Tạp chí Tin học và Điều khiển học, T.21, S.3 (2005), 208–215

A MODIFIED META-CONTROLLED BOLTZMANN MACHINE

TRAN DUC MINH[1], LE HAI KHOI[1], JUNZO WATADA[2], TERUYUKI WATANABE[3]

[1] *Institute of Information Technology, VAST*
[2] *Graduate School of Information, Production and Systems, Waseda University, Japan*
[3] *Osaka Institute of Technology, Japan*

Abstract. The Meta-controlled Boltzmann machine was proposed by J. Watada et. al for solving the optimal quadratic programming problem. It is shown that this model converges more efficiently than a conventional Boltzmann machine. In this paper, we propose a modified version of this model and compare it with the original model in solving a quadratic programming problem, the portfolio selection problem.

Tóm tắt. Mô hình Meta-controlled Boltzmann machine được đề xuất bởi J. Watada và cộng sự nhằm giải quyết bài toán tối ưu bậc hai. Các nghiên cứu đã chứng tỏ rằng mô hình này hội tụ nhanh hơn mô hình Boltzmann chuẩn. Trong bài này, chúng tôi đề xuất mô hình mới và so sánh với mô hình cũ về hiệu năng trong việc giải một bài toán tối ưu bậc hai, bài toán đầu tư vốn.

1. INTRODUCTION

J. Watada et al proposed a model called Meta-controlled Boltzmann Machine to solve quadratic programming problems [10, 13]. Their model employed a Hopfield Network as the controller and a Boltzmann machine, which is controlled by the Hopfield layer, as the lower layer to make the whole system converge more efficient than a Conventional Boltzmann Machine. The quadratic programming problem here is understood as minimizing a quadratic function, which are the energy functions of the two layers mentioned above since in the operation, both the two layers will try to minimize their own energy functions to get to a possible lowest point of the energy surface.

The Meta-controlled Boltzmann machine operates by deleting the units, which are not selected in the Meta-controlling layer (Hopfield layer) in its execution, of the lower layer. Then the lower layer is restructured by using the selected units. Because of this feature, a Meta-controlled Boltzmann machine converges more efficiently than a conventional Boltzmann machine [11]. The key to solve a minimization problem using this model is to transform its objective function into the energy functions of the two layers since the Hopfield and Boltzmann Machines are ensured to converge at the minimum point of the energy function. The algorithm of Meta-controlled Boltzmann Machine is described in Figure 1.

However, when the inner behaviors of this model are evaluated in solving one of the most famous quadratic programming problem, a portfolio selection problem, we see that when the disturbs from the lower layer to the Meta layer is small enough, we will have a more "stable" system [8]. So we can achieve the same result without using disturb values from the lower layer to the Meta layer by replacing deterministic neurons by stochastic neurons in the Meta layer. It means that the Meta layer does not need to be disturbed by the lower layer in execution. It should be noted that the Meta layer can be changed to a Boltzmann machine-like model since a Hopfield network can also be changed to use the probabilistic update rule in units whose states can only be 0 or 1 [2, 7].

2. PORTFOLIO SELECTION PROBLEM

In 1952, H. Markowitz [5] proposed a method to allocate an amount of funds to plural stocks for investment. The method was named a portfolio selection problem. Based on time-series data of return rate, it theoretically decides the best investing rate to each of stocks, which minimizes the risk, i.e. the variance of the profits in keeping the expected return rate that a decision maker desires. It should be noted that the model can reduce the risk by means of allocating the amount of funds to many stocks. The model is excellently concise for real problems. Since then, researches have been pursued on various aspects of the model, such as realizing efficient calculation [4, 9, 12, 13].

When the selected number of stocks is limited out of the huge number of stocks, the portfolio selection problem can be formulated by a zero-one mixed-integer programming problem as described below:

FORMULATION

$$\text{Maximize } \sum_{i=1}^{n} \mu_i m_i x_i,$$

$$\text{Minimize} \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} m_i x_i m_j x_j,$$

$$\text{Subject to } \sum_{i=1}^{n} m_i x_i = 1,$$

$$\sum_{i=1}^{n} m_i = S,\ m_i \in \{0,1\},\ i = 1, ..., n,$$

where σ_{ij} denotes a covariance between stocks i and j, μ_i is an expected return rate of stock i, x_i is investing rate to stock i, n denotes the total number of stocks and S denotes the number of stocks selected, and finally, m_i denotes a selection variable of investing stocks.

