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

Efficiency of Statistical Measures to Estimate Network Structure of Chaos Coupled Systems

Yuta Ueoka*, Tomoya Suzuki*, Tohru Ikeguchi† and Yoshihiko Horio‡
*Department of Information System Design, Faculty of Science and Engineering, Doshisha University
1-3 Tatara-Miyakodani, Kyotanabe-shi, Kyoto 610-0321, Japan
†Department of Information and Computer Sciences School of Engineering, Saitama University
225 Shimo-Ohkubo, Saitama, 338-8570, Japan
‡Department of Electrical Engineering, Tokyo Denki University
2-2 Kanda-Nishiki-cho, Chiyoda-ku, Tokyo 101-8457, Japan

***Abstract*— Real systems often show complex behavior due to interaction among many elements and compose networks. To model and predict these systems, we estimate network structures by using only time-series data as behavior of systems. Because elements interacting each other show similar behavior, it is useful to apply the correlation coefficient, the partial correlation coefficient, the mutual information, and the transfer entropy. However, according to network topology and strength of interaction among nodes, the efficiency of these statistical measures and the optimum measure must be different. In this study, we constitute chaos coupled systems and examine the efficiency of these measures, changing model parameters. Moreover, we discuss given results on the basis of the degree of synchronization among elements and the Lyapunov exponents of system.**

I. Introduction

There are many nonlinear dynamical systems in the real world. These systems often constitute networks, for example, neural networks, stock markets, chain of earthquake, etc., whose elements interact each other and show complex behavior. To analyze such complex systems, it is important to estimate network structures only by observed behavior from networks. If we know network structures, we can model systems more accurately and can use the model to predict future behavior. As behavior of a network, we obtain time-series data observed in each node. Because connected nodes interact each other, each behavior might be similar. We consider that it is possible to estimate network structure by estimating the similarity by statistical measures, such as the correlation coefficient, the partial correlation coefficient, the mutual information, and the transfer entropy[1]. The correlation coefficient is suitable to estimate linear relationship. However, there is the problem of false correlation in clustered network. The partial correlation is useful for reducing false correlation. On the other hand, the mutual information and the transfer entropy are suitable to estimate nonlinear relationship. However, the mutual information has the similar problem to the correlation coefficient. In our study, to examine each optimum measure according to network structures, we perform some simulations by using the Watts and Strogatz (WS) model[2]. The WS model can make various network structures from the regular network to the small-world network and the random network. Moreover, we compose chaos coupled systems[3] having various network topology made by the WS model, and drive temporal behavior of the systems as time-series data observed in each node. Then, we estimate network structures by using above statistical measures.

II. Chaos Coupled Systems

Using the WS model proposed by D. Watts ans S. Strogatz[2], we can compose a network whose topology can be simply changed with a parameter. First, we prepare the regular network where all nodes are connected only locally. Although the network has many clustered nodes, the distance between nodes is very large. If we randomly rewire all edges of the network, the network becomes the random network where local clusters are destroyed and the distance between nodes becomes short. Then, according to the rewiring probability p, we can realize the small-world network, which is an intermediate graph between the regular network and the random network, and realize large cluster and small distance between nodes, simultaneously. If we set $p = 0$, the topology of the WS model is the regular network. If we set $p = 1$, the topology of the WS model becomes the random network. By setting $0 < p < 1$, we can realize the small-world network.

As a numerical model of complex systems, the coupled map lattice was proposed by K. Kaneko[3]. However, each element of the model interacts only locally like the regular network. In this study, to change the topology of interaction, we modify the coupled map lattice by using the WS model. Namely, in the modified coupled map lattice, each element interacts in the regular network, the random network or the small-world network.

This model is defined by

$$x_{t+1}^i = f\left((1-\epsilon)x_t^i + \frac{\epsilon}{N_i}\sum_{j\in G_i(p)}^{N_i} x_t^j\right) \tag{1}$$

where x_t^i is a time-series value of the i-th node, and N_i is the number of nodes connected with the i-th node. These nodes are represented by j. The $G_i(p)$ is a set of node j connected with the i-th node and depends on network structure decided by the rewiring probability p. The strength of interaction in $G_i(p)$ is denoted by ϵ. Then, as a dynamics in the coupled map lattice, we use the logistic map:

$$f(x_t) = 1 - ax_t^2, \tag{2}$$

which is widely analyzed as a sort of chaotic maps.

III. Analyses of Network Behavior

In this section, we examine behavior of the chaos coupled system introduced in Sec.2. We set the number of nodes n to 100, and set the length of each node's time-series data to $2,000$ after transient states of $3,000$ times.

