2008 International Symposium on Nonlinear Theory and its Applications
NOLTA'08, Budapest, Hungary, September 7-10, 2008

Lazy Self-Organizing Map Considering Lazy-Neuron Rate for Effective Self-Organization

Taku Haraguchi†, Haruna Matsushita† and Yoshifumi Nishio†

†Tokushima University, 2-1 Minami-Josanjima, Tokushima, Japan
Email: {taku, haruna, nishio}@ee.tokushima-u.ac.jp

Abstract—In the previous study, we have proposed the Lazy Self-Organizing Map (LSOM), whose learning rate depends on only each neuron's character, and investigated its behaviors. In this study, we propose an improved LSOM, whose learning rate depends on each neuron's character, lazy-neuron rate and time. We apply the improved LSOM considering lazy-neuron rate to various input data set and investigate its effectiveness using three measurements.

1. Introduction

In data mining, clustering is one of typical analysis techniques and is studied for many applications, such as a statement, a pattern recognition, a image analysis and so on. Then, the Self-Organizing Map (SOM) [1] has attracted attention for the study on clustering [2] in recent years. SOM is an unsupervised neural network introduced by Kohonen in 1982 and is a simplified model of the self-organization process of the brain. SOM can retain topological feature, which is association between neurons, and according to its advantage, SOM can classify similar data. Meanwhile, in the world, the amount and the complexity of data increase from year to year. Therefore, it is important to classify various data exactly.

In the previous study, we have applied ant world to the conventional SOM. There is a report that about 20% of worker ants are "lazy" [3], furthermore, there also is another report that the ants group, which contains the lazy ants at food collections, can collect more foods than the group which contains only the worker ants. From these reports, we have proposed a new type of SOM algorithm, which is called Lazy SOM (LSOM) algorithm [4]. The important feature of LSOM is that three kinds of neurons exist; *worker neurons*, *lazy neurons*, which do not work, and *indecisive neurons* which are the neighborhoods of the lazy neurons. The learning rate of the lazy neurons is smaller than that of the worker neurons. The learning rate of the indecisive neurons becomes small due to the lazy neurons. The learning rate of the previous LSOM depends on each neuron's character. For this reason, the previous LSOM can obtain the map reflecting the distribution state of the input data more effectively than the conventional SOM, however, it tend to obtain a strongly twist map.

In this study, we propose an improved LSOM resembling the learning rate of the conventional SOM in order to carry out more exact self-organization than the previous LSOM. We investigate efficacy of the lazy-neuron rate of the improved LSOM and apply it to various input data set. We confirm that the improved LSOM containing the lazy neurons, which is from 10% to 20% of the total, can obtain the most effective and exact map reflecting the distribution state of the input data than the conventional SOM.

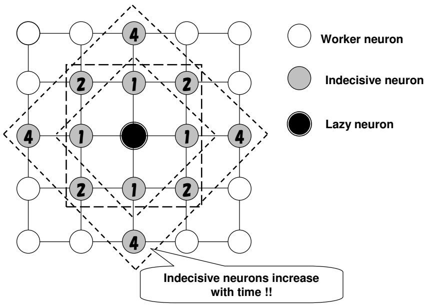


Figure 1: LSOM contains three kinds of neurons: *worker neuron*, *lazy neuron* and *indecisive neuron*. Each number denotes neighborhood distance between lazy neuron and each neuron.

2. Lazy Self-Organizing Map (LSOM)

In our previous research, we have proposed LSOM containing three kinds of neurons (as Fig. 1): *worker neurons*, *lazy neurons*, which do not work, and *indecisive neurons* which are neighborhoods of the lazy neuron. However, the previous LSOM has a defect. The learning rate of the previous LSOM depends on each neuron's character. For this reason, the previous LSOM can obtain the map reflecting the distribution state of the input data more effectively than the conventional SOM, however, it tend to obtain a strongly twist map. In this study, we propose that the improved LSOM which can retain association between neurons with the feature of the previous LSOM remaining. In consequently, the learning rate of the improved LSOM is decided by each neuron's character, a lazy-neuron rate, which is the percentage of lazy neurons of the total, and time. It means that the improved LSOM keeps features of the conventional SOM than the previous LSOM.

