proved by a geometrical formulation combined with the use of Marotto's theorem. The analysis has provided the ranges of

parameters for the chaotic phase and convergent phase. Recently, Chen and Shih (2007) further extended the TCNN to the setting with piecewise linear activation function, which not only improves the performance of computation, but also admits more succinct and crystal mathematical description on the chaotic phase than the TCNN with the logistic activation function. Such a setting fits into the revised version of theorem in (Marotto, 2005) pertinently.

Let us describe the model equation for the TCNN.

$$x_i(t+1) = \mu x_i(t) - \omega_{ii}(t)[y_i(t) - a_{0i}] + \alpha[\sum_{j=1, j\neq i}^{n} \omega_{ij} y_j(t) + v_i], \tag{14}$$

$$|\omega_{ii}(t+1)| = |(1-\beta)\omega_{ii}(t)|, \tag{15}$$

for i=1,…,n, $t \in \mathbf{N}$, (positive integers), where x_i is the internal state of neuron i; y_i is the output of neuron i, which corresponds to x_i through an activation function; μ with $0 < \mu < 1$ is the damping factor of nerve membrane; ω_{ii} is the self-feedback connection weight; a_{0i} is the self-recurrent bias of neuron i; ω_{ij} is the connection weight from neuron j to neuron i; v_i is the input bias of neuron i; β with $0 < \beta < 1$, is the damping factor for ω_{ii}; α is a positive scaling parameter. Equation (15) represents an exponential cooling schedule in the annealing procedure. The activation function adopted in (Chen & Aihara, 1995; 1997; 1999) is the logistic function given by

$$y_i(t) = 1/[1 + \exp(-x_i(t)/\varepsilon)],$$

which is depicted in Fig. 3 (b).

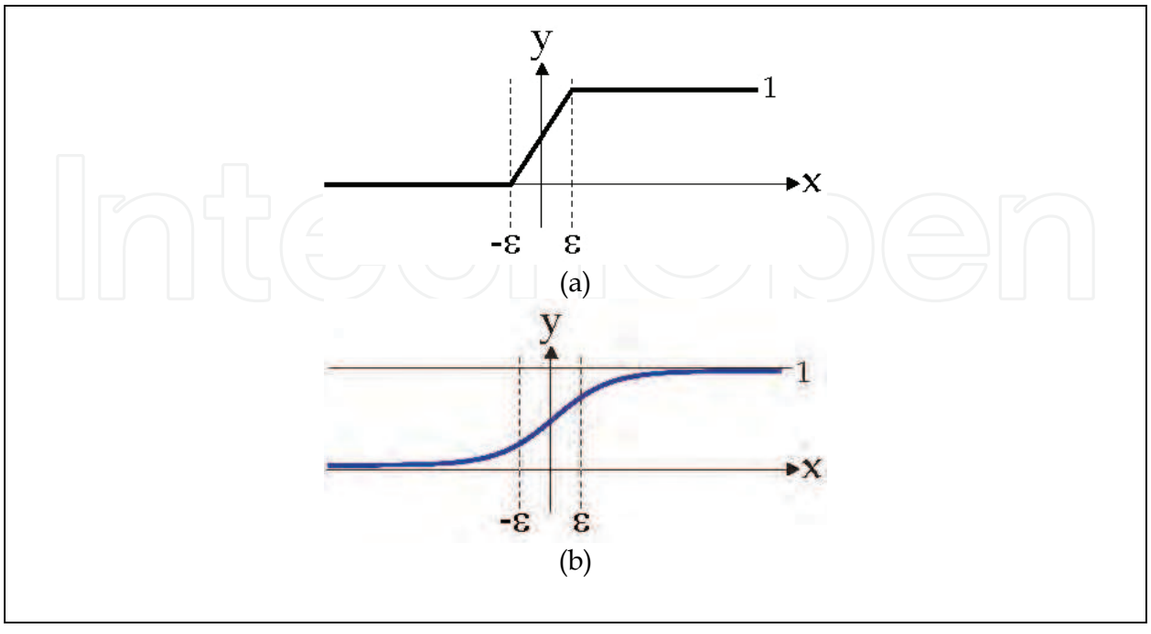

Fig. 3. The graphs of (a) the piecewise linear and (b) the logistic activation function.

One may also consider more general annealing process:

$$x_i(t+1) = \mu x_i(t) + (1-\beta)^{q(t)}\omega[y_i(t) - a_{0i}] + \alpha[\sum_{j=1, j\neq i}^{n} \omega_{ij} y_j(t) + v_i], \tag{16}$$

where i=1,…,n, $0 < \beta < 1$; $q(t)$ satisfies the condition that there exists an $n_1 \in \boldsymbol{N}$ such that $q(t) - t \geq 0$ for all $t > n_1$. The standard annealing process simply takes $q(t) = t$.