We calculate the degree of synchronization in the chaos coupled system by

$$S = \frac{1}{{}_n\mathrm{P}_2} \sum_{i \neq j} |C_{i,j}| \tag{3}$$

where $C_{i,j}$ is the correlation coefficient between two time series x^i and x^j.

Figure 1 shows the results of S. If p is small, that is, the topology of network becomes the regular graph, the degree of synchronization is small. However, if p becomes larger, that is, the topology of network becomes the random graph, the degree of synchronization becomes larger. In addition, when $\epsilon = 0.2$, most of nodes are synchronized regardless of p.

Next, we calculate the Lyapunov exponents of the chaos coupled system. The Lyapunov exponents quantify the dependence on a initial condition and the nonlinearity of systems. To calculate the Lyapunov exponents, we construct a attractor Z_t in n-dimensional state space:

$$Z_t = \{x_t^1, x_t^2, \cdots, x_t^n\}. \tag{4}$$

By calculating the difference δD_t enhanced from the initial difference δD_0 between the attractor Z_t and its nearest attractor, we estimate the Lyapunov exponents λ_d, $d = 1, \cdots, n$, as follows:

$$\lambda_d = \lim_{t \to \infty} \frac{1}{t} \log \frac{|\delta D_t|}{|\delta D_0|}. \tag{5}$$

Then, $\max\{\lambda_d\}$ is estimated as the maximum Lyapunov exponent λ_M.

Figure 2 shows the results of λ_M. When $\epsilon = 0.1$, λ_M is very large. Then, as p is smaller, that is, the topology of network becomes the regular graph, λ_M becomes larger.

IV. Estimation of Network Structure

To understand a complex system in detail, it is important to examine its network structure. This information helps us to image what happens on networks. In this study, we try to detect network structure only from time-series data derived by chaos coupled systems. To examine whether two nodes interact each other or not, that is, whether two nodes are connected or not, we used the following statistical measures. If these measures between two nodes are very large, we considered that these nodes are connected. Then, we equalized the number of estimated connections with the number of connections of the original network. To calculate the accuracy of estimated connections, we use the index:

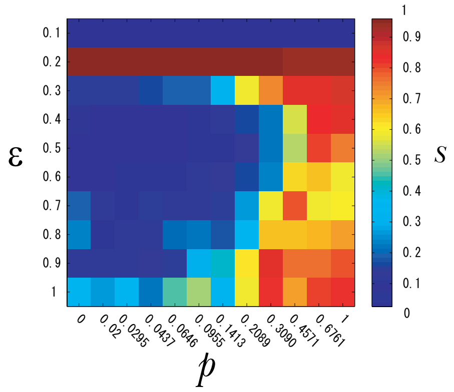

Fig. 1. The degree of synchronization S in the chaos coupled system defined by Eq.(1).

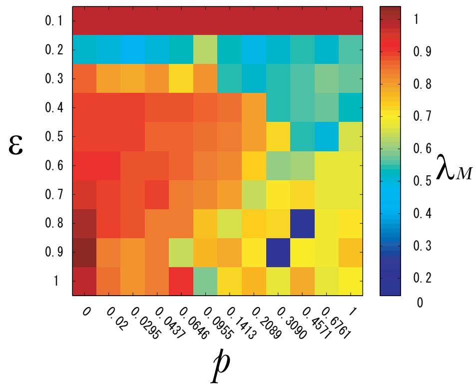

Fig. 2. The maximum Lyapunov exponent λ_M of the chaos coupled system defined by Eq.(1).

$$E = \frac{|P \cap N|}{|N|} \tag{6}$$

where P means the set of estimated connections, and N means the set of the original connections. That is, $E = 1$ means perfectly estimated the original connections.

A. The Correlation Coefficient

The correlation coefficient estimates linear correlation between two time series x^i and x^j.

$$C_{i,j} = \frac{\left\langle (x_t^i - \langle x^i \rangle)(x_t^j - \langle x^j \rangle) \right\rangle}{\sigma_{x^i} \sigma_{x^j}} \tag{7}$$

However, this measure cannot estimate nonlinear relation, and includes indirect correlation due to the common input into both node i and j.

The accuracy E by the correlation coefficient is shown in Fig.3. If p is larger, the accuracy E by the correlation

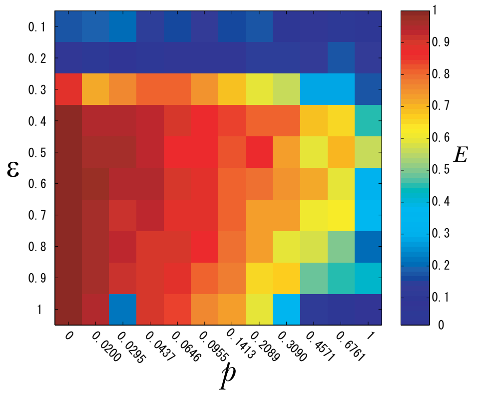

Fig. 3. The accuracy E of network structures estimated by the correlation coefficient.