We explain the learning algorithm of the improved LSOM in detail. LSOM has a two-layer structure of the input layer and the competitive layer as the conventional SOM. In the input layer, there are d-dimensional input vectors $\boldsymbol{x}_j = (x_{j1}, x_{j2}, \cdots, x_{jd})$ $(j = 1, 2, \cdots, N)$. In the competitive layer, M neurons are arranged as a regular 2-dimensional grid. Each neuron has a weight vectors $\boldsymbol{w}_i = (w_{i1}, w_{i2}, \cdots, w_{id})$ $(i = 1, 2, \cdots, M)$ with the same dimension as the input vector. $(p \times M)$ neurons are classified into a set of the lazy neurons S_{lazy} at random. i.e., p denotes lazy-neuron rate.

(LSOM1) An input data $\boldsymbol{x}_j$ is inputted to all the neurons at the same time in parallel.
(LSOM2) We find a winner neuron by calculating the distances between the input vector $\boldsymbol{x}_j$ and the weight vector $\boldsymbol{w}_i$

of the neuron i. The winner neuron c is the neuron with the weight vector nearest to the input vector $\boldsymbol{x}_j$;

$$c = \arg\min_i \{\|\boldsymbol{w}_i - \boldsymbol{x}_j\|\}, \tag{1}$$

where $\|\cdot\|$ is the distance measure, in this study, we use Euclidean distance. If $c \in S_{\text{lazy}}$, we perform (LSOM3). If not, we perform (LSOM4).

(LSOM3) The lazy neuron, which is the winner c, is excluded from the set of the lazy neuron S_{lazy}, and a neuron f is selected to become a member of S_{lazy}. f is farthest from the input data $\boldsymbol{x}_j$ and is not in S_{lazy};

$$f = \arg\max_i \{\|\boldsymbol{w}_i - \boldsymbol{x}_j\|\}, \quad i \notin S_{\text{lazy}}. \tag{2}$$

In other words, the lazy neuron becomes worker if it becomes the winner c, and another neuron f becomes lazy.

(LSOM4) We find the indecisive neurons. A set of the indecisive neurons N_{lazy} are the neighborhoods of each lazy neuron l in S_{lazy}.

$$N_{\text{lazy}} = \{i \mid \quad \|\boldsymbol{r}_i - \boldsymbol{r}_l\|^2 \leq D(t), \; i \neq c, \; i \notin S_{\text{lazy}}, \; l \in S_{\text{lazy}}\}, \tag{3}$$

where $\|\boldsymbol{r}_i - \boldsymbol{r}_l\|$ is the neighborhood distance between map nodes i and l on the map grid, and $D(t)$ corresponds to the neighborhood size. $D(t)$ increases with time according to the following equation;

$$D(t) = \left[D_{\max} \frac{t}{T} \right], \tag{4}$$

where $[\cdot]$ denotes the Gauss' notation and $D_{\max}$ is a fixed parameter deciding the max value of $D(t)$, t is the learning step, T is the maximum number of the learning.