The disadvantage of using the logistic activation function is that the output values for some neurons may be neither close to one nor to zero, as demonstrated in Fig. (4). Although it is possible to avoid such a situation by choosing high gain of the logistic activation function, i.e., small ε, taking the piecewise linear output function (Fig. 3 (a)) leads to much better performance.

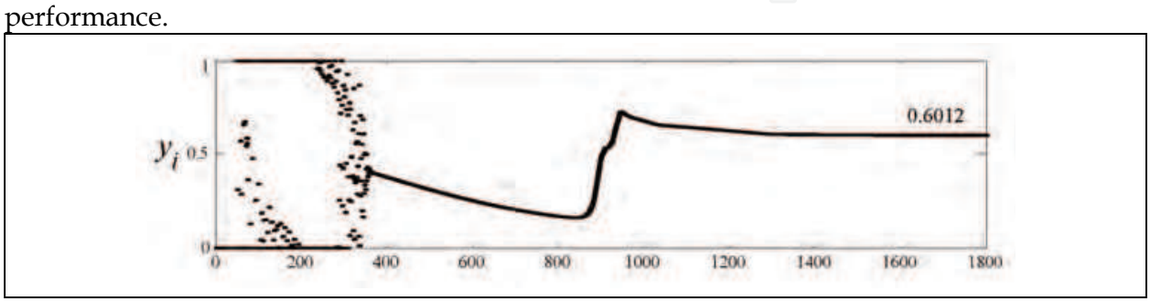

Fig. 4. An example that the TCNN with the logistic activation function has an infeasible solution, i.e., there exists an output entry y_i which approaches 0.6012, neither close to 1 nor to 0, after 1400 iterations.

4.1 Piecewise linear activation function

Chen & Shih (2007) proposed a transiently chaotic neural network (TCNN) with piecewise linear activation function, instead of the logistic one, as follows:

$$x_i(t+1) = \mu x_i(t) + (1-\beta)^{q(t)}\omega[y_i(t) - a_{0i}] + \sum_{j=1}^{n} w_{ij} y_j(t) + v_i, \tag{17}$$

where i=1,…,n, and x_i and y_i satisfy the following relation

$$y_i(t) = g_\varepsilon(x_i(t)) := [2 + |\frac{x_i(t)}{\varepsilon} + 1| - |\frac{x_i(t)}{\varepsilon} - 1|]/4, \varepsilon > 0, \tag{18}$$

That is, we consider the following time-dependent map on R^n:

$$F_i(t, \mathbf{x}) = \mu x_i + (1-\beta)^{q(t)}\omega[g_\varepsilon(x_i) - a_{0i}] + \sum_{j=1}^{n} \omega_{ij} g_\varepsilon(x_j) + v_i. \tag{19}$$

Corresponding to this piecewise linear activation function, for a fixed $\varepsilon > 0$, we partition the real line into the left (ℓ), middle (m), right (r) parts; namely,

$$\Omega_\ell := (-\infty, -\varepsilon), \Omega_{\mathrm{m}} := [-\varepsilon, \varepsilon], \Omega_{\mathrm{r}} := (\varepsilon, \infty). \tag{20}$$

Consequently, R^n can be partitioned into the following subsets:

$$\Omega_{q_1\cdots q_n} = \{(x_1,\cdots,x_n) \in R^n \mid x_i \in \Omega_{q_i};\ q_i = ``\mathrm{r}'', ``\mathrm{m}'', ``\ell''; i = 1,\cdots,n\}, \tag{21}$$

as illustrated in Fig. 5 for n=2. We may call $\Omega_{\mathrm{m}\cdots\mathrm{m}}$ the *interior region*, each $\Omega_{q_1\cdots q_n}$, with $q_i = ``\ell'', ``\mathrm{r}''$, for all i, an *saturated region*; each $\Omega_{q_1\cdots q_n}$, with $q_i = ``\ell''$, or $``\mathrm{r}''$, for some *i*, and $q_j = ``\mathrm{m}''$ for some *j*, a *mixed region*.

