A New Stochastic Binary Neural Network Based on Hopfield Model and Its Application

Taichi Nakamura, Akio Tsuneda, and Takahiro Inoue

Department of Electrical and Computer Engineering, Kumamoto University
2–39–1 Kurokami, Kumamoto 860-8555, Japan
Tel: +81-96-342-3853, Fax: +81-96-342-3630
E-mail: taichi@st.eecs.kumamoto-u.ac.jp, tsuneda@eecs.kumamoto-u.ac.jp

Abstract: This paper presents a new stochastic binary neural network based on the Hopfield model. We apply the proposed network to TSP and compare it with other methods by computer simulations. Furthermore, we apply 2-opt to the proposed network to improve the performance.

1. Introduction

Neural networks are models of the brain's cognitive process and have a potential to realize advanced information processing. It is well known that the Hopfield neural network can be used for solving combinatorial optimization problems based on its energy function decreasing monotonically. However, in general, it is difficult to get the global minimum, which corresponds to the optimal solution, because the energy function has many local minima that we finally get starting from several initial states. Namely, it completely depends on the initial state of networks whether we can get the global minimum or not.

The Boltzmann machine is a stochastic neural network which is obtained by replacing the deterministic output function of neurons in Hopfield model with a stochastic one. Due to its stochastic behavior, the Boltzmann machine can reach the global minimum without converging to local minima. There are also attempts to add some noise in order to lead the network to the global minimum escaping from local minima [1]. Furthermore, research efforts have been devoted to chaotic neural networks [2], in which each neuron can exhibit chaotic behavior, that is, it behaves randomly in spite of its deterministic output function. It should be noted that the important common feature of the above neural networks is their stochastic or random nature.

In this paper, we propose a new stochastic binary neural network based on the Hopfield model, which are different from conventional ones [3]. We analyze characteristics of the proposed neural network concerning the performance for solving the traveling salesman problem (TSP) which is one of well-known combinatorial optimization problems. Some comparisons with other methods are also given.

2. Hopfield Neural Network

In this paper, we consider neural networks with n neurons. Here, i-th neuron is linked to the j-th neuron with a synaptic weight w_{ij}. Let $u_i(t)$ and $X_i(t)$ be the state and output of the i-th neuron at the time t in a network, respectively. Each of $u_i(t)$ and $X_i(t)$ is 1 or 0. The state of the i-th neuron is updated as

$$u_i(t+1) = \begin{cases} 1 & (if \;\; \sum_j w_{ij}X_j(t) > 0) \\ 0 & (if \;\; \sum_j w_{ij}X_j(t) < 0) \\ u_i(t) & (if \;\; \sum_j w_{ij}X_j(t) = 0). \end{cases} \quad (1)$$

In the Hopfield network, $X_i(t) = u_i(t)$ holds. The energy function in a network is defined by

$$E(t) = -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}u_i(t)u_j(t). \quad (2)$$

In the Hopfield network, it is known that energy function decreases monotonically while the state of each neuron is updated asynchronously. It enables us to solve combinatorial optimization problems. However, in general, the energy function given by eq.(2) has many local minima. Hence we cannot get the global minimum from many initial states.

3. A New Stochastic Binary Neural Network

In this paper, we propose a new stochastic binary neural network based on the Hopfield model. The state of each neuron is updated in the same way as the Hopfield model. However, the output is define by

$$X_i(t) = \begin{cases} u_i(t) & \text{(with probability } 1-p) \\ 1-u_i(t) & \text{(with probability } p), \end{cases} \quad (3)$$

as shown in Fig.1. Namely, the output is stochastic. Hence, even if we reach a local minimum, we can get out of it. Thus, we apply this model to TSP and investigate its perfomance.