The Meta-controlled Boltzmann machine mentioned above converted the objective function into the energy functions of the two components that are Meta-controlling layer (Hopfield Network [6]) and the Lower-layer (Boltzmann Machine) as described below:

Meta-controlling layer

$$E_u = -\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} S_i S_j + K_u \sum_{i=1}^{n} \mu_i S_i.$$

Lower layer

$$E_l = -\frac{1}{2} \left\{ \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} x_i x_j + 2 \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j \right\} + 2 \sum_{i=1}^{n} x_i + K_l \sum_{i=1}^{n} \mu_i x_i,$$

where K_u, K_l are weight of the expected return rate for each layer and s_i is the output value of the ith unit of the Meta-controlling layer [10, 12, 13].

3. MODIFIED META-CONTROLLED BOLTZMANN MACHINE

We should note that the change into a higher level in energy function of the Hopfield layer was made by the values disturbed from the lower layer. However, the disturb values from the lower layer made the operation of the Meta layer somewhat similar to a Boltzmann machine.

Step 1. Set each parameter to its initial value.
Step 2. Input the values of K_u and K_l.
Step 3. Execute the Meta-controlling layer.
Step 4. If the output value of a unit in the Meta-controlling layer is 1, add some amount of value to the corresponding unit in the lower layer. Execute the lower layer.
Step 5. After executing the lower layer the constant number of times, decreases the temperature.
Step 6. If the output value is sufficiently large, add a certain amount of value to the corresponding unit in the Meta-controlling layer.
Step 7. Iterate from Step 3 to Step 6 until the temperature reaches the restructuring temperature.
Step 8. Restructure the lower layer using the selected units of the Meta-controlling layer.
Step 9. Execute the lower layer until reaching at the termination.

Figure 1. Algorithm of the Meta-controlled Boltzmann Machine

Beside it, the algorithm set above will make the two layers to operate along with each other to help the Meta layer overcome the local minimum by adding some value to the unit's state. This is a good idea because the energy function's value is also depended on the units' states, thus the energy can also go up a little bit as shown in Figure 5.

However, the typical way to help the Meta layer (employed Hopfield model [6]) to overcome the local minima is to use the simulated annealing algorithm as shown in [1, 2, 3]. We can also see that the energy functions of the two layers are changed very much as well as the Meta layer's units' states. It could be overcome by replacing the deterministic neurons in the Meta layer by the stochastic neurons with probabilistic update rule. This is more similar to the Boltzmann machine model and thus, it is more typical to help the Meta layer to overcome local minima.

When the stochastic neurons are employed in the Meta layer, the energy function of this layer is retained as well as other parameters. Beside it, the purpose of the Meta-controlled layer is just to select the units (the stocks) that may give us a good solution.

Step 1. Set each parameter to its initial value.
Step 2. Input the values of *Ku* , *Kl*.
Step 3. Execute the Meta-controlling layer.
Step 4. If the output value of a unit in the Meta-controlling layer is 1, add some amount of value to the corresponding unit in the lower layer. Execute the lower layer.
Step 5. After executing the lower layer aa constant number of times, decreases the temperature.
Step 6. Iterate Step 4, 5 until the temperature reaches the restructuring temperature.
Step 7. Restructure the lower layer using the selected units of the Meta-controlling layer.
Step 8. Execute the lower layer until reaching at the termination.

Figure 2. Algorithm of the Modified Meta-controlled Boltzmann machine

So we can employ stochastic neurons in the Meta layer with a simple annealing schedule that is: $T = T_{\text{current}} \times 0.99$. And when it reaches the stop temperature, we will stop and return the result to the lower layer. After the lower layer converges, we will run the lower layer for the final time with the selected units to get the final result. It is expected to converge quickly than the original model. The algorithm should have a little bit changed as shown in Figure 2.

The reason that we still need to execute the lower layer with all the units is to make sure that it received the influence values from the Meta layer. Thus, the modified structure as well as the modified algorithm will have the same or better performance and the more correctness in refer to the value of energy function and because of the reliability of the Boltzmann Machine model.

4. SIMULATION

In this section we will show the comparison of performance between the original Meta-controlled Boltzmann machine and the modified Meta-controlled Boltzmann machine using the same parameters as shown in the following table.

Table 1. Parameter values used in experiment

Parameters and values		**Used in**
Number of stocks	15	Original net.
		Modified net.
Hop1 (Number of execution of Meta layer)	7*100	Original net.
		Modified net.
K_u	10	Original net.
		Modified net.
K_l	0.1	Original net.
		Modified net.