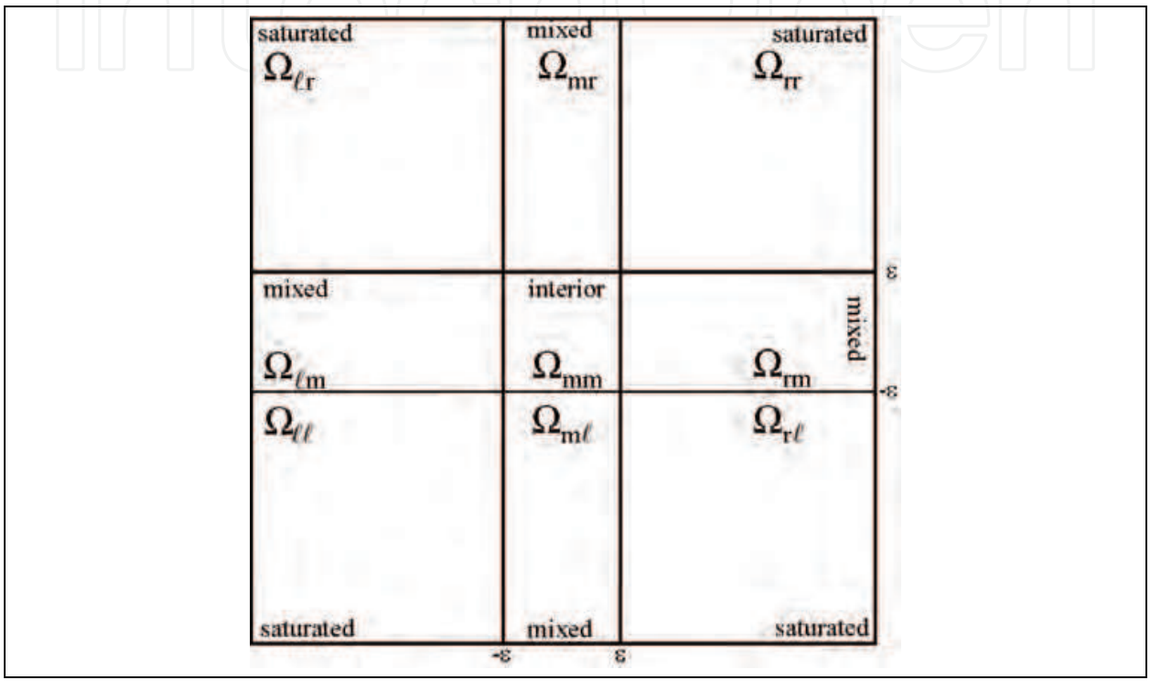

Fig. 5. Illustration of $\Omega_{q_1q_2}$ in R^2, where q_1 and q_2 are $``\ell''$ or $``\mathrm{m}''$ or $``\mathrm{r}''$

Through introducing upper and lower bounds for the map (19), the existence of snap-back repellers in each of the 3^n regions, hence Marotto's chaos, for the system can be established. Let us quote the parameter conditions and the theorem. Consider

$$\text{(A) } \omega > 0,\ 0 < \tfrac{1-\mu}{\omega} < \tfrac{1}{2\varepsilon},\ -\tfrac{1-\mu}{\omega}\varepsilon - \tfrac{h}{\omega} + a_0 > 0,\ \tfrac{1-\mu}{\omega}\varepsilon + \tfrac{h}{\omega} + a_0 < 1,$$

$$\text{(B) } \mu\varepsilon + \omega(1-a_0) - h > \tfrac{1}{\mu}[(\tfrac{1}{2}\omega - \omega a_0 + h)/(1-\mu-\tfrac{\omega}{2\varepsilon}) - \omega + \omega a_0 - h].$$

Theorem 1 (Chen & Shih, 2007). If the parameters $(\mu, \omega, \varepsilon, a_{0i}, h_i)$, satisfy (A) and (B) with a_0=a_{0i}, h=h_i, for every i=1,…,n, then there exist snap-back repellers for the TCNN with activation function (18).

On the other hand, system (17) admits a time-dependent Lyapunov function

$$V(t,\mathbf{x}) = -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\omega_{ij}y_iy_j - \sum_{i=1}^{n}v_iy_i - (\mu-1)\varepsilon\sum_{i=1}^{n}(y_i^2 - y_i) + c^t, \tag{22}$$

where $y_i = g_\varepsilon(x_i)$, i=1,…,n, and $0< c <1$. Note that V is globally Lipschitz, but not C^1. Let $W = [\omega_{ij}]_{n\times n}$. By applying the LaSalle's invariant principle, the following convergent theorem can be derived.

Theorem 2 (Chen & Shih, 2007). If $0 \le \mu \le 1$, $\varepsilon > 0$, $|\frac{1-\beta}{c}|<1$ and the matrix $W + 2\varepsilon(1-\mu)I$ is positive-definite, then there exists $n_2 \in \mathbf{N}, n_2 > n_1$ so that $V(t+1, \mathbf{x}(t+1)) \le V(t, \mathbf{x}(t))$ for $t \ge n_2$ and V is a Lyapunov function for system (17) on $N_2 \times R^n$.

The conditions for chaotic and convergent dynamical phases for the TCNN are all computable. The range of the parameters satisfying these conditions can also be depicted numerically. There are other advantages in adopting the piecewise linear activation function. Note that the feasible and optimal solutions lie in the saturated regions $\Omega_{q_1 \cdots q_n}$, with $q_i = ``\ell'', ``\mathrm{r}''$, for all i. One can further impose conditions so that the fixed points in the interior and mixed regions are unstable. Accordingly, iterations from almost any initial value converge to outputs with component equal to either 0 or 1. Details for these discussions can be found in (Chen & Shih, 2007).