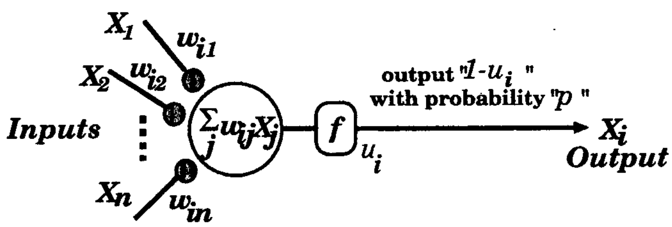


Figure 1: Proposed neuron model

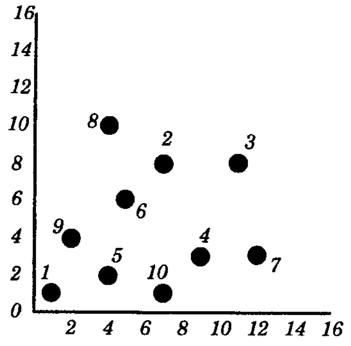
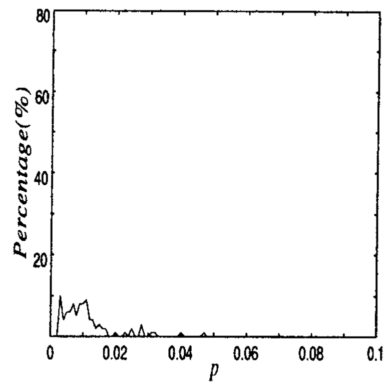
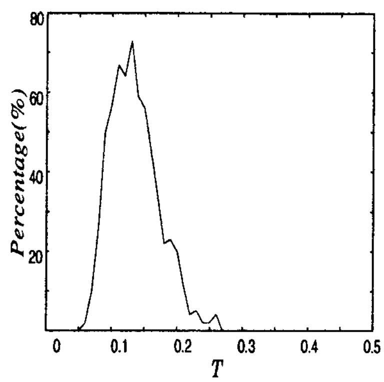


Figure 2: A 10-city model



(a)Proposed model



(b)Boltzmann machine

Figure 3: Percentages of the number of initial states which leads to the optimal solutions at least once.

4. Traveling Salesman Problem(TSP)

TSP(Traveling Salesman Problem) is one of well-known combinatorial optimization problems. In the TSP, the objective function is given by

$$\begin{aligned}\phi &= \frac{A}{2}\sum_i^N(\sum_j^N u_{ij}-1)^2+\frac{B}{2}\sum_j^N(\sum_i^N u_{ij}-1)^2\\ &+ \frac{D}{2}\sum_i^N\sum_k^N\sum_j^N d_{ik}u_{ij}(u_{k,j+1}+u_{k,j-1}), \qquad (4)\end{aligned}$$

where, u_{ij} is binary neurons, and $u_{ij} = 1(u_{ij} = 0)$ represents that city i is (not) visited at the j-th order $(i,j = 1,\ldots,N)$, d_{ik} is the distance (or traveling cost) between city i and city k, and A, B, and D are adjusting parameters related to convergence rate. In eq.(4), the first term and the second term on the righthand side denote the constraints such that

C1. The salesman visits each of all cities once.

C2. The salseman can visit only one city at once.

In this case, the energy function is given by

$$E = -\frac{1}{2}\sum_{ij}\sum_{mn} W_{ij,mn}u_{ij}u_{mn} - \sum_{ij} h_{ij}u_{ij} \qquad (5)$$

where $W_{ij,mn}$ is a weight between u_{ij} and u_{mn}, and h_{ij} is a threshold respectively given by

$$\begin{aligned}W_{ij,mn} &= -A\delta_{im}(1-\delta_{jn}) - B\delta_{jn}(1-\delta_{im})\\ &\quad - Dd_{im}(\delta_{n,j+1}+\delta_{n,j-1}), \qquad (6)\\ h_{ij} &= A + B, \qquad (7)\end{aligned}$$

where, δ_{ij} denotes the Kronecker's delta defined by

$$\delta_{ij} = \begin{cases} 1 & (i = j) \\ 0 & (i \neq j). \end{cases} \qquad (8)$$

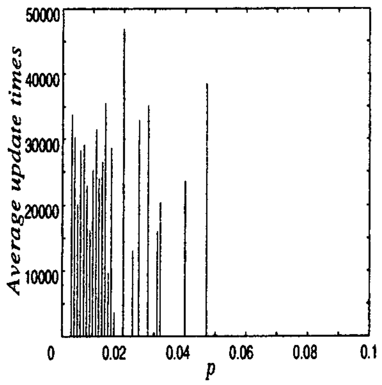
5. Computer Simulation

We take a 10-city model as an example which is shown in Fig.2. By computer simulation, we try to solve the TSP of this 10-city model using the proposed stochastic neural network. For comparison, we also solve the TSP by using the Boltzmann machine which is one of conventional stochastic neural networks.

