



Multi-Layer Perceptron Decided Learning Neurons by Regular Output Glia

Chihiro Ikuta[†], Yoko Uwate[†], Yoshifumi Nishio[†], and Guoan Yang[‡]

[†]Dept. of Electrical and Electronics Eng., Tokushima University
2-1 Minami-Josanjima, Tokushima, Japan

Email: {ikuta, uwate, nishio}@ee.tokushima-u.ac.jp

[‡]Dept. of Automation Science and Technology,
School of Electronic and Information Engineering, Xi'an Jiaotong University
No.28 Xianning West Road, Xi'an City, Shaanxi Province, China
Email: gayang@mail.xjtu.edu.cn

Abstract—A glia is nervous cell which is existing in a brain. This cell changes a Ca^{2+} concentration and this ion affects a neuron learning. In the biological system, when the glia does not increase the Ca^{2+} concentration, the neuron cannot increase the response. From these features, we propose a Multi-Layer Perceptron (MLP) decided learning neurons by regular output glia. The neurons are separated to some groups. Each group changes a learning term and a non-learning term. We consider that a performance of the MLP improves by having two terms. By two different simulations, we confirm a learning ability and a characteristics of the proposed MLP.

1. Introduction

Nervous cells compose a higher brain function. The nervous cell has two kinds of cells which are a neuron and a glia. The neuron has been investigated to a biological function and its applications. However, the glia was not noted, because the glia had been considered to a static cell. Recently, some researchers discovered the novel glial functions [1]. The glia uses ions concentration to a transporter of signals. The ions are a Ca^{2+} , a glutamate acid, an adenosine triphosphate, and so on [2]. Among them, we note the Ca^{2+} , because The Ca^{2+} transmits to wide range in the brain and relate closely to the neuron works.

N. Takatas reported that the Ca^{2+} affects the synaptic Long-Term Potentiation (LTP) by an experiment on living animals [3]. In that research, they used two mice. One mouse has normal glia. On the other hand, the mouse has the additional genetic defect glia which cannot become detached to the Ca^{2+} . In the genetic defect mouse, they could not observe the LTP. In the normal mouse, the researchers observed the LTP from the increase of the response. Moreover, the increase of the D-serine was observed in the increase of the Ca^{2+} concentration. The D-serine is important ion for LTP. Thereby, we can say that the glia relate closely to the LTP of the synapse.

In this study, we propose a Multi-Layer Perceptron (MLP) decided learning neurons by regular outputs glia which is inspired from function of the biological glia. In the biological system, the LTP is observed in increasing

the Ca^{2+} , moreover this ion is propagated to the wide range. We attract these features and apply these features to the MLP. The hidden-layer neurons are separated to some groups. The neurons in same group learn at same time. The learning groups are changed in rotation. The MLP learns by a Back Propagation (BP) algorithm [4]. However, the standard MLP is often falls into local minimum, because the BP uses the steepest decent method. The proposed MLP has learning neurons and non-learning neurons. We consider that the non-learning neurons become to noise generators. The non-learning neurons change to the learning neurons with time. The hidden-layer neurons repeat the change of the learning term and the non-learning term. We confirm the performance of the proposed MLP by two different tasks.

2. Proposed Method

The MLP is a famous feed forward neural network. This network is composed of layers of neurons and its outputs are tuned by the weights of connections. In general, it is learned by the BP which was proposed by D.E. Rumelhart [4]. The MLP can be applied to a pattern recognition, a pattern classification, a data mining, and so on. However, the MLP has the local minimum problem, because the BP uses the steepest decent method. For the local minimum problem, we often add a noise to any part of the MLP.