On the other hand, the following objective function is considered in (Chen & Aihara, 1995)

$$E(\mathbf{y}) = \frac{\gamma_1}{2}[\sum_{i=1}^{N}(\sum_{k=1}^{N} y_{ik} - 1)^2 + \sum_{k=1}^{N}(\sum_{i=1}^{N} y_{ik} - 1)^2]$$

$$+ \frac{\gamma_2}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{N} d_{ij} y_{ik}(y_{j(k-1)} + y_{j(k+1)}), \tag{23}$$

where γ_1 and γ_2 are parameters which are selected to reflect the relative importance of the constraints and the cost function of the tour length, respectively.

To apply the TCNN to the TSP, we reformulate the setting of TSP with two-dimensional indices into the one-dimensional form. Restated, by letting $s(i,j)=j+(i-1)N$, where N is the number of cities for the TSP, Eq. (23) becomes

$$E(\mathbf{y}) = -\frac{1}{2}\mathbf{y}\overline{W}\mathbf{y}^T - 2\gamma_1 I_{N^2\times N^2}\mathbf{y} + \gamma_1 N, \tag{24}$$

where, $I_{N^2\times N^2}$ is the identity matrix of size $N^2 \times N^2$, $\mathbf{y} = (y_1, \cdots, y_{s(i,j)}, \cdots, y_{N^2})$ and

$$\overline{W} = -\gamma_1[I_{N\times N} \otimes 1_{N\times N} + 1_{N\times N} \otimes I_{N\times N}] - \gamma_2 D \otimes B; \tag{25}$$

$1_{N\times N}$ is the matrix whose entries are all one, $D = [d_{ij}]^T$ and $B = [b_{ij}]$ with $b_{i,j} = 0$ except that $b_{i,i+1} = b_{i,i-1} = b_{1,N} = b_{N,1} = 1$; the $N^2 \times N^2$ block matrix $A \otimes B$ is defined by the formula $[A \otimes B]_{ij} = [a_{ij}B]$, where $A = [a_{ij}]$ and $B = [b_{ij}]$. The TCNN for the TSP is adjusted to

$$x_i(t+1) = \mu x_i(t) + (1-\beta)^{q(t)}\omega[y_i(t) - a_{0i}] + \sum_{j=1}^{N^2} W_{ij} y_j(t) + 2\gamma_1, \tag{26}$$

where $W=[W_{ij}]:=\overline{W}-\mathrm{diag}[\overline{W}]/2$, i=1,…,N^2. According to previous discussions, there is a Lyapunov function for Eq. (17):

$$V(t,\mathbf{y})=-\frac{1}{2}\sum_{i=1}^{N^2}\sum_{j=1}^{N^2}W_{ij}y_iy_j-\sum_{i=1}^{N^2}2\gamma_1y_i-(\mu-1)\varepsilon\sum_{i=1}^{N^2}(y_i^2-y_i)+c^t. \tag{27}$$

Notice that the Lyapunov function (27) can be compared to a constant shift of the objective function (24) when $|c|<1$, ε is sufficiently small and as t is large.

5. Numerical simulations

Let us describe the method to suitably choose the parameters in the numerical simulation. Due to the deterministic nature for the TCNN (17), parameters are selected such that its dynamical behaviors have some stable properties. Therefore, we take a parameter μ with 0< μ < 1 for boundedness of iterations for the TCNN. Set ω=0, and choose ε, a_{0i}, h, where $h=\max\{h_i: i=1,\cdots,N^2\}$, $h_i=\sum_{k=1}^{N^2}\{|W_{ik}|+2|\gamma_1|\}$, $i=1,\cdots,N^2$, so that the TCNN with these parameters is in convergent phase. In convergent phase, any iteration is going to settle at a fixed point. Next, we let $|\omega|$ increase from 0 to see if parameters $(\mu,\varepsilon,\omega,a_{0i},h)$ enter the chaotic regime which has been justified in (Chen & Shih, 2002; 2007). These computations can be assisted by a computer programming. If the output matrix does not tend to a permutation matrix, one can enlarge slightly the parameter γ_1 in Eq. (25).

In this section, we quote the numerical simulations in (Chen & Shih, 2007) to demonstrate the computation performance of using the TCNN (17) to find the optimal route of the TSP. We consider the five cities $\{1,2,3,4,10\}$ with coordinates in Table 1. These are data from the ten-city TSP problem in (Hopfield & Tank, 1985). The positions of the ten cities and the optimal route are presented in Fig. 6.

