第 39 卷 第6(A)期
2000 年 11 月

中山大学学报（自然科学版）
ACTA SCIENTIARUM NATURALIUM
UNIVERSITATIS SUNYATSENI

Vol.39 No.6 (A)
Nov. 2000

文章编号：0529-6579（2000）06-0285-05

A Generalized Updating Rule for Discrete Hopfield Neural Network with Delay*

QIU Shen-shan, LIU Yong-qing, DENG Fei-qi
(College of Electronic and Information Engineering,
South China University of Technology, Guangzhou 510640, China)

Abstract: In the paper, the Hopfield neural network with delay is studied from the standpoint of taking it as optimization computational model. A generalized updating rule for the network (GURD) is established. It is proved that in any sequence of updating rule modes, the GURD monotonously converges to a stable state of the network. All ordinary HNND algorithms are instances of the GURD. It is shown that the convergence conditions of GURD expand into more loose conditions in applications and break through original constrained conditions. New updating rule mode and restrictive conditions can guarantee the network to achieve local maximum of the energy function.

Keywords: neural network; updating rule; optimization problem; convergence

CLC number: TP 39 **Document code**: A

1 Introduction

Recently the convergence of various connectionist models including discrete Hopfield neural network (DHNN)[1~3], continuous Hopfield neural network (CHNN)[1] and the network with delay[4] has been investigated by researchers in many fields. Since Hopfield[1] first applied DHNN to traveling salesman problems, DHNN (which have a very simple and efficient hardware implementation) has been shown to have the capability to provide powerful approaches for a wide variety of combinatorial optimization problems. However, Hopfield's approach has been subjected to severe criticism, concerning the quality of solutions, especially for large-size combinatorial optimization problems. In fact, there have been many negative and positive discussions, notably from Wilson and Pawley[2]. The other aspects of the theory research concerned with the computing complexity problem[5], computational power, finite automata equivalency and dynamic system simulation by DHNNs, which are beyond the scope of this paper.

It is well known that the DHNN without delay has an important property, i.e., it always converges to a stable state when operating in a serial mode and to a cycle of length at most 2 when operating in a parallel mode[1~3,6]. However, in most optimization problems, the corresponding matrix and updating rule mode do not satisfy the above convergence conditions. So convergence problem is consid-

* 基金项目：This work was supported in part by the national nature sciences foundation and nature sciences foundation of Guangdong province and nature sciences foundation of South China University of Technology (69934030; 990595).

收稿日期：2000-09-18；作者简介：邱深山（1962～），男，讲师，博士后；研究方向：机器学习、数据挖掘、神经网络.

ered one of the important theory problems of the neural networks. To correct the deficiencies of the Hopfield network, various attempts such as to optimally select parameters and to eliminate infeasible solutions have been made. Among these attempts, some are to maintain the structure of the Hopfield network, the others are to modify the network structure. In [8] we constructed a DHNN with delay and presented a updating rule which is similar to one in Hopfield network without delay[1~3,6]. Unfortunately, the corresponding to convergence conditions is constrained strictly, therefore its application region was abated. We have concluded from these results that the energy function is dependent on the updating rule mode. Whether the constrained convergence conditions can be expanded into more loose is regarded as an open problem[8~10]. It is well known that a detailed specification for the convergence of DHNN is very important to investigate most real applications.

In fact, the convergence, network architectures, network initial conditions as well updating modes are intimately related. In this paper we extend convergence conditions[8] and give a clear answer to what updating rule mode can make energy function be decreased monotonically with respect to evolution time, i.e., based on the modifying energy function (bivariate energy function) by means of the modified DHNN as well as its updating mode, the convergence of DHNN with delay is obtained .

The present paper has the following organization. Section 1 is the introduction. Section 2 gives a brief review for the DHNN without and with delay. Section 3 introduces some definitions and notations used in this paper. Section 4 investigates the convergence of DHNN with delay and gives our main result. The last section offers the conclusion of this paper.