In this study, we propose the MLP decided learning neurons by regular output of glia. The glia is one of nervous cells in the brain. This cell can change the Ca^{2+} concentration which is propagated to neighboring glia. The Ca^{2+} relate closely the LTP of the synapse. In order to increase the response amplitudes of the synapse by the learning, the glia must increase the Ca^{2+} concentration [3]. In the proposed MLP, the hidden-layer neurons are separated some learning groups. The neurons in the same group learn by the BP algorithm at same time. The groups repeat the two states which are the learning term and the non-learning term. During the learning term, every neuron are learned as same as the standard MLP. During the non-learning term the weights of connections between the input-layer neurons and the hidden-layer neurons are not updated. We show an

example model of the propose MLP in Fig. 1. The connecting neuron with the excited glia can learn the weight of connection between the hidden-layer and the input-layer. The excitation state is changed to neighboring glia.

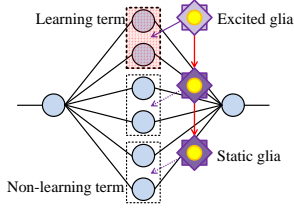


Figure 1: Proposed MLP.

2.1. Flow of the proposed MLP

We show the flow of the learning of the proposed MLP in Fig. 2. In this example, one group is composed of two neurons. Every group has same time length of the learning term and the non-learning term. The learning term and the non-learning term are changed with time. During the learning term, the neurons are learned by BP algorithm. On the other hand, the weights of connections between the hidden-layer neuron and the input-layer neurons are not updated. First group goes into the learning term. Another groups are the non-learning term. Second group goes into the learning term, then the first group remain the learning term. In the middle of the learning term of the second group, the first group goes into the non-learning term. After that, every group repeats changing the learning term and the non-learning term.

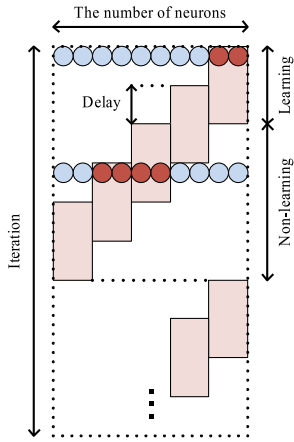


Figure 2: Flow of the learning of the proposed MLP.

2.2. Updating rule of neuron

The neuron has multi-inputs and a single output. We can change the neuron output by tuning weights of connections. The updating rule of the neuron is defined by Eq. (1).

$$y_i(t+1) = f\left(\sum_{j=1}^n w_{ij}(t)x_j(t) - \theta_i(t)\right), \quad (1)$$

where y is an output of the neuron, w is a weight of connection, x is an input of the neuron, and θ is a threshold of neuron. In this equation, the weights of the connections and the thresholds of neurons are learned by BP algorithm. We used sigmoidal function to an activating function which is described by Eq. (2).

$$f(a) = \frac{1}{1 + e^{-a}} \quad (2)$$

where a is an inner state.

3. Simulations

In this simulations, we use four kinds of the MLPs;

1. The standard MLP
2. The MLP with random noise.
3. The MLP decided learning neurons by output of glias (random sequence)
4. The MLP decided learning neurons by regular output of glias (continuous sequence)

In the MLP with random noise, we add the uniformed random noise to the thresholds of the hidden-layer neurons in the MLP. We use two kinds of the proposed MLP. The random sequence model does not decide to the group of going into the learning term. The group of going into the learning term is decided at random. In the continuous sequence model, the learning term group changes to the neighboring group. We use a Mean Square Error (MSE) an error function. The MSE is described by Eq. (3).

$$MSE = \frac{1}{N} \sum_{n=1}^N (T_n - O_n)^2, \quad (3)$$

where T is a target point and O is an output of the MLP. We confirm the performance of the MLPs by two different benchmarks. We use an approximation of a logistic map and a Two-Spiral Problem to the benchmarks. Every simulation result is obtained from 100 trials. We change the parameters which are the number of neurons in one group, the delay of changing the group of the leaning term, and a the length of the learning term.