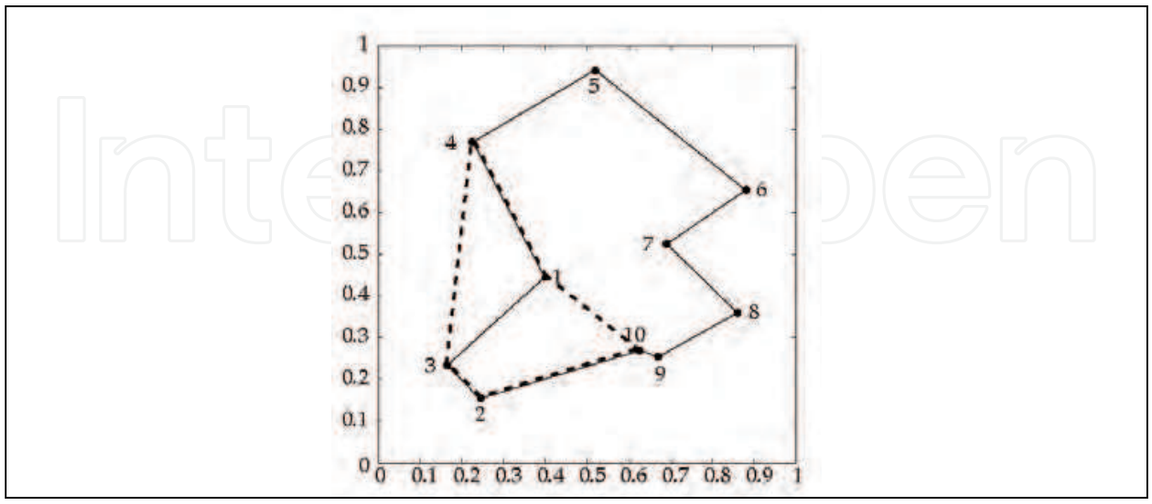

Fig. 6. Illustration of the locations of 10 cities for the Hopfield-Tank original data. The best way to travel for ten (resp. five) cities is in terms of the solid line (resp. dotted line) connection.

In our simulation, the parameters for the TCNN (17) are set as

$$\mu = 0.9\,;\ \beta = 0.005\,;\ \varepsilon = 0.01\,;\ a_{0i} = 0.65\,;$$

$$\omega = -0.08\,;\ \gamma_1 = 0.015\,,\ \gamma_2 = 0.015\,;\ q(t) = t\,.$$

Recall that coefficients γ_1 and γ_2 reflect the relative strength of the constraint and the tour length energy terms (23). An optimal route trajectory is demonstrated in Fig. 7. Our simulation indicates that the order of the best route for the TSP is $4 \mapsto 1 \mapsto 10 \mapsto 2 \mapsto 3$. The other best routes include $1 \mapsto 10 \mapsto 2 \mapsto 3 \mapsto 4$ and $4 \mapsto 3 \mapsto 2 \mapsto 10 \mapsto 1$. Actually, all of them represent the same loop. Three diagrams in Fig. 8 are plotted to show the evolutions of constraint part and tour length part in the objective function.

City No.	x-axis	y-axis
1	.4	.445
2	.245	.155
3	.165	.235
4	.225	.77
10	.625	.27

Table 1. Coordinates of positions for 5 cities.

