



Transformation of growing networks to time series and its nonlinear time series analysis

Yutaka Shimada[†], Yuta Haraguchi[†], and Tohru Ikeguchi^{†‡}

[†]Graduate School of Science and Engineering, Saitama University
255 Shimo-ohkubo, Sakura-ku, Saitama, 338-8570 Japan

[‡]Saitama University Brain Science Institute

Email: sima@nls.ics.saitama-u.ac.jp, haraguchi@nls.ics.saitama-u.ac.jp, tohru@mail.saitama-u.ac.jp

Abstract—In this paper, we propose a novel method to generate a time series from a complex network to identify an underlying evolution process behind the complex network. Based on the adjacency information between two nodes, we first arrange the nodes in the complex network in the Euclidean space, and show that their arrangements clearly correspond to their network structure. We further show that if the complex network is a growing network, namely complex network with temporal evolutions, we can regard the obtained arrangements as a time series. Thus, we can transform the complex network to a time series and analyze the time series using the methods based on the nonlinear dynamical system theory.

1. Introduction

In the real world, there exist several complex networks, for example, human relationships, synaptic connections in neural systems, the world-wide-web, ecological food webs, and so on. These networks have complex structures and recent researches on real complex networks have been clarified that some of these structures are led by temporal evolutions of the networks [1, 2, 3]. In such networks, nodes are added to the network at each temporal step, then the network grows with its complex structure formed. Thus, it is an important issue to clarify how the networks evolve.

Here, we propose a novel method which elucidate an underlying evolution process of growing complex networks. In the proposed method, we use geometric arrangement of nodes and temporal information when the nodes are added to the network. First, the nodes are arranged in the Euclidean space based on their adjacency relationships by using the classical multidimensional scaling[4]. Here, if the networks have temporal information, the nodes have a temporal order. Then, we can recognize values of the nodes as a time series and thereby can transform the complex networks to the time series. We analyze the obtained time series by using the time-series analysis methods based on the nonlinear dynamical system theory.

In numerical simulations, we show that the arrangements of the nodes in the complex networks clearly correspond to their network structures. Using the obtained arrangements and the temporal information, we elucidate the evolution

process of the complex networks generated by mathematical models.

2. Method

In our method, using the classical multidimensional scaling(CMDS)[4], we obtain arrangements of nodes in a complex network in a d -dimensional Euclidean space. If distances d_{ij} between two any objects i and j ($i, j = 1, \dots, N$) are given, the CMDS gives an arrangement of these objects in the d -dimensional Euclidean space so that their distance relations are preserved as correctly as possible. Here, our method is originated from the method proposed by Hirata et al.[5] in which an original time series can be reproduced only from its recurrence plot. We extend this concept to the complex network analysis. Using the obtained arrangements and temporal information, we can transform the network to a time series. Our method consists of the following three steps:

1. Generating a distance matrix.
2. Applying the CMDS to the distance matrix.
3. Transforming the network to a time series by using the arrangement of nodes and the temporal information.

2.1. Generation of distance matrix

To obtain an arrangement of nodes, we first need to define distances between any two nodes. Here, we introduce two methods for generating the distance matrix $D = (d_{ij})$.

The type-1 method is originally proposed by Hirata et al. that determines the distances between two nodes based on their adjacency relationships[5]. In the type-1 method, the distance d_{ij} between two adjacent nodes i and j is determined by the following equation:

$$d_{ij} = 1 - \frac{|G_i \cap G_j|}{|G_i \cup G_j|}, \quad (1)$$

where G_i is the set of labels of adjacent nodes of the node i , $|G|$ is the number of elements in the set G , and \cap and \cup are the union and intersection of two sets. The distances between non-adjacent nodes are determined by the shortest

path length by using the distances between the adjacent two nodes.

The type-2 method uses the adjacency information of each nodes directly [6]. If the i th and j th nodes are connected, these nodes are similar and we set d_{ij} to unity but if two nodes are not connected, these two nodes are dissimilar and we set d_{ij} to $w(> 1)$. If $i = j$, we set d_{ij} to zero. Then, in the type-2 method, the distance matrix has only three values.

2.2. Classical multidimensional scaling

Next, we apply the CMDS to the distance matrix generated from the complex networks. In the CMDS, there are mainly two procedures[4]: the centering of matrix $A = \frac{1}{2}D^{(2)} = (-d_{ij}^2/2)$ by the centering matrix J and its eigenvector decomposition, namely $JAJ = V\Lambda V^T$, where $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_d)$, $V = (v_1 \cdots v_d)$, and λ_i and v_i are the i th eigenvalue and the i th eigenvector. Here, let the coordinate value of the i th node be x_i and $X = (x_1 \cdots x_N)$. From the relationship between the innerproduct and the distance[4], $X = V\Lambda^{1/2}$, where $\Lambda^{1/2} = \text{diag}(\lambda_1^{1/2}, \dots, \lambda_d^{1/2})$. Then the dimension, d , of the Euclidean space in which nodes are arranged is the rank of X . By the above mentioned procedure, we can obtain the coordinate values of the nodes X only from the distance matrix D .

