2010 International Symposium on Nonlinear Theory and its Applications
NOLTA2010, Krakow, Poland, September 5-8, 2010

Discrete Higher Order Inverse Function Delayed Network

Takahiro Sota†, Yoshihiro Hayakawa‡, Shigeo Sato† and Koji Nakajima†

†Research Institute of Electrical Communication, Tohoku University
2-1-1 Katahira, Aoba-ku, Sendai-shi, 980-8577, Japan
‡Sendai National College of Technology
4-16-1 AyashiChuoh, Aoba-ku, Sendai-shi, 989-3128, Japan
Email: sota@nakajima.riec.tohoku.ac.jp, hayakawa@sendai-nct.ac.jp, shigeo@riec.tohoku.ac.jp, hello@riec.tohoku.ac.jp

Abstract—The Inverse function Delayed model (ID model) is a neuron model that has dynamics effected by negative resistance. The negative resistance can destabilize local minimum states that are undesirable network responses. Actually, we have demonstrated that the ID network can perfectly remove all local minima with N-Queen problems or 4-Color problems where stationary states are only correct answers. Meanwhile, about the case of Traveling Salesman Problems or Quadratic Assignment Problems, we also applied the ID network to them with the same method by introducing quartic-form energy function and higher order connections.

In this paper, we introduce the discrete ID model with higher order connections to achieve the hardware implementation of the network.

1. Introduction

Hopfield et al. proposed a neural network that had dynamics of moving along the gradient of the quadratic-form energy function [1]. This network can find the solutions of combinatorial optimization problems (COP) by assigning the optimal solution to the global minima of the energy function. However, the network state is often trapped into local minima except for the global minima, and it is known as the local minimum problem.

As an improvement way of this problem, the method of using the Inverse function Delayed model (ID model) [2] has been proposed. One of important properties of ID model is a negative resistance. The region of the negative resistance is controllable, then the effect of the negative resistance can destabilize the undesirable local minimum states of the energy function with appropriate setting. Especially, N-Queen problems or 4-Color problems are able to be solved with 100% success rate by using the ID model[3].

Meanwhile, in the case of Traveling Salesman Problems (TSP) or Quadratic Assignment Problems(QAP), we have proposed the Higher-order Connection ID model (HC-ID model) and the quartic-form energy function for these problems[4]. The HC-ID model is the ID model with the higher-order synaptic connections, and the HC-ID network with 3rd-order synaptic connections has a quartic-form energy function. This network can destabilize only the undesirable local minimum states of TSP or QAP by introducing the quartic-form energy function. Actually, optimal solutions of some 4-city TSP's and 4-size QAP's are obtained in numerical experiments. However, it is difficult to apply the HC-ID network with computer simulation to large size problems because the HC-ID network requires much computation time to update neuron states. Consequently, we aim to implement the HC-ID network on hardware to avoid this problem.

In this paper, we introduce the discrete HC-ID network instead of the continuous HC-ID network to achieve the hardware implementation. The discrete network is superior for hardware implementation. We propose the discrete time network, at first, and then derive the binary output network from the discrete time network. Finally, we confirm the expected result by solving a TSP and a QAP numerically.

2. Discrete HC-ID Network

2.1. Continuous HC-ID Network

The continuous type ID model with 3rd order synaptic connections is described by following equations[4]:

$$\begin{aligned}\tau_u \frac{du_i}{dt} =& \sum_{j=1}^{N}\sum_{k=1}^{N}\sum_{l=1}^{N} w_{ijkl}x_j x_k x_l \\ &+ \sum_{j=1}^{N}\sum_{k=1}^{N} w_{ijk}x_j x_k + \sum_{j=1}^{N} w_{ij}x_j + h_i, \qquad (1)\\ \tau_x \frac{dx_i}{dt} =& u_i - g(x_i), \qquad (2)\end{aligned}$$

where N is the number of neurons, and u_i , x_i and h_i are the internal state, the output and the bias of neuron i, respectively. $w_{ijk\cdots}$ is the synaptic weight from neurons $j, k, \cdots$ to neuron i. τ_u and $\tau_x (\ll \tau_u)$ are the time constants of the internal state and the output, respectively.