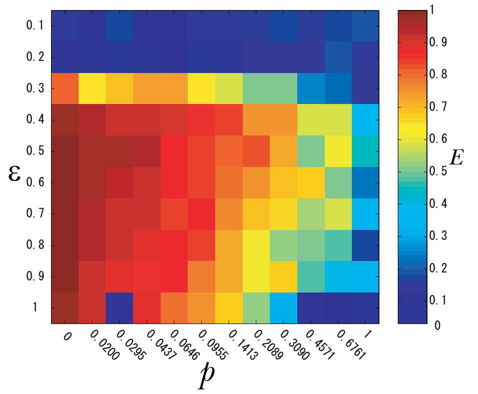

Fig. 5. The accuracy E of network structures estimated by the mutual information.

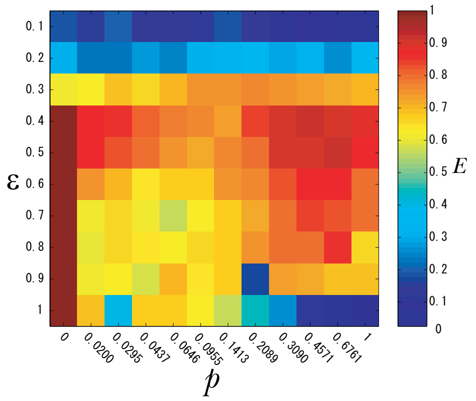

Fig. 4. The same as Fig.4, but the partial correlation coefficient.

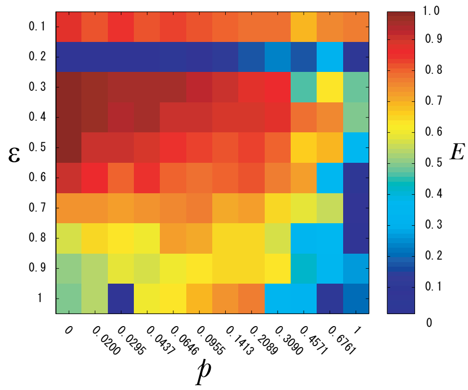

Fig. 6. The same as Fig.5, but the transfer entropy.

coefficient is worse. The reason is that synchronization among nodes is strong as shown in Fig.1, and it is difficult to remove indirect correlations. If p is small, that is, the network structure becomes the regular graph, the accuracy E becomes better because the synchronization is weak. Then, as ϵ is larger, the interaction among nodes becomes stronger and the behavior of systems is more affected by network topology; hence the accuracy E is greatly affected by p. However, when ϵ is small, because the synchronization is very strong in $\epsilon = 0.2$ and the maximum Lyapunov exponent is very large in $\epsilon = 0.1$, it is difficult to estimate such a network structure by the coefficient coefficient.

B. The Partial Correlation Coefficient

To remove indirect correlations, it is efficient to use the partial correlation coefficient:

$$P_{i,j} = \frac{-C'_{i,j}}{\sqrt{C'_{i,i}C'_{j,j}}} \tag{8}$$

where $C'_{i,i}$ is the inverse matrix of $C_{i,i}$.

The accuracy E by the partial correlation coefficient is shown in Fig.4. Even if the network structure is random and the synchronization is strong, we can estimate network structures more accurately by removing false correlations.

C. The Mutual Information

The mutual information can measure nonlinear relations. However, the mutual information has a similar problem to the correlation coefficient, that is, indirect relations are included. The mutual information is denoted as follows:

$$M_{i,j} = \sum_t p(x_t^i, x_t^j) \log \frac{p(x_t^i, x_t^j)}{p(x_t^i)p(x_t^j)}, \tag{9}$$

where x_t^i is a time series of the i-th node, $p(x_t^i, x_t^j)$ is a joint probability. We use the following equation to calculate joint probability:

$$p(x_t^i, x_t^j) = \frac{1}{L}\sum_{t'} \theta\left(F_s\gamma - \left|\begin{pmatrix} x_t^i \\ x_t^j \end{pmatrix} - \begin{pmatrix} x_{t'}^i \\ x_{t'}^j \end{pmatrix}\right|\right) \tag{10}$$

where θ is a step function, L is the length of time series data, F_s is a full scale range of the time series, and γ is a resolution rate. In this study, we set $\gamma = 0.2$.

Figure 5 shows the accuracy E by the mutual information. These results are similar to the case of the correlation coefficient.