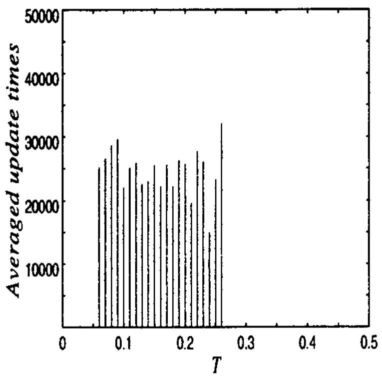
We perform such simulations changing the parameters p and T which are the probability in our model and the temperature in the Boltzmann machine, respectively. Note that each network tends to the Hopfield model as p, $T \to 0$, Simulation conditions are as follows.

- We use 100 different initial states for each p or T.
- We update states of neurons 50,000 times from each initial state.
- $A = B = 1.0$, $D = 4.4$ (determined empirically)

Fig.3 shows percentages of the number of initial states which lead to the optimal solutions at least once for each value of (a) the parameter p in the proposed network and (b) the parameter T in Boltzmann machine.



(a)Proposed model



(b)Boltzmann machine

Figure 4: Averaged update times till the network reach the optimal solution

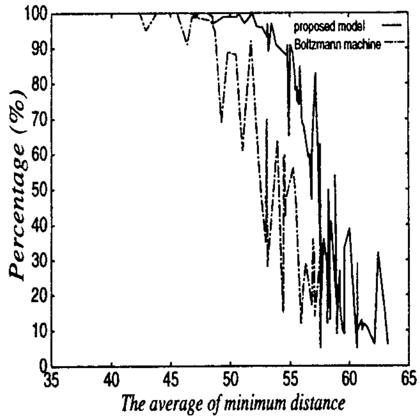


Figure 5: Percentages of initial states which lead to states satisfying constraints C1 and C2

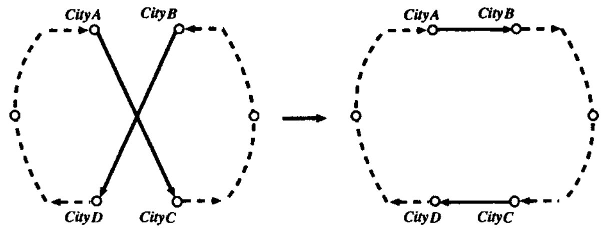


Figure 6: Illustrative description of the algorithm of 2-opt, where $d_{AC} + d_{BD} > d_{AB} + d_{CD}$.

We can find that the percentage is about 10% for the best parameter p in the proposed network and is about 74% for the best T in the Boltzmann machine.

Fig.4 shows averaged update times till the network reach the optimal solution first for each initial state which leads to optimal solutions for each value of p and T.

Futhermore, Fig.5 shows percentages of initial states which lead to states satisfying constraints C1 and C2 versus the average of the minimum distance obtained from each initial state for each p or T.

From these figures, we can find that the proposed network is inferior to the Boltzmann machine if we got the "good" parameters in advance. However, when we use "bad" parameters, the proposed network is superior to the Boltzmann machine in the sense that the constraints C1 and C2 are satisfied with higher probability.

6. Application of 2-opt

6.1 2-opt

We can improve a solution by using 2-opt [4] which can eliminate crossing branches as shown in Fig.6.

The algorithm of 2-opt as follows. Let d_{AB} be the distance between nodes (cities) A and B. Assume there are two paths $A \to C$ and $B \to D$, where $C \to \cdots \to B$. If $d_{AC} + d_{BD} > d_{AB} + d_{CD}$, we change the paths to $A \to B$ and $C \to D$, where $B \to \cdots \to C$. Thus, we can shorten the total length of the route. Applying this algorithm to every pair of paths, we can get a better solution. Generally, this algorithm starts from an initial route which is randomly given.

It is known that the average of the total length of route obtained by 2-opt tends to double of the shortest route as the number of cities increases [4]. Of course, we need enormous computation time for a large number of cities.

6.2 Applicattion of 2-opt to Stochastic Neural Network

We apply 2-opt to the proposed network in order to improve the solution. For comparison, we also apply 2-opt to the Boltzmann machine.