2 Hopfield networks

DHNN can be viewed as a graph in which each node represents a neuron in the network. The order of the network is the number of neuron in the network. The connection strength between each pair of nodes (neurons) is represented by a weight associated with the edge connecting the corresponding pair of nodes. Formally, let $N=(W,\theta)$ be a DHNN with n nodes denoted by $x_1,x_2,\cdots,x_n$. In the pair (W,θ), $w=(w_{ij})$ is a $n\times n$ real matrix with w_{ij} representing the connection weight from node (neuron) x_j to x_i; $\theta=(\theta_i)$ is a n-dimensional real vector with θ_i representing the threshold attached to neuron x_i. The state of the i^{th} node at time t is denoted by $x_i(t), t=0,1,2,\cdots$. Each neuron (node) is assumed to have two possible states, 1 or -1. The state of the network at time t is the vector $X(t)=[x_1(t),x_2(t),\cdots,x_n(t)]^{\mathrm{T}}$. In general, the state of the network at time $t+1$ is a function of $\{W,\theta\}$ and the state of network at time t. The network is thus completely determined by the parameters $\{W,\theta\}$, the initial value $X(0)$ of the states, and the manner in which the nodes are updated (evolved). If at time step t, node $x_i$is chosen to update, then at the next step

$$x_i(t+1) = \mathrm{sgn}(h_i(x(x(t)))) = \begin{cases} 1, & h_i(x(t)) \geqslant 0 \\ -1, & h_i(x(t)) < 0 \end{cases} \tag{1}$$

where $h_i(t)=\sum_{j=1}^{N} w_{ij}x_j(t)-\theta_i$.

If only one neuron is allowed to change state at each step by performing the state evolving equation (1) while the states of the other nodes are unchanged, then network is said to be operating in a serial mode. If all the nodes are chosen to be updated at each step by performing the state evolving equation (1), then network is said to be operating in a parallel mode[1~3,6].

A DHNN with delay is a computation model, similar to a multilayer percetron in which all connections feed forward. The difference between a DHNN with delay and a multilayer percetron is as follows. The inputs to any node x_i for DHNN with delay can consist of the outputs of earlier nodes not only during the current step t, but also during some number d of previous time $(t-1, t-2, \cdots, t-d)$. Following Bruck[3] and Orponen[5], the state of a neuron is updated according to the following equation:

$$x_i(t+1) = \operatorname{sgn}(H_i(x(t))) = \begin{cases} 1, & H_i(x(t)) \geqslant 0 \\ -1, & H_i(x(t)) < 0 \end{cases} \tag{2}$$

where $H_i(x(t)) = \sum_{k=0}^{d}\sum_{j=1}^{n} w_{ij}^k x_j(t-k) - \theta_i(t)$.

For the convenience of the later analysis in this paper, we only discuss special case $d = 1, \theta(t) \equiv 0$ and use the triple $N = (w^0, w^1, \theta)$ to represent the model (2), although the following is different from here in output function (not necessary constrained sign form function) and the updating rule mode. For a special case (i.e., $w^1 = O_{n\times n}$), $N = (w^0, w^1, \theta)$ will represent DHNN model[1~3,6].

The DHNN with delay without feedback connections is called TDNN which has successfully been used in a number of real world applications.

3 Updating rule of HNN with delay

Definition and notation

Let $B^n \equiv \{-1, 1\}^n$ represent the set of all vectors $X = (x_1, x_2, \cdots, x_n)^{\mathrm{T}}$, where each component x_i takes value $+1$ or -1. We propose a new updating rule for Hopfield neural network with delay in order to extend the convergence condition. The following is the definition of a generalized updating rule.

Definition 1 Let $N = (w^0, w^1, \theta)$ represent DHNN with delay, starting from any initial vector $X(1) = X(0) \in B^n$, given a sequence of neural index subsets $L(k) \subseteq \{1, 2, \cdots, n\}$ for $k \geqslant 1$, the neural network update in the following form:

When $k = 1$, $L(1) \subseteq (1, 2, \cdots, n\}$, is neural index subset by means of randomly selected (in general $L(1) \equiv \{1, 2, \cdots, n\}$), $x_i(k+1) = x_i(k)$, if $i \notin L(k)$ and

When $k \geqslant 1$, 1 For $i \in L(k)$, the following is updating rule

$$x_i(k+1) = \begin{cases} 1, \text{ if } x_i(k) = -1 \text{ and } H_i(k) > \sum\limits_{j\in L(k)} |w_{ij}^0| + 2\sum\limits_{j\in L(k-1)} |w_{ij}^1| \\ -1, \text{ if } x_i(k) = 1 \text{ and } H_i(k) < -\sum\limits_{j\in L(k)} |w_{ij}^0| - 2\sum\limits_{j\in L(k-1)} |w_{ij}^1| \\ \qquad x_i(k), \qquad\qquad \text{otherwise} \end{cases} \tag{3}$$

Where $H_i(k) = \sum_{j=1}^{n}(w_{ij}^0 x_j(k) + w_{ij}^1(k-1) - \theta)$.