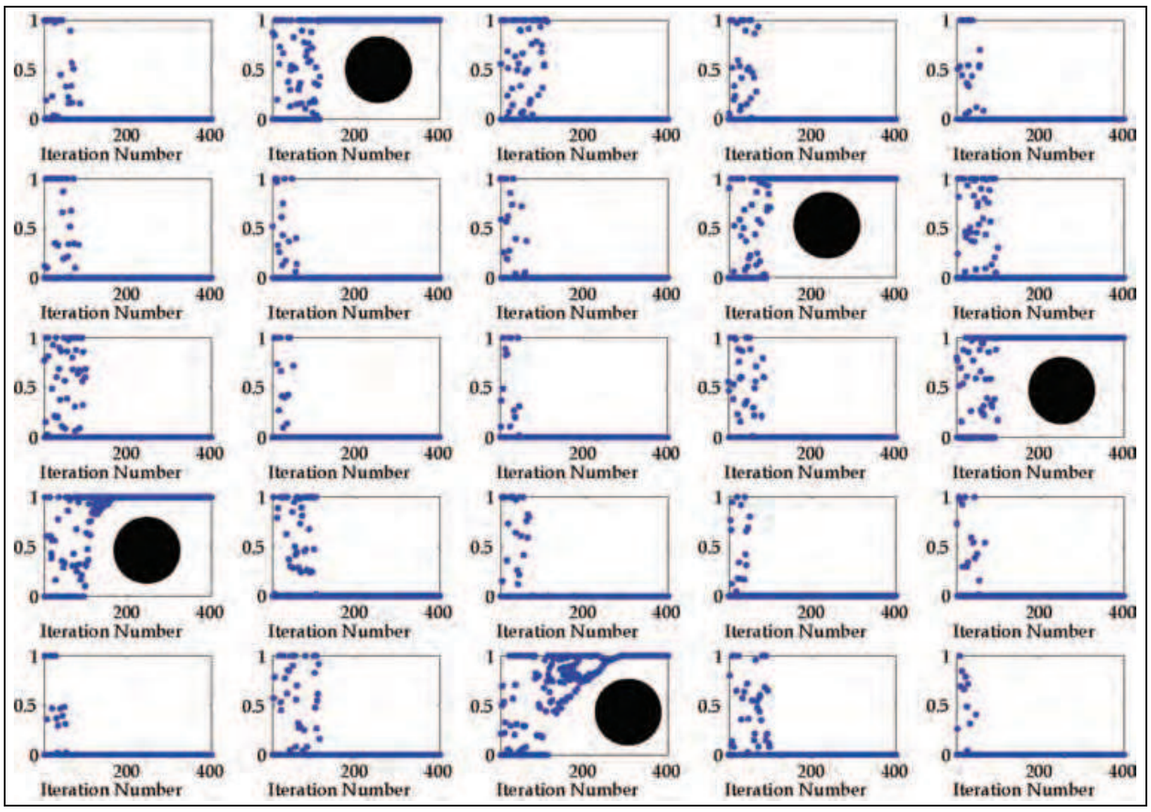

Fig. 7. Evolutions of outputs y_{ij} in Eq. (17). The trajectory approaches one, in the subfigures with a black point.

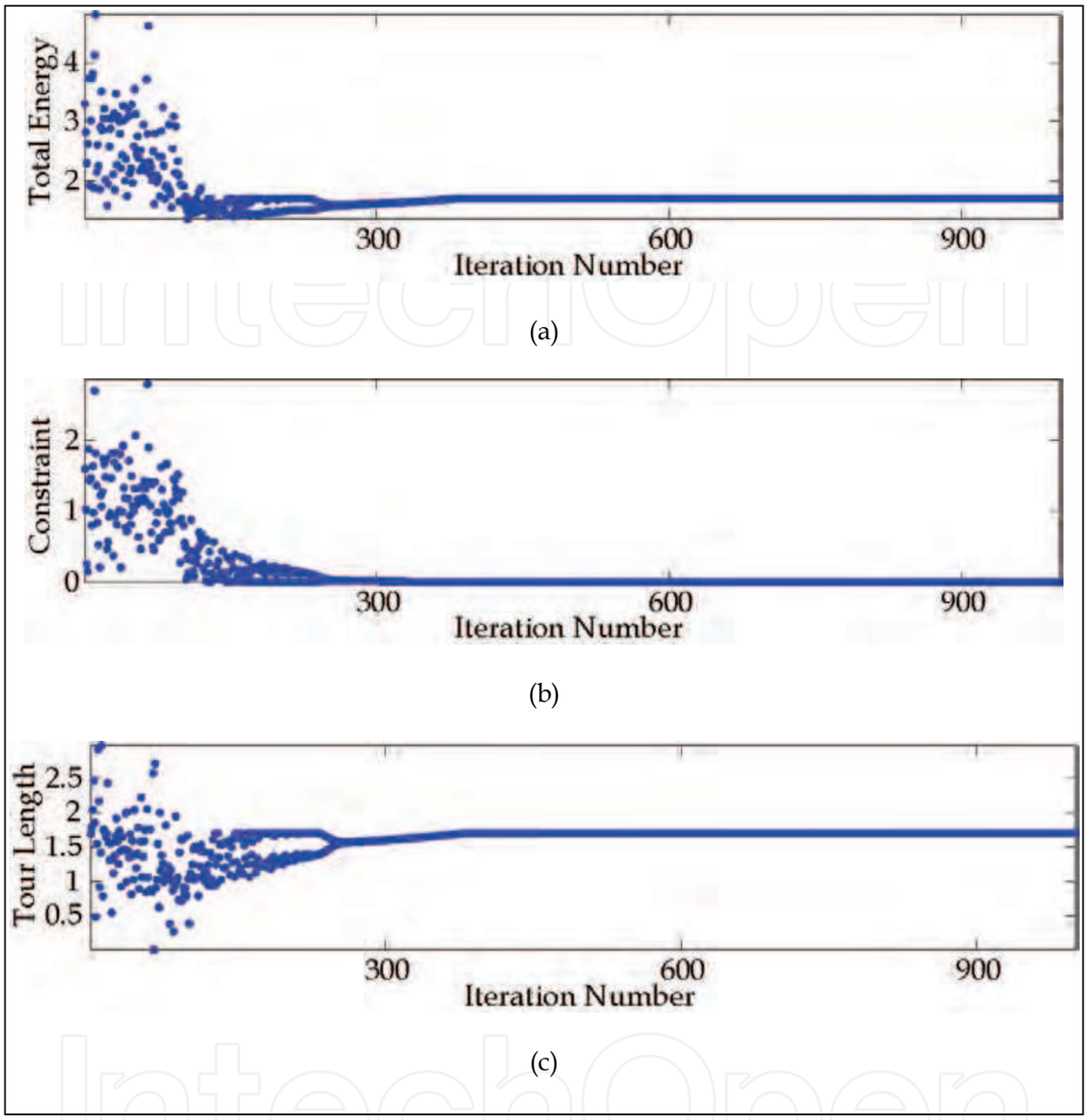

Fig. 8. Evolutions of (a) E, (b) the constraint term; (c) the tour-legnth term, in Eq. (23).