The algorithm is as follows.

A1. Update of state of the network continues until the constrains C1 and C2 are satisfied.

A2. We apply 2-opt to the state satisfying the constrains.

A3. After improvement by 2-opt, return to A1.

Table 1: Simulation results. (I) the percentages of initial states which leads to the optimal solution. (II) average of the number of update times until we get the optimal solution.

	Proposed	Boltzmann +2-opt	2-opt only +2-opt
(I)	99.64%	49.32%	92.56%
(II)	1702.2	1064.2	52.2

In this way, we can apply 2-opt to an initial state better than random states.

6.3 Computer Simulation

Fig.7 shows an example of city configurations which are intractable for 2-opt. For example, if we apply 2-opt to an initial route as shown in Fig.8, we never get the optimal solution as in Fig.7.

Thus, for the city configuration, we perform computer simulations of the algorithm described in the subsection 6.2. For the proposed network and the Boltzmann machine, we update their states 50,000 times from 10,000 random initial states, where the parameters are given by $p = 0.003, T = 0.13, A = B = 5.0$, and $D = 1.0$.

Table 1 shows (I) the percentages of initial states which lead to the optimal solution and (II) averages of the number of update times until we get the optimal solution. For comparison, Table 1 also shows the results obtained by 2-opt only, where 10,000 random initial routes are used.

We can find that the proposed network with 2-opt can reach the optimal solution with much higher probability than the Boltzmann machine with 2-opt. But, the average number of update times in the proposed method is larger than the Boltzmann machine.

We also find that the proposed network with 2-ot can reach the opimal solution with somewhat highter probability but its average number of update times is much larger than 2-opt only. This is considered to be a kind of trade-off. Note that we obtained similar results for several sets of parameters.

7. Conclusion

We have proposed a new stochastic binary neural network. We investigated its performance of solving the TSP, and compared it with the Boltzmann machine. As a result, we found that the proposed network is inferior to the Boltzmann machine if we have appropriate network parameters. However the proposed network can satisfy the constraints with higher probability even for bad parameters than the Boltzmann machine.

Furthermore, we showed that the proposed network with 2-opt can reach the optimal solution with higher probability than the Boltzmann machine with 2-opt. We will apply the proposed method to other various configurations with more cities in future.

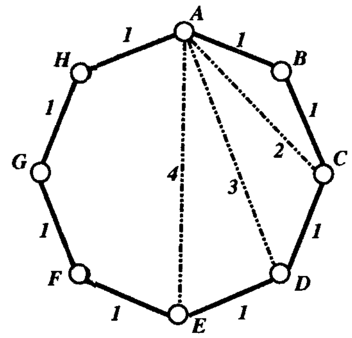


Figure 7: An example of city configurations intractable for 2-opt.

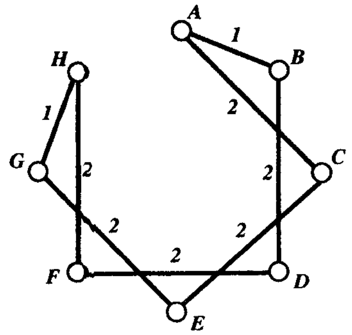


Figure 8: An example of initial routes which never reach the optimal solution in Fig.7.

References

[1] Y. Hayakawa, A. Marumoto, and Y. Sawada, "Effects of the Chaotic Noise on the Performance of a Neural Network Model for Optimization Problems," *Physical Review E*, vol.51, no.4, pp.R2693-R2696, 1995.

[2] K. Aihara, T. Takabe and M. Toyoda, "Chaotic Neural Networks," *Physics Letters A*, vol.144, no.6,7, pp.333-340, 1990.

[3] A. Tsuneda, D. Kajiwara, T. Inoue, "A New Probabilistic Neural Network and Its Characterristics", *Proc. of the 23rd Symposium on Information Theory and Its Application (SITA2000)*, pp.415-418, Kumamoto, Japan, October, 2000. (in Japanese)

[4] D. J. Rosenkrantz, R. E. Stearns and P. M. Lewis, "An analysis of several heuristics for the traveling salesman problem," *SIAM Journal on Computing* no.6, pp.563-581, 1977.