3.1. Approximation of the chaotic function

First, we show the approximation result of the chaotic function. The chaotic function is the logistic map which is described by Eq. (4).

$$x(t+1) = \alpha x(t)\{1 - x(t)\}, \quad (4)$$

where α is an operation coefficient of the generation pattern of the time series. In this simulation, we use $\alpha = 4$. Then, the logistic map becomes fully-developed chaos. The number of neurons in the MLP are arranged of 1-12-1 and 1 trial

has 10000 iterations. The MLP is inputed $x(t)$ and learns $x(t + 1)$. We change the parameters of the proposed MLP and obtain the results.

Figures 3 (a)-(c) show results for each condition. In this simulation, the proposed MLP is better than the standard MLP and the MLP with random noise in every condition. Moreover, the MLP with random noise is worse than the standard MLP. The standard MLP has the good performance when the simulation task is easy. Because the percentage of falling into the local minimum become small. The noise impede the learning of the MLP in the easy task. The proposed MLP can decrease error to under the standard MLP result. Even if the additional noise is not efficiency case, the proposed MLP has good performance. When the number of the neurons in the group become worse, the average of error is increased. And also, when the time length of the learning term is longer, the average of error become larger. We consider that the neurons has better performance when each neuron separately learn. Because, we consider that this simulation task is easy for the MLP and we give too many number of the hidden-layer neurons.

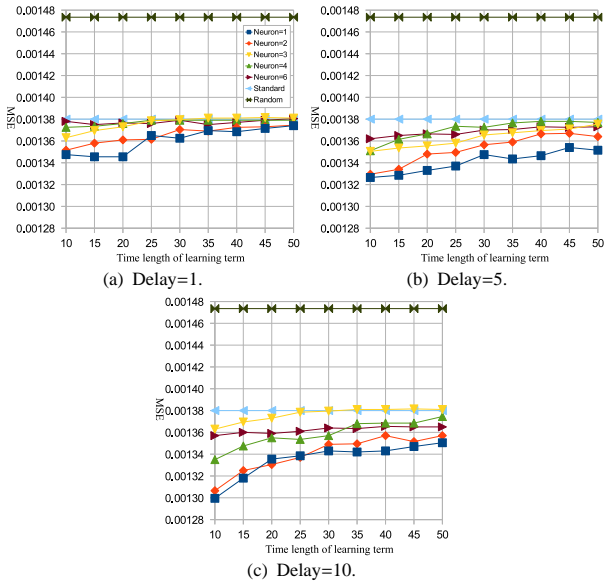


Figure 3: Approximation results of random sequence.

Next, we show the continuous sequence case in Fig. 4. In these results, the trend of the graphs are similar to the random sequence case. However, the difference of the number of neurons is more clearly.

3.2. Two-spirals problem

Next, we use the TSP to the benchmark. The TSP is a famous benchmark for the artificial neural network. It is known to the high non-linearity problem. In this benchmark, the MLP learns the coordinates of two different spirals. After learning, the MLP classifies the unknown coordinates to either one of the two-spirals. In this simulation, the number of neurons in the MLP are arranged of 2-40-1 and the iterations are 100000.

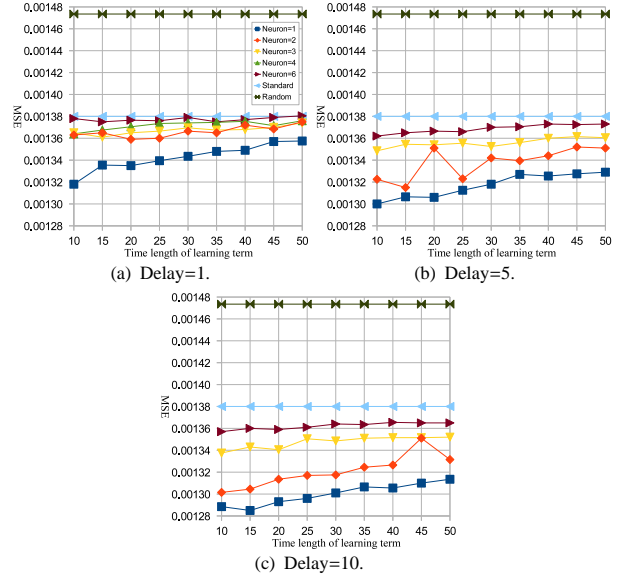


Figure 4: Approximation results of continuous sequence.