2.3. Complex networks to time series

To transform the network to a time series, we use the obtained arrangements and the temporal information when each node is added to the network. We consider that this temporal order is described as labels of the nodes. For example, if a node is the first node added to the network, the label of the node is one, if the second, the label is two, if the third, the label is three, and so on. Thus, considering the coordinate values of each node as the amplitude of the time series and the labels as the temporal order, we can transform the complex networks to the time series. In this paper, for the sake of simplicity, we consider that only one node is added to the network at each temporal step.

3. Experimental settings

To evaluate the validity of our method, we conduct two numerical experiments. First, we confirm how does the nodes in the networks are arranged in the Euclidean space. We generated the networks by using three models: the Watts and Strogatz (WS) model [7], the Barabási and Albert (BA) model [1], and the Dorogovtsev, Goltsev, and Mendes (DGM) model [8].

In the WS model, introducing the link-rewiring probability p from zero to unity, we varied the network structure from the regular network to random networks. In the numerical simulations, we generated the networks with the WS model, which have 1,000 nodes with the average degree of ten. We set the rewiring probability p to zero in the

regular, 0.01 in the small world, and unity in the random networks. In the BA model, the network starts with a complete graph consisted of m_0 nodes and only one node with l links is added to the network at each temporal step. The additional nodes are easy to be connected to nodes with high degree. We set both m_0 and l to four and generate a network with 1,000 nodes in the numerical simulations. The DGM model generates the network having a pseudo-fractal structures [8]. In the numerical simulations, we generate a network with 1,095 nodes by the DGM model.

Next, we transform the networks to time series and analyze them. To transform the networks to time series, we need temporal information. The BA model satisfies this condition, because each node is added to the network at temporal step and thereby the nodes have the temporal order. Here, modifying the DGM model, we can generate a model which deterministically grows and satisfies the condition, namely only one nodes is added to the network at each temporal step. In the DGM model, an initial graph($t = 0$) consists of two nodes with one link. At each temporal step, one node for each link is added to the network. Thus, the number of nodes which are added to the network at time t is $2 \cdot 3^{t-1}$ ($t > 0$). We modified the DGM model as follows:

1. Start with an initial graph consisted of two nodes and one link. Here, the label of one node is zero, that of another node is one, and that of the link is zero. We set time t to zero. Let the link be l^* , which has the smallest label.
2. Add one new node with two links to the network and increase time t by one. This new node is connected to two nodes at both ends of the link l^* .
3. Update the label of the link by $l^* \leftarrow 3t - 2$.
4. Set labels of the new additional node and two links. We set the label of the new node to $t + 1$. We set one of two links which is connected to the node having the smaller label to $l^* + 1$ and the larger label to $l^* + 2$.
5. Set the link l^* to the link with the smallest label.
6. Repeat steps 2–5.

Because the modified DGM model deterministically grows with the pseudo-fractal structure formed in a bounded region, its evolution process might show a chaotic-like behavior. On the other hand, the BA model grows based on the stochastic rules. Thus, these two networks have different evolution processes. In the numerical simulations, we generate growing networks from the BA and the modified DGM models and transform these growing networks to time series. By using the time series prediction, we elucidate the difference of the evolution processes between them[9].

To determine the evolution processes of networks by the prediction, we use the first eigenvector obtained from the

networks and reconstruct it in the s -dimensional state space by the time delay coordinate[10]. In the numerical simulations, we set s to eight and determine the temporal delay τ by mutual information [11].

Let $\mathbf{x}(t)$ be a point in the s -dimensional reconstructed state space at time t . In the prediction, we estimate the point $\hat{\mathbf{x}}(t + T_p)$, the T_p steps after $\mathbf{x}(t)$ by the following equation:

$$\hat{\mathbf{x}}(t + T_p) = \frac{\sum_{i=1}^M \exp(-d_i) \mathbf{x}(m_i + T_p)}{\sum_{i=1}^M \exp(-d_i)}, \quad (2)$$

where $\mathbf{x}(m_i)$ is the i th nearest neighbor of the $\mathbf{x}(t)$, d_i is the Euclidean distance between $\mathbf{x}(t)$ and $\mathbf{x}(m_i)$, and M is the number of the nearest neighbors.

We use the first half of the time series as the database for the predictions and predict the last half. To evaluate the prediction accuracy, we investigate the correlation coefficient between the true time series, namely the last half of the time series, and its predicted time series.

4. Results and discussions

Figure 1 shows examples of the arrangements of nodes in the networks. To describe the arrangements in the plane, we use the two eigenvectors with the first and second largest eigenvalues. Figures 1(a), (b), and (c) are the arrangements for the regular, small world, and random networks generated from the WS model. Figure 1(d) is the arrangements for the scale free network generated by the BA model. Figure 1(e) is the arrangements for the scale free network generated by the DGM model. In Fig. 1, the left column shows the results for the type-1 method, the right column shows those for the type-2 method.

From Figure 1, in both the type-1 and type-2 methods, we can see clear correspondences of the arrangements to the network structures; the arrangement of the regular network is regular, that of the random network is random, and that of the small world is in between them (Figs. 1(a), (b), and (c)). As shown in the random network (Fig. 1(c)), the network generated from the BA model also has the random arrangement because each additional node randomly connects to other nodes and the connections depends on the probability determined by the degree of each node. Because the network generated from the DGM model has the pseudo-fractal structures[8], its arrangements shows a characteristic structure.

Here, the significant differences between the type-1 and the type-2 methods are shown in the scale free network. In the results for the type-2 method, a few points are arranged far from other points (Figs. 1(d) and (e) in the right column). These points correspond to hub nodes which have significantly higher degree than the other nodes. In the type-1 method, because the adjacency relationship is directly used, the differences between the hub nodes and the

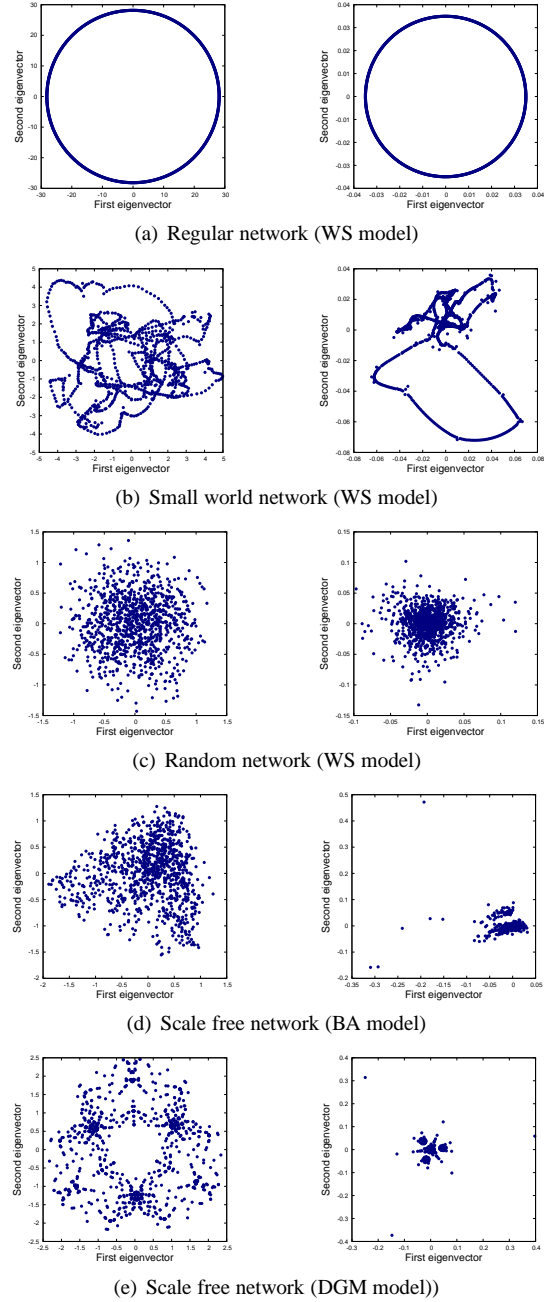
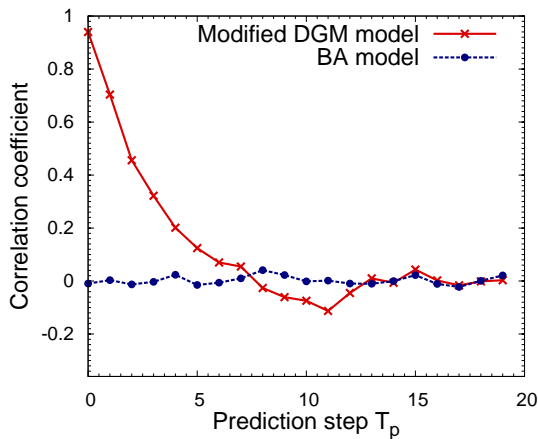


Figure 1: Examples of the arrangements of nodes in the (a) regular ($p = 0$), (b) small world ($p = 0.01$), (c) random ($p = 1$), and (d),(e) scale free networks. The left column shows the results for Type-1 method, The right column shows that for Type-2 method. The networks generated from (a),(b),(c) the WS model, (d) the BA model, and (e) the DGM model.

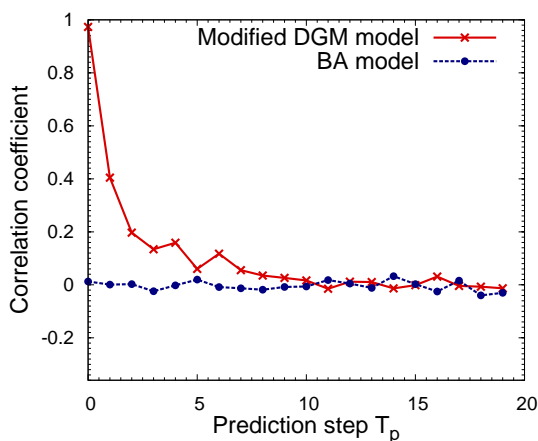
other nodes are emphasized in their arrangements. On the other hand, in the arrangement by the type-1 method, the hub nodes and the other nodes are indistinguishable due to indirect use of the adjacency relationship (Figs. 1(d), (e) in the left column).

Figure 2 shows the prediction results for time series ob-

tained from the scale free networks generated from the BA and the modified DGM models. Figures 2(a) and (b) are the results for the type-1 and type-2 method, and both methods show the same tendency. From Fig. 2, the correlation coefficients for the BA model are low for all prediction steps T_p . On the other hand, those of the modified DGM model are also low in the large T_p , but high in the small T_p . These tendencies indicate that the time series obtained from the BA model might be random and that from the modified DGM model might not be random but chaotic. These results imply that the time series obtained from the networks include the information of evolution processes of the original networks, and our method can clarify the underlying information of evolutions of the complex networks by only using static network structures and the temporal information.



(a) Type-1 method



(b) Type-2 method

Figure 2: Prediction results for the DGM model (the red line with crosses) and the BA model (the dark blue dashed line with circles). The results for (a) the type-1 method and (b) the type-2 method.

5. Conclusion

In this paper, we proposed a novel method which elucidate underlying evolution processes of growing complex networks. In the proposed method, using geometric arrangements of nodes and temporal information when the nodes are added to the network, we can transform the networks to time series.

In numerical simulations, we evaluate the proposed method with mathematical models. In addition, to generate the network which deterministically grows, we modified the model proposed by the Dorogovtsev et al. We first showed that the arrangements of nodes in the complex networks clearly correspond to their network structures, and further showed that using the obtained arrangements and the temporal information, we can determine the evolution processes of the complex networks generated by the mathematical models.

The research of Y.S. was supported by JSPS research follow(No.22-8148). The research of T.I. was partially supported by a Grant-in-Aid for Exploratory Research (No.20650032) from JSPS.

References

- [1] A.-L. Barabási and R. Albert, *Science*, 1999 Vol.286, pp.509–512.
- [2] M. E. J. Newman, *SIAM Review*, 2003, Vol.45, pp.167–256.
- [3] S. Boccaletti, V.Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, *Physics Reports*, 2006, Vol.424, pp.175–308.
- [4] T. F. Cox and M. A. A. Cox, *CHAPMAN & HALL/CRC*, 2000.
- [5] Y. Hirata, S. Horai S, and K. Aihara, *Eur. Phys. J. Special Topics* 2008, Vol.164, pp.13–22.
- [6] Y. Haraguchi, Y. Shimada, T. Ikeguchi, and K. Aihara, *LNCS*, 2009, Vol.5769, pp.325–334.
- [7] D. J. Watts and S. H. Strogatz, *Nature*, 1998, Vol.393, pp.440–442.
- [8] S. N. Dorogovtsev, A. V. Goltsev A V, and J. F. F. Measures, *Phys Rev E*, 2002, Vol.65, pp.066122.
- [9] G. Sugihara and R. May, *Nature*, 1998, Vol.344, pp.734–741.
- [10] N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw, *Phys. Rev. Lett.*, 1980, Vol.45, pp.712–716.
- [11] A. M. Fraser and H. L. Swinney, *Phys. Rev. A*, 1986, Vol.33, pp.1134–1140.