As another observation on the convergent and chaotic phases, we compute the Lyapunov exponents for the one-neuron equation:

$$x(t+1) = \mu x(t) + \omega[g_\varepsilon(x(t)) - a] + h,\ \omega = (1-\beta)^t \omega_0 \tag{28}$$

with parameters set as

$$\mu = 0.9;\ a = 0.65;\ \varepsilon = 0.01;\ h = 0.$$

Let us consider $\omega \in [\omega_0, 0]$ with $\omega_0 = -0.08$. If the system possesses a chaotic behavior, its maximal Lyapunov exponent is positive along the chaotic trajectory, and vice versa. The maximal Lyapunov exponent means the average of the maximal eigenvalue for linear part of the system along the chaotic trajectory in the ergodic sense. If the maximal Lyapunov exponent is negative, the system corresponds to stable phase. Notably, for a one-neuron map, there is only one Lyapunov exponent. The bifurcation diagram of Lyapunov exponent for the map (28) with parameters $\omega \in [\omega_0, 0]$ is shown in Fig. (9). It follows from Fig. (9) that there is a bifurcation point near $\omega_0 = -0.04$. In other words, the behavior changes near the point, and transforms from chaotic phase to stable phase. However, since our algorithm is based on $\omega = (1-\beta)^t \omega_0$, we also present the correspondence between iteration number t and parameter ω in Fig. (10). Similar computations can be applied to the multidimensional systems.

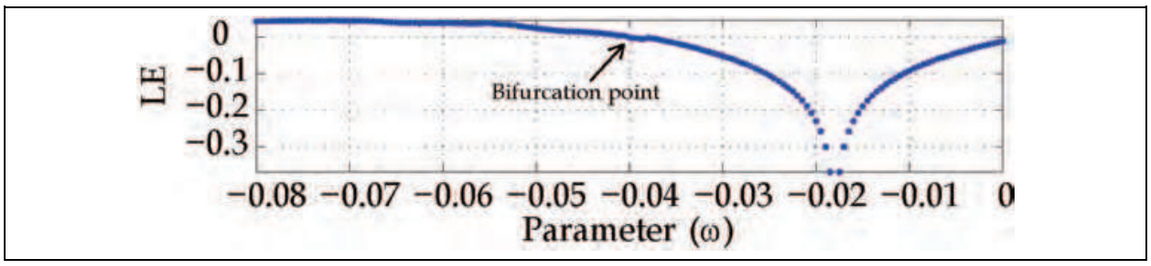

Fig. 9. Bifurcation diagram of Lyapunov exponent for one-dimensional map (28).

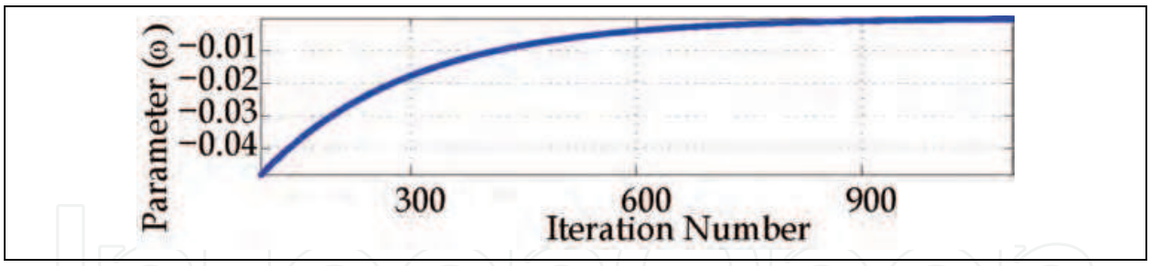

Fig. 10. Correspondence between iteration number *t* and parameter ω.

6. Conclusions

It has been more than two decades since artificial neural networks were employed to solve TSP. Among the efforts in improving the performance of this computational scheme, substantial achievements have been made in incorporating chaos into the system and developing mathematical analysis for finding the parameters in the chaotic regime and convergent regime. There are several advantages in employing the piecewise linear activation function. We have observed that the TCNN with piecewise linear activation function has better performance than with the logistic activation function in the

applications. In addition, the parameter conditions derived in this framework are much simpler than the ones for logistic activation functions.
There are certainly some further improvements to be developed; for example, in the decision of timing to cool down the process from the chaotic phase; observing and realization of the rotational symmetry and reversal symmetry in the solution structure, as well as conditions for stability of feasible solutions and instability of infeasible solutions.