The updating mode is called the generalized updating rule for discrete neural network with delay, it is abbreviated to GURD (the Generalized Updating Rule for DHNN with Delay).

A sequence of neural index subsets $L(k) \subseteq \{1, 2, \cdots, n\}$ satisfying the above conditions is generalized in evolution mode. It includes a serial mode and parallel mode for neural network with delay when sequence of neural index subsets is chosen satisfying $|L(k)| = 1$ *or* n[1~3,6~7].

Definition 2 The function $E(u, v)$ defined below is called bivariate energy function on B^n for $N = (w^0, w^1, \theta)$:

$$E(u, v) = u^{\mathrm{T}} w^0 u + 2u^{\mathrm{T}} w^1 v - 2u^{\mathrm{T}}\theta \tag{4}$$

where $u = x(t)$, $v = x(t-1)$. In short, the bivariate energy function is written as $E(t) \equiv E(x(t), x(t-1))$.

4 Main results

Based on the GURD mode, the neural network with delay has the following convergence property.

Theorem Let $N = (w^0, w^1, \theta)$ be a DHNN with delay, operating in GURD mode. If w^1 is a row-diagonal dominant matrix, i.e. $w^1_{ii} \geqslant \sum_{j \neq i} |w^1_{ji}|$, for any $k \geqslant 1$, $\Delta x(k) \neq 0$ or$\Delta x(k-1) \neq 0$, it is guaranteed that

$$\Delta E(k) \equiv E(x(k+1), x(k)) - E(x(k), x(k-1)) \geqslant 0$$

The equality holds if only and if $x(k+1) = x(k) = x(k-1)$.

Proof According to definition 2

$$E(u, v) = u^{\mathrm{T}} w^0 u + 2u^{\mathrm{T}} w^1 v - 2u^{\mathrm{T}} \theta$$

where $u = x(k), v = x(k-1), \theta$ is a threshold vector $\theta = (\theta_1, \theta_2, \cdots, \theta_n)^{\mathrm{T}}$, then

$$|E(k)| \equiv |E(x(k), x(k-1))| \leqslant \sum_{i=1}^{n} \sum_{j=1}^{n} (|w^0_{ij}| + 2|w^1_{ij}|) + 2\sum_{i=1}^{n} |\theta_i| \qquad (5)$$

Let $\Delta x_i(k) \equiv x_i(k+1) - x_i(k), \Delta E(k) \equiv E(k+1) - E(k), (\Delta E(k))_i \equiv (E(k+1) - E(k))|_{\Delta x_i(k) \neq 0}$

Then

$$\Delta E(t) \equiv x^{\mathrm{T}}(k+1) w^0 x(k+1) + 2x^{\mathrm{T}}(k+) w^1 x(k) = 2x^{\mathrm{T}}(k+1)\theta - x^{\mathrm{T}}(k) w^0 x(k) - 2x^{\mathrm{T}}(k) w^1 x(k-1) + 2x^{\mathrm{T}}(k)\theta$$

For any $k \geqslant 2, i \in L(k) \subseteq \{1, 2, \cdots, n\}$ and operating the GURD mode, we have

$$\Delta E(k) = \sum_{i=1}^{n} 2\Delta x_i(k) H_i(x(k)) + \sum_{i=1}^{n} \sum_{j=1}^{n} (\Delta x_i(k))(\Delta x_i(k)) w^0_{ij} + 2\sum_{i=1}^{n} \sum_{j=1}^{n} x_i(k+1) \Delta x_i(k-1) w^1_{ij} \qquad (6)$$

(1) If $\Delta x_i(k) > 0$, necessarily $x_i(k) = -1$ and $x_i(k+1) = 1$.

$H_i(k) > \sum_{j \in L(k)} |w^0_{ij}| + 2 \sum_{j \in L(k-1)} |w^1_{ij}|$ is a sufficient condition for $(\Delta E(k))_i > 0$

(2) If $\Delta x_i(k) < 0$, necessarily $x_i(k) = 1$ and $x_i(k+1) = -1$.