Substituting Eq. (2) into Eq. (1), the following equation is obtained:

$$\tau_x \frac{d^2 x_i}{dt^2} + \eta(x_i)\frac{dx_i}{dt} = -\frac{\partial U}{\partial x_i}, \qquad (3)$$

where

$$\eta(x_i) = \frac{dg(x)}{dx}\bigg|_{x=x_i}, \tag{4}$$

$$\begin{aligned} U = &-\frac{1}{4\tau_u}\sum_i\sum_j\sum_k\sum_l w_{ijkl}x_ix_jx_kx_l \\ &-\frac{1}{3\tau_u}\sum_i\sum_j\sum_k w_{ijk}x_ix_jx_k \\ &-\frac{1}{2\tau_u}\sum_j w_{ij}x_ix_j - \frac{1}{\tau_u}\sum_i h_i. \end{aligned} \tag{5}$$

$\eta(x_i)$ and U denoting a friction coefficient and a potential, respectively, Equation (3) shows that the HC-ID model operates as the particle in potential U. Moreover the negative resistance appears in the area where the gradient of the function g is negative. These dynamics of HC-ID model are similar to the normal ID model, except the potential U of 3rd HC-ID model is a quartic-form function and there is no effect of dissipation.

2.2. Derivation of Discrete Time Model

The discrete time HC-ID model can be derived by applying Euler method to Eqs. (1) and (2). The basic equations of the discrete time HC-ID model are

$$\begin{aligned} u_i(t+1) =& u_i(t) + T_\Delta\Bigg(\sum_{j=1}^{N}\sum_{k=1}^{N}\sum_{l=1}^{N} w_{ijkl}x_j(t)x_k(t)x_l(t) \\ &+\sum_{j=1}^{N}\sum_{k=1}^{N} w_{ijk}x_j(t)x_k(t) + \sum_{j=1}^{N} w_{ij}x_j(t) + h_i\Bigg), \end{aligned} \tag{6}$$

$$x_i(t+1) = x_i(t) + T_\Delta\frac{\tau_u}{\tau_x}\Big(u_i(t) - g(x_i(t))\Big), \tag{7}$$

where

$$T_\Delta = \Delta/\tau_u \leq 1, \tag{8}$$

and Δ is a time increment.

2.3. Limitation of Output and Inner State

Although the output range of HC-ID model has no limitation normally, the range should be between 0 and 1 to apply to COP. Hence the discrete time HC-ID model restricts the range of the output as follows when calculating the next step output:

$$x(t+1)\begin{cases} = 0 & \text{if } x(t+1) < 0 \\ = 1 & \text{if } x(t+1) > 1 \end{cases}. \tag{9}$$

Moreover, the inner state u is also limited as follows due to the output restriction:

$$u(t+1) = u(t)\begin{cases} & \text{if } x(t) = 0 \text{ and } u(t+1) < u(t) < 0 \\ & \text{if } x(t) = 1 \text{ and } u(t+1) > u(t) > 0 \end{cases}. \tag{10}$$

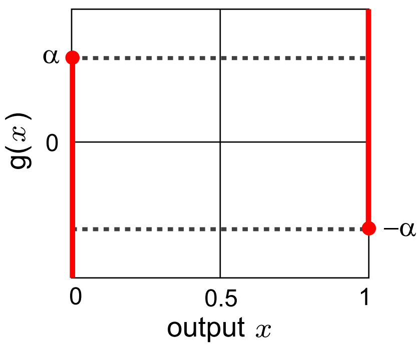

Figure 1: The g-function used in the binary output model.

2.4. Derivation of Binary Output Model

The output of discrete time HC-ID model becomes binary values when the parameter τ_x in Eq. (7) approximates to 0. The output is obtained as follows under the limitation of $\tau_x \to 0$:

$$x_i(t+1) = \begin{cases} 1 & (u_i(t) - g(x_i(t)) > 0) \\ 0 & (u_i(t) - g(x_i(t)) < 0) \\ x_i(t) & (u_i(t) - g(x_i(t)) = 0) \end{cases}. \tag{11}$$

Moreover, the g-function for binary output is defined by following equation:

$$g(x)\begin{cases} \leq \alpha & (x = 0) \\ \geq -\alpha & (x = 1) \end{cases}, \tag{12}$$

where α is a positive parameter. This g-function no longer has negative resistance effect, it has only the hysteresis effect. That effect increases with increasing α. And the output of the binary model is described as follows from Eqs. (11) and (12):

$$x_i(t+1) = \begin{cases} 1 & (u_i(t) > \alpha) \\ 0 & (u_i(t) < -\alpha) \\ x_i(t) & \text{otherwise} \end{cases}. \tag{13}$$

We use this discrete HC-ID model afterward.