First, we show the approximation performances for learning the two-spirals in Figs 5 and 6. These results show that the MLP with random noise is better than the standard MLP. Thus, the standard MLP falls into the local minimum in the simulation. In the proposed MLP, when the number of learning term neurons are smaller, the average of errors are worse than the results of the MLP with random noise. However, when the number of learning neurons as same time are large, the approximation performances improve. From these results, we can say that the neurons are learned at same time when the MLP has difficult task.

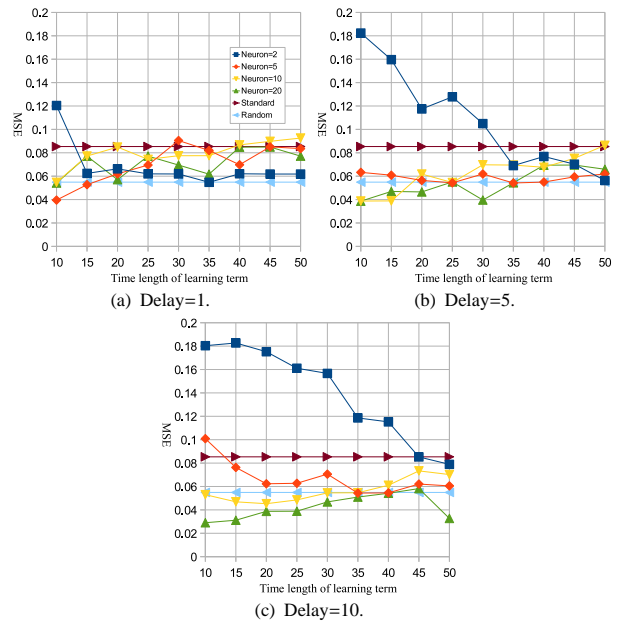


Figure 5: Approximation results of random sequence.

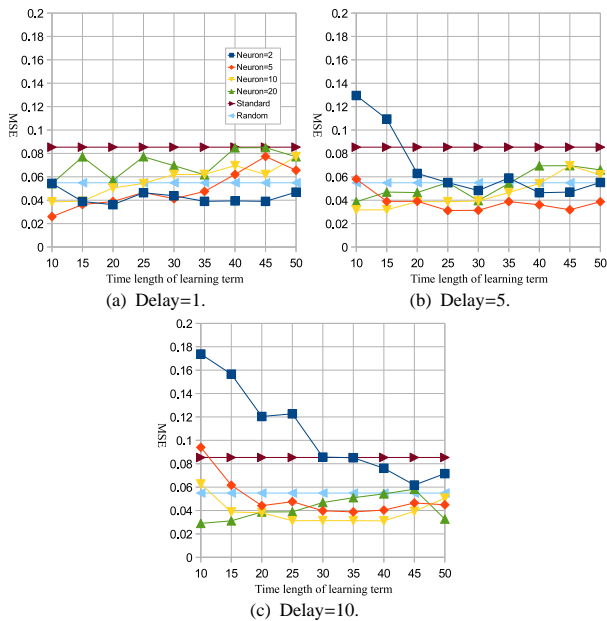


Figure 6: Approximation results of continuous sequence.

Finally, we show the classification performances in Figs. 7 and 8. The approximation results are sometime over the average of the standard MLP. However, all classification results are under the average of the standard MLP. Thereby, the proposed MLP has a better generalization capability than the standard MLP. Generally, the MLP with random noise has better generalization capability than the standard MLP. The generalization capability of the proposed MLP is better than the MLP with random noise in some conditions. Thus, we consider that the two states term give the effect of similar to noise doping.