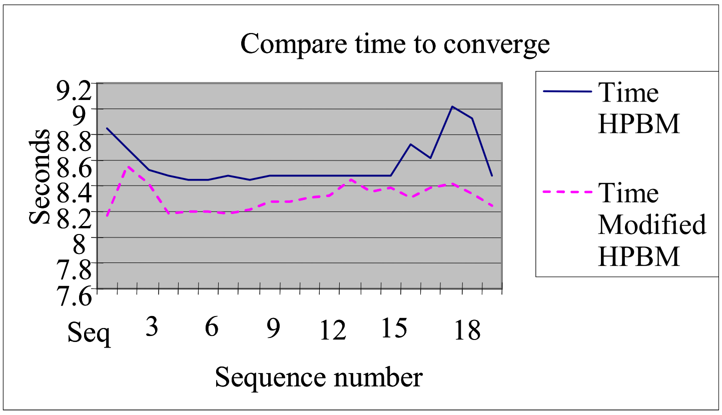

Figure 3. Compare the time to converge

The above comparison between two kinds of neural network models employed in solving the portfolio selection model is done on the time of executing the two systems as well as the energy levels that they reached. However, to get a more precise result, we should set the parameter **hop1**, which stands for the execution times of the Meta layer, in both the modified as well as the original version of Meta-controlled Boltzmann machine to at least 3 times of the total number of units in a layer since the more times the Meta-layers is executed, the more chance to get more units selected. This is what the H. Markowitz's model was designed for.