2.5. States Stability of Discrete HC-ID Network

Eq. (6) is able to rewrite as Eq. (14) by using the potential U:

$$u_i(t+1) = u_i(t) - T_\Delta\frac{\partial U}{\partial x_i}\bigg|_{x_i=x_i(t)}. \tag{14}$$

Therefore, the inner state u_i keeps decreasing if $\partial U/\partial x_i$ is positive, and finally the output x_i becomes 0. In contrast, the output becomes 1 eventually if $\partial U/\partial x_i$ is negative.

3. Applying Quartic Energy Function

3.1. Quartic Energy Function

The quartic energy function for combinatorial optimization problems is [4]

$$\begin{aligned} E_{4\mathrm{TH}} = & \frac{A}{2}\sum_{i=1}^{n}\Big(\sum_{x=1}^{n} x_{xi} - 1\Big)^2 + \frac{A}{2}\sum_{x=1}^{n}\Big(\sum_{i=1}^{n} x_{xi} - 1\Big)^2 \\ & + \frac{B}{2}\sum_{x=1}^{n}\sum_{i=1}^{n}\sum_{y=1}^{n}\sum_{j=1}^{n} b_{xi,yj}x_{xi}x_{yj}(1 - x_{xi}x_{yj}) \\ & + \frac{C}{2}\Big(\sum_{x=1}^{n}\sum_{i=1}^{n}\sum_{y=1}^{n}\sum_{j=1}^{n} b_{xi,yj}x_{xi}x_{yj}\Big)^2, \end{aligned} \tag{15}$$

where n is the size of a problem and A, B and C are coefficients that have positive value. $b_{xi,yj} \in \boldsymbol{b}^{[N\times N]}$ is a cost value when the neuron (x, i) and (y, j) fire, and the cost matrix $\boldsymbol{b}$ has to be modified by the applied problem. The first and second terms of Eq. (15) have minimum value if only one neuron fires in each row and column, and the third term of Eq. (15) is minimized when all of output values are 0 or 1. These three terms express constrained conditions, and they will be 0 when the conditions are satisfied. Under these conditions, the fourth term expresses a squared value of cost shown by the network state.

Moreover, the partial differentiation of $E_{4\mathrm{TH}}$ with respect to a output x_{ah} is expressed as follows if the constrained conditions are satisfied:

$$\begin{aligned} \frac{\partial E_{4\mathrm{TH}}}{\partial x_{ah}} = & \sum_{y}\sum_{j} b_{ah,yj}x_{yj}\times \\ & \left\{2B\Big(\frac{1}{2} - x_{ah}\Big) + 2C\sum_{z}\sum_{k}\sum_{w}\sum_{l} b_{zk,wl}x_{zk}x_{wl}\right\} \\ = & \begin{cases} \sum_{j} b_{ah,yF[y]}\cdot 4C\left\{\frac{B}{4C} + c_{\mathrm{sol}}(\boldsymbol{x})\right\} & (x_{ah} = 0) \\ \sum_{j} b_{ah,yF[y]}\cdot 4C\left\{-\frac{B}{4C} + c_{\mathrm{sol}}(\boldsymbol{x})\right\} & (x_{ah} = 1) \end{cases}, \end{aligned} \tag{16}$$

where $F[y]$ is the column index of a firing neuron in row y and

$$c_{\mathrm{sol}}(\boldsymbol{x}) = \frac{1}{2}\sum_{z}\sum_{w} b_{zF[z],wF[w]}\,, \tag{17}$$

that is, $c_{\mathrm{sol}}(\boldsymbol{x})$ expresses the cost value obtained by using the solution represented by network state $\boldsymbol{x}$. Equation (16) shows that the sign of $\partial E_{4\mathrm{TH}}/\partial x_{ah}$ depends on the relationship between the coefficient $B/4C$ and the cost of solution $c_{\mathrm{sol}}(\boldsymbol{x})$ when $x_{ah} = 1$. That is to say, $\partial E_{4\mathrm{TH}}/\partial x_{ah}$ has negative value if $B/4C < c_{\mathrm{sol}}(\boldsymbol{x})$ holds, and $\partial E_{4\mathrm{TH}}/\partial x_{ah}$ is positive if $B/4C > c_{\mathrm{sol}}(\boldsymbol{x})$ holds.