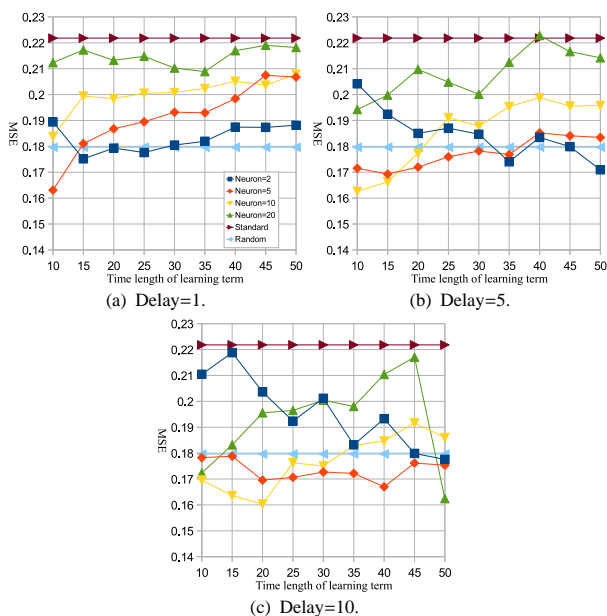


Figure 7: Classification results of random sequence.

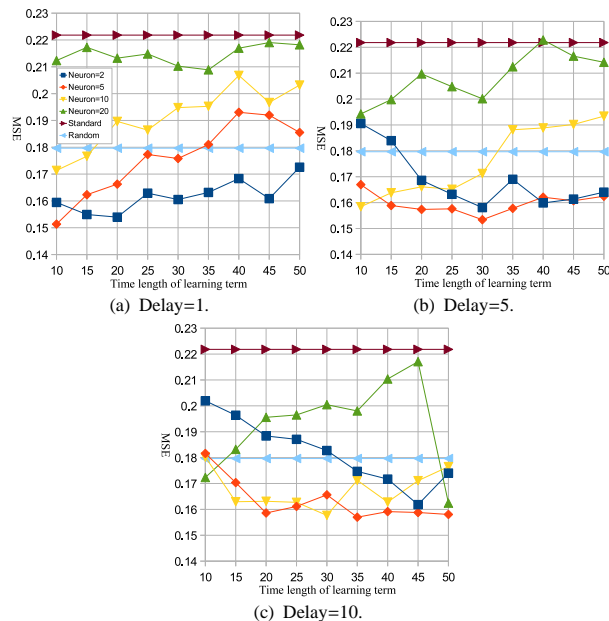


Figure 8: Classification results of continuous sequence.

4. Conclusions

In this study, we have proposed the MLP decided learning neurons by regular output of glia. We added the biological feature of glia to the MLP. In this model, the hidden-layer neurons have two different states which are the learning term and the non-learning term. The hidden-layer neurons are separated to some group. The neurons in the group are learned at same time. The learning group is changed in the rotation. By the simulation, we showed that the changing two terms gives the good influence to the MLP. It improved the approximation performance and the classification performance. Moreover, we confirmed the characteristic of the proposed MLP.

Acknowledgment

This work was partly supported by MEXT/JSPS Grant-in-Aid for JSPS Fellows (24-10018).

References

- [1] P.G. Haydon, "Glia: Listening and Talking to the Synapse," *Nature Reviews Neuroscience*, vol. 2, pp. 844-847, 2001.
- [2] S. Ozawa, "Role of Glutamate Transporters in Excitatory Synapses in Cerebellar Purkinje Cells," *Brain and Nerve*, vol. 59, pp. 669-676, 2007.
- [3] N. Takata, T. Mishima, C. Hisatsune, T. Nagai, E. Ebisui, K. Mikoshiba, and H. Hirase, "Astrocyte Calcium Signaling Transforms Cholinergic Modulation to Cortical Plasticity In Vivo," *The Journal of Neuroscience*, vol. 31, pp. 18155-18165, Dec. 2011.
- [4] D.E. Rumelhart, G.E. Hinton and R.J. Williams, "Learning Representations by Back-Propagating Errors," *Nature*, vol. 323-9, pp. 533-536, 1986.