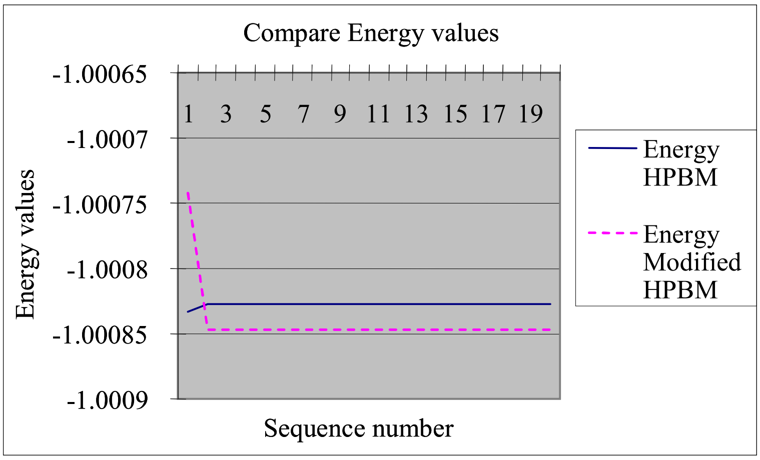

Figure 4. Compare the energy values when converged

5. CONCLUDING REMARKS

Meta-controlled Boltzmann machine behaviors

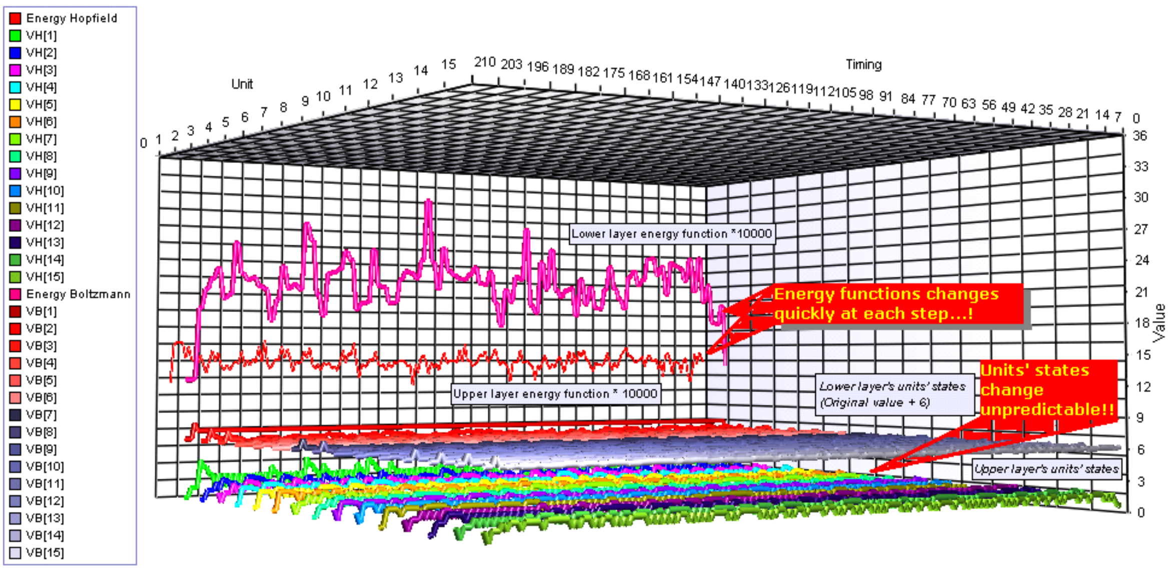

Figure 5. Inner behaviors of the Meta-controlled Boltzmann machine

In evaluating the time to converge as well as the energy level to which this model reached and comparing to the original Meta-controlled Boltzmann machine in applying in the portfolio selection problem, we can see that the performance of the modified system is better. This is caused by the new Meta layer, which employed stochastic neurons replacing the deterministic neurons. The new Meta layer removes the “discouraged units” while selects the “encouraged units” which is passed to the lower layer more efficiently that improved the performance of the whole system. Nevertheless, it is recommended that in using this model, do not just try one time to get the result but get the system run for at least 2 times or more with various values of the times to execute the Meta layer, **hop1**, because the experiments show that the energy function’s value obtained after second run may be lower than the first run (see Figure 4) as well as when the times to execute the Meta layer is increased. However, we will need time to evaluate and use this model in a wider range of problems that the conventional Boltzmann machine had been successed.

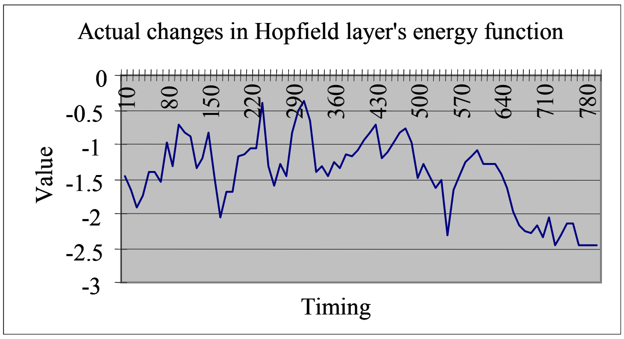

Figure 6. Actual changes in middle portion of Meta layer’s energy function (in Figure 8)

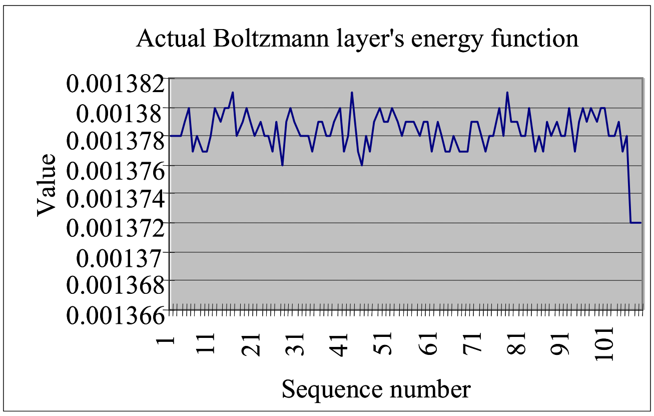

Figure 7. Actual changes in middle portion of lower layer’s energy function (in Figure 8)