3.2. Relationship between Network States Stability and Coefficients of Quartic Energy Function

When the energy function is applied to the HC-ID network, the weight matrix and bias value of the neural network are constructed to make the potential U equal to the energy function. Hence a condition $\partial U/\partial x_i = \partial E_{4\mathrm{TH}}/\partial x_i$ is satisfied if the quartic energy function is applied to the HC-ID network. It means that $x_{ah} = 1$ is stable if $\partial E_{4\mathrm{TH}}/\partial x_{ah}$ is negative, and $x_{ah} = 1$ is unstable if $\partial E_{4\mathrm{TH}}/\partial x_{ah}$ is positive. Hence only the state whose cost is smaller than the value of $B/4C$ is stable, according to the discussion of the preceding section. Consequently we can obtain solutions whose cost is smaller than $B/4C$. The solutions contain both of the global and local minimum states. The local minimum state solutions decrease with decreasing $B/4C$. Finally, only the global minimum states are stabilized and all local minimum states are destabilized by using the discrete HC-ID network when following condition is satisfied:

$$c_{\mathrm{sol}}(\boldsymbol{x}_0) < B/4C < c_{\mathrm{sol}}(\boldsymbol{x}_1), \tag{18}$$

where $c_{\mathrm{sol}}(\boldsymbol{x}_0)$ and $c_{\mathrm{sol}}(\boldsymbol{x}_1)$ are the cost value of optimal solution and the 2nd optimal solution, respectively. This condition being the same as the one of continuous[4], therefore the discrete HC-ID network can solve COP by the same way as the continuous network.

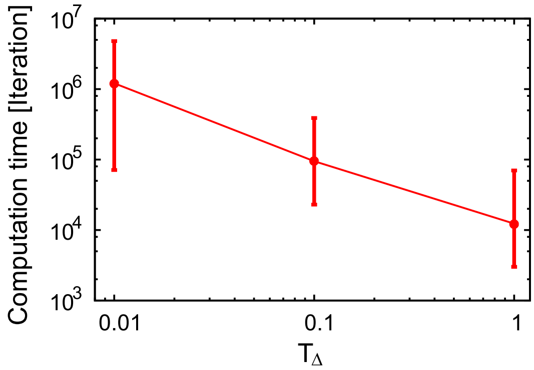

Figure 2: Computation time as a function of T_Δ in a 6-TSP $A = 22.7$, $B = 1.50$, $C = 0.10$, and $\alpha = 400.0$.

4. Simulation Result

This section shows the simulation results of 6-city TSP and 4-size QAP. In these simulation random initial states are used, and the result is obtained by 48 simulations. The coefficient A of the energy function is set to satisfy the constrained condition, and $B/4C$ is set to satisfy Eq. (18).

4.1. Success Rate and Computation Time Dependency on Parameter T_Δ

At first, it is investigated how the computation time depends on the parameter T_Δ by numerical experiments of a 6-city TSP. The success rate is always 100% for any value of T_Δ, and the variation of computation time is shown in Fig. 2. The solid circles in Fig. 2 show the average of computation time. This result suggests that T_Δ has no impact to

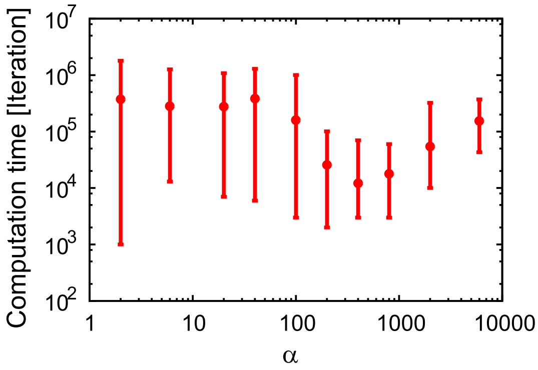

Figure 3: Computation Time as a function of α in a 6-TSP $A = 22.7$, $B = 1.50$, $C = 0.10$, and $T_\Delta = 1$.