(LSOM5) The weight vectors of all the neurons are updated as

$$\boldsymbol{w}_i(t+1) = \boldsymbol{w}_i(t) + h_{Lc,i}(t)(\boldsymbol{x}_j - \boldsymbol{w}_i(t)), \tag{5}$$

where $h_{Lc,i}(t)$ is the neighborhood function of LSOM as

$$h_{Lc,i}(t) = \alpha(t) \exp\left(-\frac{\|\boldsymbol{r}_i - \boldsymbol{r}_c\|^2}{2\sigma^2(t)} \right), \tag{6}$$

where $\alpha(t)$ is the learning rate of the improved LSOM and is decided by each neuron's character, the lazy-neuron rate p and time;

$$\alpha(t) = \begin{cases} \alpha_{\text{lazy}} \left(1 - p\frac{t}{T}\right), & \text{if } i = l, \; l \in S_{\text{lazy}} \\ \alpha_{\text{N}} \left(1 - (1-p)\frac{t}{T}\right), & \text{if } i \in N_{\text{lazy}} \\ \alpha_{\text{w}} \left(1 - (1-p)\frac{t}{T}\right), & \text{otherwise,} \end{cases} \tag{7}$$

where α_{w} is the learning rate of the worker neurons, α_{lazy} is the learning rate of the lazy neurons and α_{N} is the learning rate of the indecisive neurons, namely $\alpha_{\text{lazy}} \leq \alpha_{\text{N}} \leq \alpha_{\text{w}}$.

(LSOM6) The steps from (LSOM1) to (LSOM5) are repeated for all the input data.

3. Experimental Results

3.1. For Target data

We consider a 2-dimensional input data: Target data set shown in Fig. 2(a). This data contains 770 points which has a clustering problem of outliers [5]. The conventional SOM and the previous LSOM and the improved LSOM have 100 neurons (10×10), respectively. 10% neurons of the total, namely $p = 0.1$, are classified into a set of the lazy neurons. We repeat the learning 20 times for all input data, namely $T = 15400$. The parameters for the learning are chosen as follows;

$$\alpha_{\text{w}} = 0.5, \; \alpha_{\text{lazy}} = 0.05, \; \alpha_{\text{N}} = 0.25$$

$$\alpha(0) = 0.5, \; \sigma(0) = 4.0, \; D_{\max} = 5.$$

where we use $\alpha_{\text{w}}(t)$ corresponding to learning rate of the conventional SOM and $\sigma(0)$ corresponding to width of neighborhood function of the conventional SOM for the comparison and the confirmation of the lazy neurons effect.

The simulation results of the conventional SOM, the previous LSOM and the improved LSOM are shown in Figs. 2(b), (c) and (d), respectively. In Fig. 2(b), we can see that the conventional SOM does not self-organize up to all the outliers of the input data. In Fig. 2(c), even though the previous LSOM can self-organize up to all the corner data, it has some twists. The twists are the phenomenon that the association between the neurons do not retain, and they produce a false recognition that the distance between the neurons is near though it is far actually. Besides, the twists become a cause that we can not extract the exact data reflecting the topology of the input data in the extraction. In Fig. 2(d), the improved LSOM can self-organize up to all the corner data than the conventional SOM and does not have the twists. This results are because the following reasons. The worker neurons tend to gather at the area where the input data are concentrated in easily, the indecisive neurons and the lazy neurons can discover the outside area easily. Furthermore, the lazy neurons are hardly updated at all and have the feature of (LSOM3) in Section 2. For these features, the improved LSOM can self-organize outliers more effectively than the conventional SOM. Additionally, a reason why the improved LSOM does not have the twists is that the learning rate of the improved LSOM resembles the conventional SOM than the previous LSOM whose learning rate does not contain time. For these reasons, we can confirm the improved LSOM effectiveness.

Furthermore, in order to compare the learning performance of LSOM with other methods numerically, we use the following well-used three measurements.

Quantization Error Qe [1]: This measures the average distance between each input vector and its winner. Thus, Qe near 0 is more desirable.

Topographic Error T_e [6]: This describes how well the SOM preserves the topology of the the learned data set. Thus, Te near 0 is more desirable.

Neuron Utilization U [7]: This measures is the percentage of neurons that are the winner of one or more input vector in the map. Thus, U near 1 is more desirable.