D. The Transfer Entropy

The transfer entropy[1] can measure nonlinear relations between two nodes without indirect relations. The transfer entropy is denoted as follows:

$$T_{i\to j} = \sum_t p(x^j_{t+1}, x^j_t, x^i_t) \log \frac{p(x^j_{t+1}|x^j_t, x^i_t)}{p(x^j_{t+1}|x^j_t)} \qquad (11)$$

where $p(x^i_{t+1}|x^i_t, x^j_t)$ is a conditional probability. To simplify directional interactions between two nodes, we use the mean value of $T_{i\to j}$ and $T_{j\to i}$ for estimations:

$$T_{i,j} = \frac{1}{2}(T_{i\to j} + T_{j\to i}) \qquad (12)$$

as a statistical measure.

Figure 6 shows the accuracy E of network structure estimated by the measure $T_{i,j}$. In the regions where Lyapunov exponents are very large and nonlinearity is extremely strong, especially $\epsilon = 0.1$, the measure works better.

E. The Optimum Measure for Estimating Each Network Structure

To summarize the results shown in Figs.3–6, although the difference among these results are not very clear, we show the maximum accuracy E of network structure estimated by each statistical measure in Fig.7, and show the optimum measures to maximize the accuracy E according to modeling parameters p and ϵ in Fig.8. In Fig.7, the estimated accuracy is worse in $\epsilon = 0.2$ because the synchronization among nodes is very strong and time-series data of each node are almost the same.

As shown in Fig.8, the correlation coefficient is most efficient measure if synchronization of system is weak. As p becomes larger, the partial correlation coefficient works better. This measure can remove the false correlations made by strong synchronization among nodes. Moreover, if systems have the large Lyapunov exponents, the mutual information and the transfer entropy are efficient for estimating nonlinear interactions. In particular, the mutual information is more efficient when ϵ is larger, and the transfer entropy is more efficient when ϵ is smaller.

V. Conclusion

In this study, we estimated network structures by four statistical measures only from time series observed in chaos coupled map. Then, it was shown that the optimum measure which estimates network structure most accurately changes according to network parameters: the strength of interaction between connected nodes and network topologies. Moreover, to discuss the reason, we also estimated the degree of synchronization and the Lyapunov exponents of each system.

As results, in the regular network, the correlation coefficient is more suitable because each node is not almost synchronized and the false correlation dose not appear. However, because, in the random network, each node is synchronized strongly and the false correlation appears, the partial correlation coefficient becomes more suitable. Furthermore, as the Lyapunov exponents become larger, nonlinear measures – the mutual information and the transfer entropy become more suitable. In particular, if connected nodes interact weakly, the transfer entropy works well.

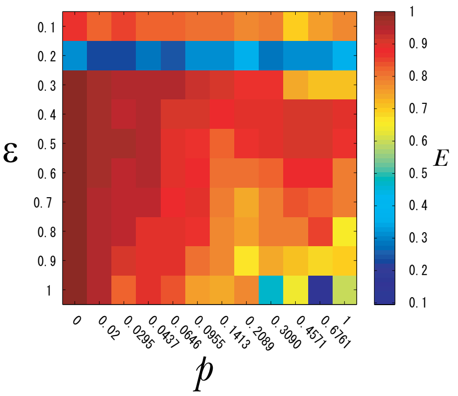

Fig. 7. The maximum accuracy E of network structure estimated by each statistical measure.

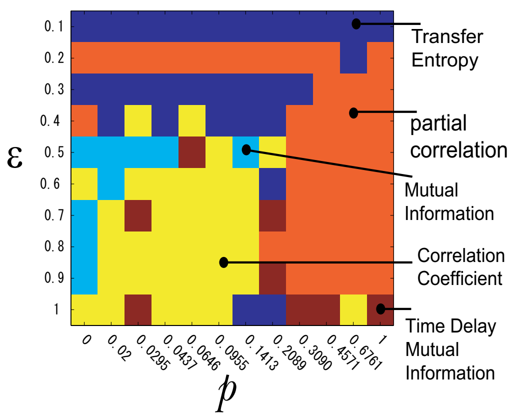

Fig. 8. The optimum measures to maximize the accuracy E of estimated network structure.

This research was partially supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Young Scientists Start-Up, 18800061, 2006–2007.

References

[1] T. Schreiber: "Measuring Information Transfer," Physical Review Letters, vol.85, pp.461–464, 2000.

[2] D. Watts and S. Strogatz: "Collective dynamics of 'small-world' networks." Nature, vol.393, pp.440–442, 1998.

[3] K. Kaneko: "Simulating Physics with Coupled Map Lattices – Pattern Dynamics, Information Flow, and Thermodynamics of Spatiotemporal Chaos" in Formation, Dynamics, and Statistics of Patterns, World. Sci., 1990.