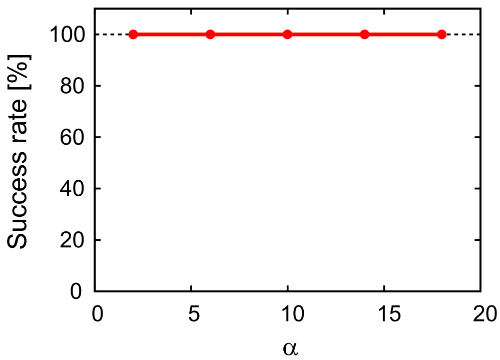

Figure 4: Success Rate as a function of α in a 4-QAP $A = 376.7$, $B = 4.68$, $C = 0.10$, and $T_\Delta = 1$.

the success rate, and the computation time decreases with increasing the parameter α.

4.2. Success Rate and Computation Time Dependency on Parameter α

Next, computation time dependency on the parameter α is investigated. At this time T_Δ is set to 1 from the foregoing result. In all α range, the optimal solution is obtained with 100% success rate. The computation time is shown as a function of α in Fig. 3. The computation time takes a minimum value at $\alpha \cong 400$, it is important to set an appropriate parameter α for solving COP with smaller iterations .

4.3. The Result of 4-QAP

Finally, we compare the success rate of the discrete network with that of the continuous network by solving a 4-size QAP. Figure 4 shows the success rate dependency on the parameter α in the 4-QAP. It shows that the discrete HC-ID network can solve the 4-QAP with 100% success rate at any α value.

However, in case of using continuous HC-ID network, it

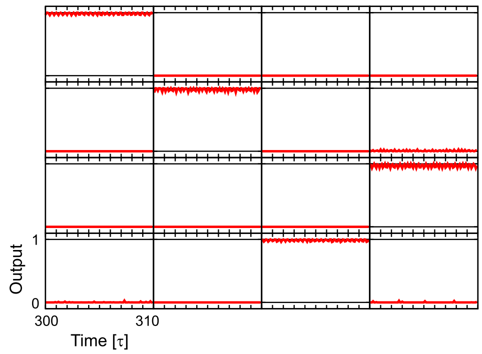

Figure 5: The output of continuous HC-ID network applied for a 4-QAP $A = 376.7$, $B = 4.68$, $C = 0.10$. One cell shows the output of one neuron.

is not always that the network state reaches to global minimum states of the QAP [4]. Figure 5 shows the output of continuous HC-ID network applied for a 4-QAP. It shows the local minimum states of a 4-QAP. If the effects of negative resistance is small, the continuous HC-ID network often cannot reach the global minimum states because the output of firing neuron oscillate around $x = 1$ as shown in Fig. 5. The output of discrete HC-ID network, meanwhile, does not take intermediate value between 0 and 1, consequently the discrete HC-ID network can avoid the oscillation state with any α.

5. Conclusion

We introduced the discrete HC-ID model instead of the continuous HC-ID model to achieve the hardware implementation. The introduced model operate in discrete time, and the output of neurons are binary. Moreover, we showed analytically that the only optimal solution states are obtained by using the discrete HC-ID network with the energy function similar to the continuous HC-ID network. Finally, we applied the discrete HC-ID network to solving a 6-TSP and a 4-QAP in numerical experiments.

References

[1] J. J. Hopfield and D. W. Tank, "“neural” computation of decisions in optimal problems", *Biol. Cybern.*, vol. 52, pp. 141–152, 1985.

[2] K. Nakajima and Y. Hayakawa, "Characteristics of inverse function delayed model for neural computation", *Proc. NOLTA'02*, pp. 861–864, 2002.

[3] Y. Hayakawa and K. Nakajima, "Design of the inverse function delayed neural network for solving combinatorial optimization problems", *IEEE Trans. on Neural Networks*, vol. 21, no. 2, pp. 224–237, 2010.

[4] T. Sota, Y. Hayakawa and K. Nakajima, "The quartic form energy function for general combinatorial optimization problems", *Proc. NOLTA'09*, pp. 527–530, 2009.