The calculated measures are shown in Table 1. The quantization error Qe of the previous LSOM is the smallest value in three algorithms and the neuron utilization U of the previous LSOM is the largest value. However, the topographic error Te of the previous LSOM is the largest value, and by using the previous LSOM, Te becomes worse 128.4672% from using the conventional SOM. Unlike the quantization error Qe, Te considers the structure of the map. For a strangely twisted map, the topographic error is big even if the quantization error is small. Namely, no matter what Qe and U have improved from using the conventional SOM, if Te becomes worse, its results are not effective. Meanwhile, although the neuron utilization U of the improved LSOM is the same as the conventional SOM, the quantization error Qe of the improved LSOM

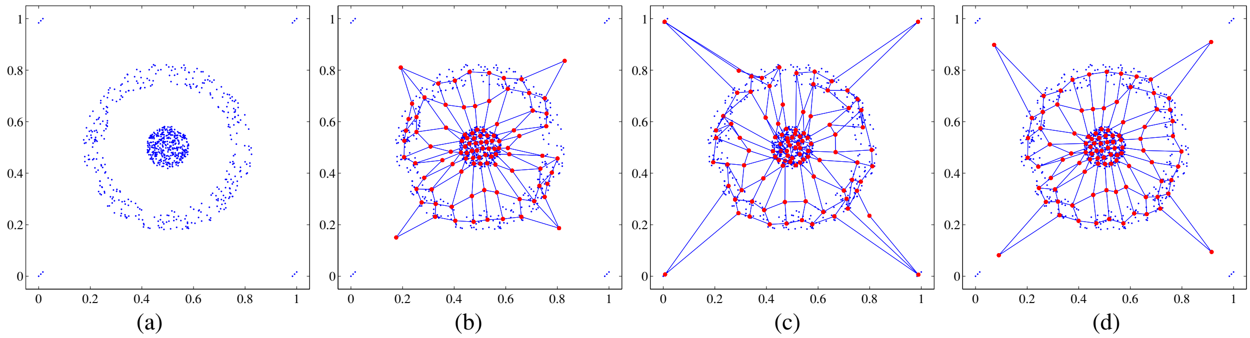


Figure 2: Simulation results for Target data. (a) Input Data. (b) Conventional SOM. (c) Previous LSOM. (d) LSOM.

is smaller, and by using the improved LSOM, Qe has improved by 13.6970% than the conventional SOM. Furthermore, the topographic error Te of the improved LSOM is smaller than the conventional SOM, and by using the improved LSOM, Te has improved by 54.7445%. For these results, we can see that the improved LSOM is effective than the previous LSOM and obtains the best results.

Table 1: Quantization error Qe, Topographic error Te and Neuron utilization U for Target data.

	Qe	Te	U
Conventional SOM	0.0190	0.1779	0.8100
Previous LSOM	0.0144	0.4065	0.8300
LSOM	0.0164	0.0805	0.8100
Improvement rate of LSOM from SOM	13.6970%	54.7445%	0%

Figure 3 shows distances between neighboring neurons and visualizes the cluster structure of the map. Black squares on this figure mean large distance between neighboring map nodes. Clusters are typically uniform areas of white squares. Each map is normalized by each distance.

From Fig. 3(b), although the outliers of the previous LSOM is clear, it is difficult to find the configuration of the input data. Meanwhile, the improved LSOM can easily confirm the outliers and the configuration of clusters, as it is clustering the input data exactly. Namely, the improved LSOM obtain the most exact map reflecting the configuration of clusters and the phase space of the input data.

3.2. For Atom data

Next, we consider a 3-dimensional input data called Atom data set shown in Fig. 4(a) [5]. This data set has clustering problems which is not linearly separable, namely, different densities and variances. The total number of the input data N is 800, and the input data has two clusters. We repeated the learning 20 times for all input data, namely T = 16000. The learning conditions are the same as those in Subsection 3.1.