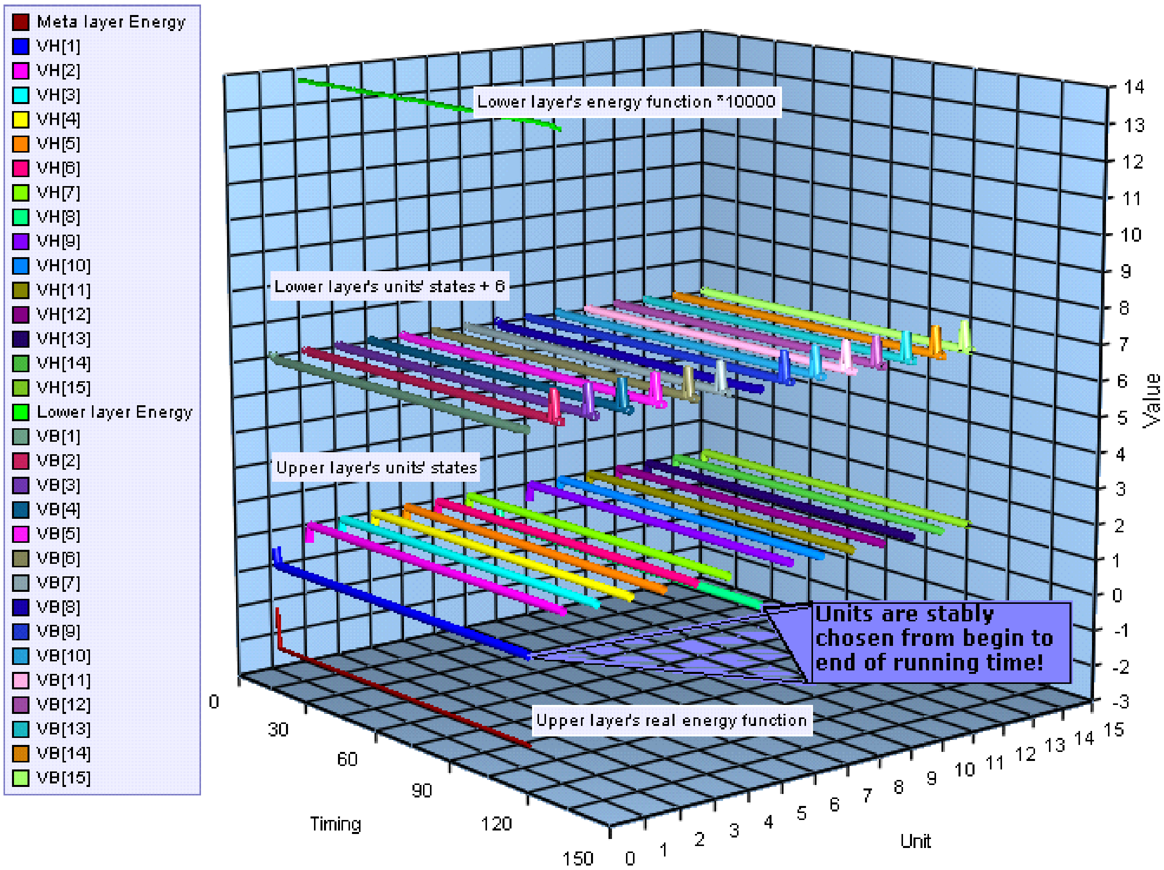

Figure 8. Modified Meta-controlled Boltzmann machine 's behaviors

REFERENCES

[1] D. H. Ackley, G. E. Hinton and T. J. Sejnowski, A learning algorithm for Boltzmann machines, *Cognitive Science* **9** (1985) 147–169.

[2] R. Bharath, J. Drosen, *Neural Network Computing*, Windcrest/McGraw-Hill, 1994.

[3] S. Haykin, *Neural Networks: A comprehensive foundation*, MacMilan Publishing Company, 1994.

[4] H. Konno, K. Suzuki, A fast algorithm for solving large scale mean-variance model by compact factorization of covariance matrices, *Journal of the Operations Research Society of Japan* **35** (1992) 93–104.

[5] H. Markowitz, Portfolio Selection, *Journal of Finance* **7** (1) (1952) 77–91.

[6] J. J. Hopfield, Neural networks and physical systems with emergent collective computational abilities, *Proceedings of National Science* (1982) 2554–2558.

[7] T. Nakamura, A. Tsuneda, and T. Inoue, A new stochastic binary neural network based on hopfield model and its application, *Proc. of International Tech. Conf. on Circuits/Systems -Computers and Communications* (2002).

[8] D. M. Tran, H. K. Le, J. Watada, T. Watanabe, "Inner behaviors of the Meta-controlled Boltzmann Machine", the 14th Soft Science Workshop, Kitakyushu, 2004 (32–37).

[9] J. Watada, "Recent Development of Soft-computing Approach to Portfolio Selection Problem," Tutorial at ECIT2002: The 2nd European Conference on Intelligent Technology, Iasi, Romania, on July 17-20, 2002 (1–22).

[10] J. Watada, K. Oda, Formulation of a two-layered Boltzmann machine for portfolio selection, *Int. Journal of Fuzzy Systems***2** (1) (2000) 39–44.

[11] J.Watada, T.Watanabe, Evaluation of Meta-Controlled Boltzmann machine and its application to re-balancing portfolio selection, *Journal of Intelligent and Fuzzy Systems* (to appear).

[12] T. Watanabe, K. Oda and J. Watada, Hierarchical decision making of strategic investment? *International Journal on Fuzziness, Uncertainty, and Knowledge-Based Reasoning (IJUFKS)* **7** (4) (1999) 429–438.

[13] T.Watanabe, J.Watada, A Meta-Controlled Boltzmann machine for rebalancing portfolio selection, *Central European Journal of Operations Research* (to appear).

Received on July 29 - 2005
Revised on October 19 - 2005