$H_i(k) < -\sum_{j \in L(k)} |w^0_{ij}| - 2 \sum_{j \in L(k-1)} |w^1_{ij}|$ is a sufficient condition for $(\Delta E(k))_i > 0$

(3) Since w^1 is a diagonal dominant matrix, then $(\Delta E(k))_i \geqslant 0$ when $\Delta x_i(k)_i = 0$, $\Delta E(k-1)_i \neq 0$. Any nonzero update $\Delta x(k) \neq 0$ or$\Delta x(k-1)_i \neq 0$ in the GRUD is case (1), (2) and (3), then $\Delta E(k)_i \geqslant 0$.

It is noted that the theorem is generalized theorem of [8 ~ 10] which extends constrained conditions of weight matrices w^0 and w^1. In short, w^0's diagonal elements are non - negative (serial mode) and w^0 is a nonnegative definite (parallel mode) is not necessary for GRUD mode. It has been shown that updating rule process of the network and activation process of neurons has been characterized by dynamic process. But many researchers have also held their courses that the active output function of neural network be confined to assign sigmoid function and the serial, parallel and partial parallel mode, is considered as important updating rule mode which is nothing more than these modes.

Corollary After a finite number of updates, under the GURD mode, the neural network with delay will converges to a fixed point.

In fact, the fixed point of DHNN with delay corresponds to a local maximum point of energy function which it is our research main purpose and potential active interest. On the other hand, the convergence property is indispensable for practical application design of DHNN with delay.

5 Conclusions

In this article we have demonstrated how a generalized updating rule can be made energy function increase monotonously for neural network with delay. The GURD mode can operate in any sequence of updating modes. We prove the neural network with delay is convergent in operating the GURD mode. The algorithm of paper [8 ~ 10], including original MHNN algorithm and partial parallel mode algorithm[1~3,6~7,11] can be easily derived from our main result.

References:

[1] HOPFIELD J J. Neural networks and physical systems with emergent collective computational ablities[J]. Proc Nat Acad Sci USA ,1982,79:2554 ~ 2558.

[2] WILSON G N, PAWLEY G S. On the stability of the traveling salesman problem algorithm of Hopfield and tank [J]. Biological Cybernetics,1988,58:63 ~ 70.

[3] BRUCK J. On the convergence properties of the Hopfield model[J]. Proceedings of IEEE,1990,78(10):1579 ~ 1585.

[4] DANIEL S CLOUSE, GILES C LEE, etc. Time-delay neural network: Representation and induction of finite-state machines[J]. IEEE Transactions on neural networks, 1997,8(5):1065 ~ 1070.

[5] ORPONEN P. Neural networks and complexity theory. Proc.17th International Symposium on Mathematical Foundations of Computer Science. Lecture Notes in Computer Science 629, Springer-Verlag, Berlin ,Germany,1992,50 ~ 61.

[6] XU ZONGBIN, HU GUOQING, KWONG CHUNGPING. Asymmetric Hopfield-type networks: theory and applications[J]. Neural Networks,1996,9(3):483 ~ 501.

[7] YI SUN. A generalized updating rule for modified Hopfield neural network for quadratic optimization[J]. Neurocomping,1998,19:133 ~ 143.

[8] QIU SHENSHAN, LIU YONGQING. Convergence of Discrete Hopfield-type Neural Networks with Delay. American control conference 2000, Chicago. Illinois, 2000,6(1):658 ~ 659

[9] QIU SHENSHAN, XU XIAOFEI, etc. Convergence of discrete Hopfield-type neural network with time-delay in a serial mode[J]. Journal of Computer Research & Development, 1999,36(5):557 ~ 563.

[10] QIU SHENSHAN, XU XIAOFEI. Matrix criterion of Dynamic analysis in discrete neural networks with delay[J]. Journal of Software. 1999,10(10):1108 ~ 1113.

[11] DONGLIANG LEE. New stability conditions for Hopfield networks in partial simultaneous update mode[J]. IEEE Transactions on Neural Networks, 1999,10(4):975 ~ 978.

延迟离散 Hopfield 神经网络的一般演化规则

邱深山*，刘永清，邓飞其

摘　要：从优化计算模型的角度研究延迟离散 Hopfield 网络，提出了它的一般演化规则(GURD)．在任何演化序列下，证明了 GURD 单调收敛到网络的稳定状态，已有的延迟神经网络的演化算法为其特例，同时 GURD 的收敛条件已经突破了已有的限制，有更大的应用空间．新的演化规则和收敛条件保证了延迟神经网络收敛到能量函数的局部最大值．

关键词：神经网络；延迟；演化规则；优化问题；收敛性

中图分类号：TP 39

* 华南理工大学电子与信息学院，广东 广州 510640