The learning results of the conventional SOM, the previous LSOM and the improved LSOM are shown in Figs. 4(b), (c) and (d), respectively. For these results, we see that the previous LSOM can self-organizes dispersed input data compared with the conventional SOM, however, it can not retain association between neurons. Meanwhile, the improved LSOM can self-organizes dispersed input data compared with the conventional SOM, furthermore, it retains exactly retain association between neurons.

Table 2: Quantization error Qe, Topographic error Te and Neuron utilization U for Atom data.

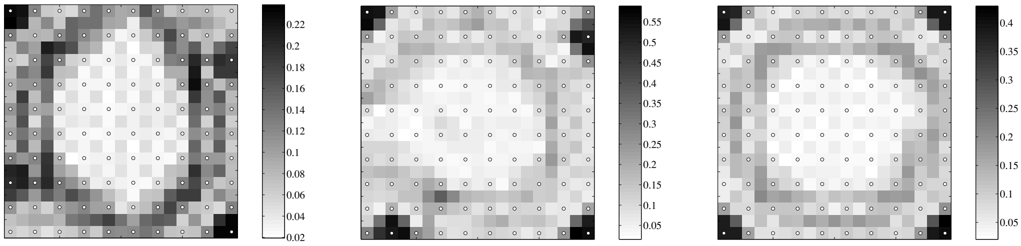
	Qe	Te	U
Conventional SOM	0.0558	0.2725	0.8300
Previous LSOM	0.0509	0.4065	0.8200
LSOM	0.0483	0.2537	0.8500
Improvement rate of LSOM from SOM	13.4116%	6.8807%	2.4096%

Table 2 shows the calculated measures. The quantization error Qe of the improved LSOM is the smallest value in the three algorithms and the neuron utilization U of the improved LSOM is also the best value. Furthermore, by using the improved LSOM, Qe has improved 13.4116% from using the conventional SOM and U has improved 2.4096%. This means that there are few errors between the input data and the neurons of learned map, in other words, the improved LSOM can be most reflecting the distribution state of the input data. Besides, 85% of the neurons of the improved LSOM are winner for one or more input data, namely, there are few inactive neurons. We consider these obtained results. The inactive neurons are on a part without the input data. Therefore, if the inactive neurons are few, more neurons can self-organize the input data of clusters. In consequence, the distance between input data and its neurons will be small respectively. Furthermore, unlike the conventional SOM, the learning rate of the improved LSOM does not decay to zero. Namely, in the closing stage of learning, neurons of the improved LSOM can update. For these reasons, the quantization error Qe will be a small value. Meanwhile, the topographic error Te of the improved LSOM is the smallest value in the three algorithms. Furthermore, by using the improved LSOM, Te has improved 6.8807% from using the conventional SOM. This is because unlike the previous LSOM, which does not depend on time, the learning rate of the improved LSOM resembles it of the conventional SOM. For these results, the improved LSOM obtain the best results.

4. Behaviors of LSOM Considering Lazy-Neuron Rate

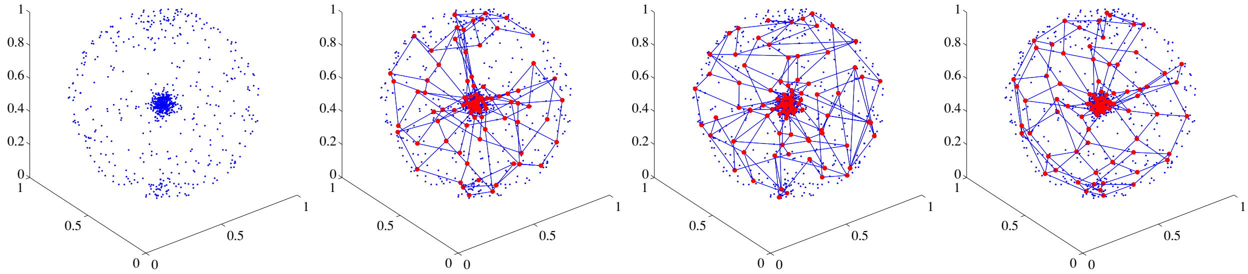
We apply LSOM to Target data set with changing the lazy rate p from 0.2 to 0.9. The learning conditions are the same as those in Subsection 3.1.

Figure 5 shows the improvement rate of LSOM using the conventional SOM when carrying out the simulation



(a) (b) (c)

Figure 3: Visualization of result for Target data. (a) Conventional SOM. (b) Previous LSOM. (c) LSOM.



(a) (b) (c) (d)

Figure 4: Simulation results for Atom data. (a) Input data. (b) Conventional SOM. (c) Previous LSOM. (d) LSOM.

with difference the lazy-neuron rate p. The quantization error Qe of LSOM has improved significantly than the conventional SOM regardless of the lazy-neuron rate p. The topographic error Te of LSOM containing 10% the lazy neurons has improved significantly and containing 20% the lazy neurons has improved slightly from using SOM. However, as the lazy-neuron rate p increase, the topographic error Te becomes worse. In addition, when LSOM contains from 50 to 90 % lazy neurons, the neuron utilization U becomes worse slightly than SOM.

These results mean that the more the lazy-neuron rate p increases, the more the association between neurons does not retain. Namely, the improved LSOM containing from 10% to 20% lazy neurons can retain the association between neurons and obtains the effective and exact map reflecting the distribution state of the input data. In other words, we can comprehensively evaluate that its condition is the best balance.

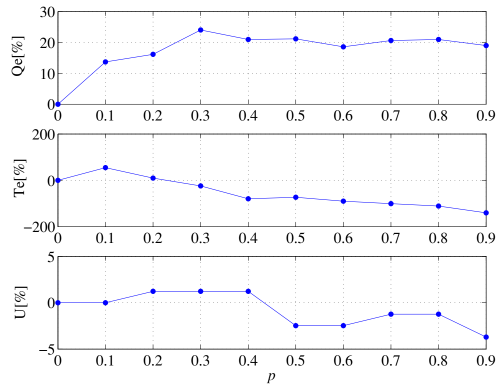


Figure 5: Lazy-neuron rate p versus improvement rate [%] of LSOM from SOM Quantization error Qe, Topographic error Te and Neuron utilization U.

5. Conclusions

In this study, We have proposed the improved Lazy Self-Organizing Map (LSOM). The learning rate of the improved LSOM is decided by each neuron's character, lazy-neuron rate and time. We have investigated efficacy of the new algorithm and lazy-neuron rate of the improved LSOM by applying these to some input data set. In consequence, we can say that the improved LSOM containing from 10% to 20% lazy neurons can obtains more effective and exact map reflecting the distribution state of the input data.

References

[1] T. Kohonen, *Self-Organizing Maps*, Berlin, Springer, vol. 30, 1995.

[2] J. Vesanto and E. Alhoniemi, "Clustering of the Self-Organizing Map," *IEEE Trans. Neural Networks*, vol. 11, no. 3, pp. 586–600, 2002.

[3] H. Hasegawa, "Optimization of GROUP Behavior," Japan Ethological Society Newsletter, no. 43, pp. 22–23, 2004.

[4] T. Haraguchi, H. Matsushita and Y. Nishio, "Lazy Self-Organizing Map and its Behaviors, Proc. of IJCNN'08, pp.2276-2281, June 2008 .

[5] A. Ultsch, "Clustering with SOM: U*C," Proc. Workshop on Self-Organizing Maps, pp.75-82, 2005.

[6] K. Kiviluoto, "Topology Preservation in Self-Organizing Maps," Proc. of International Conference on Neural Networks, pp. 294-299, 1996.

[7] Y. Cheung and L. Law, "Rival-Model Penalized Self-Organizing Map," *IEEE Trans. Neural Networks*, vol. 18, no. 1, pp. 289–295, 2007.