7. Acknowledgements

This work is partially supported by The National Science Council, and the MOEATU program, of R.O.C. on Taiwan.

8. References

Aihara, K.; Takabe, T. & Toyoda, M. (1990). Chaotic neural network. *Phys. Lett. A*, 144, pp. 333–340.

Bersini, H. (1998). The frustrated and compositional nature of chaos in small Hopfield networks. *Neural Networks*, 11, pp. 1017–1025.

Bersini, H. & Senser, P. (2002). The connections between the grustrated chaos and the intermittency chaos in small Hopfield networks. *Neural Networks*, 15 , pp. 1197–1204.

Chen, L. & Aihara, K. (1995). Chaotic simulated annealing by neural network model with transient chaos. *Neural Networks*, 8, pp. 915–930.

Chen, L. & Aihara, K. (1997). Chaos and asymptotical stability in discrete-time neural networks. *Phys. D*, 104 , pp. 286–325.

Chen, L. & Aihara, K. (1999). Global searching ability of chaotic neural networks. *IEEE Trans. Circuits Systems I Fund. Theory Appl.*, 46 , pp. 974–993.

Chen, S. S. & Shih, C. W. (2002). Transversal homoclinic orbits in a transiently chaotic neural network. *Chaos*, 12, pp. 654–671.

Chen, S. S. & Shih, C. W. (2007). Transiently chaotic neural networks with piecewise linear output functions. *Chaos, Solitons & Fractals,* doi:10.1016/j.chaos.2007.01.103.

Hänggi, M.; Reddy, H. C. & Moschytz, G. S. (1999). Unifying results in CNN theory using delta operator. *IEEE International Symposium on Circuits and Systems*, 3, pp. 547–550.

Harrer, H. & Nossek, J. A. (1992). An analog implementation of discrete-time CNNs. *IEEE Transactions on Neural Networks*, 3, pp. 466–476.

Hopfield, J. J. & Tank, D. W. (1985). Neural computation of decisions in optimization problems. *Biol. Cybernet.*, 52, pp. 141–152.

LaSalle, J. P. (1976). The stability of dynamical systems. *Regional Conference Series in Applied Mathematics*. Philadelphia, PA:SIAM.

Li, T. & Yorke, J. (1975). Periodic three implies chaos. *Am Math Monthly*, 82, pp. 985–992.

Lorenz, E. (1963) Deterministic nonperiodic flow. *J. of Atmospheric Science*, 20, pp. 130–141.

Marotto, F. R. (1978). Snap-back repellers imply chaos in R^n . *J. Math. Anal. Appl.*, 63, pp. 199–223.

Marotto, F. R. (2005). On redefining a snap-back repeller. *Chaos Solitons & Fractals*, 25, pp. 25–28.

Nozawa, H. (1992). A neural network model as a globally coupled map and applications based on chaos. *Chaos*, 2, pp. 377–386.

Smith, K. (1999). Neural networks for combinatorial optimization: a review of more than a decade of research. *INFORMS J. on Computing*, 11(1), pp. 15–34.

Yamada, T. & Aihara, K. (1997) Nonlinear neurodynamics and combinatorial optimization in chaotic neural networks. *J. Intell. Fuzzy Syst.*, 5, pp. 53–68.

Traveling Salesman Problem

Edited by Federico Greco

ISBN 978-953-7619-10-7
Hard cover, 202 pages
Publisher InTech
Published online 01, September, 2008
Published in print edition September, 2008

The idea behind TSP was conceived by Austrian mathematician Karl Menger in mid 1930s who invited the research community to consider a problem from the everyday life from a mathematical point of view. A traveling salesman has to visit exactly once each one of a list of m cities and then return to the home city. He knows the cost of traveling from any city i to any other city j. Thus, which is the tour of least possible cost the salesman can take? In this book the problem of finding algorithmic technique leading to good/optimal solutions for TSP (or for some other strictly related problems) is considered. TSP is a very attractive problem for the research community because it arises as a natural subproblem in many applications concerning the every day life. Indeed, each application, in which an optimal ordering of a number of items has to be chosen in a way that the total cost of a solution is determined by adding up the costs arising from two successively items, can be modelled as a TSP instance. Thus, studying TSP can never be considered as an abstract research with no real importance.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Shyan-Shiou Chen and Chih-Wen Shih (2008). Solving TSP by Transiently Chaotic Neural Networks, Traveling Salesman Problem, Federico Greco (Ed.), ISBN: 978-953-7619-10-7, InTech, Available from: http://www.intechopen.com/books/traveling_salesman_problem/solving_tsp_by_transiently_chaotic_neural_networks

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen