Detecting essential nodes in complex networks from measured noisy time series

Yao-Chen Hung[*] and Chin-Kun Hu[*,†]

[*]*Institute of Physics, Academia Sinica, Nankang, Taipei 11529, Taiwan*
[†]*Center for Nonlinear and Complex Systems and Department of Physics, Chung-Yuan Christian University, Chungli 320, Taiwan*

Abstract. A nonlinear measure, namely multi-interdependency, is proposed to detect the essential nodes in heterogeneously dynamical networks. The method is based upon the conceptions of the nearest conditional neighbors and singular value decomposition (SVD). Numerical results show that the value of multi-interdependency is positively correlated with the degree of nodes, which is beneficial to identify the nodes of topological and functional importance. Moreover, such a method has been demonstrated being robust against the effect of intrinsic noise.

Keywords: Complex networks; Interdependency
PACS: 05.45.Tp, 05.45.Xt, 89.75.Hc

INTRODUCTION

In complex networks, the node with large connection (namely hub) often plays an essential role. For example, in a protein-protein interaction network, the essential proteins generally have more degrees than the non-essential ones [1]. In an earthquake network, aftershocks tend to return to the neighborhood of the mainshok locus, which causes the mainshok to possess high connectivity [2]. Similar phenomena can also be observed in neural systems, food web networks, epidemiological systems, social networks and so on. From the viewpoint of dynamics, the removal of hubs would extremely alter the largest eigenvalue λ of the coupling matrix [3], and reduce the synchronizability of such a network [4]. How to detect these essential nodes in a topology-unknown network then becomes a relevant issue of potential applications.

In this paper, applying the ideas from single value decomposition method (SVD) [5], we calculate the multi-interdependency of nodes within a network to solve the aforementioned problem[1]. Numerical experiments are carried out from the simultaneously measured time series generated from coupled Hénon maps [6] on a scale-free network. We show that the interdependencies of the highly connected nodes are relatively larger than others, despite the fluctuation of intrinsic noise. The result provides us a useful judgement to distinguish the hubs from other merely connected nodes. Moreover, the variation of λ against the removal of nodes is applied to quantify our results further. Some discussions including limitations about this method are also presented.

[1] There are also other techniques available to define the multi-interdependency. The comparison of different definitions will be reported elsewhere.

CP922, *Noise and Fluctuations, 19th International Conference,* edited by M. Tacano, Y. Yamamoto, and M. Nakao

MULTI-INTERDEPENDENCY

Consider a network consisting of M nodes with the scale-free topology. Each node stands for a dynamical system, and links in the network represent the interactions (or coupling) between nodes. Let us denote by $\mathbf{X}^i$ the dynamical system of the ith node, and by x^i the measured scale time series of $\mathbf{X}^i$, where $i = 1, 2, ..., M$. From the time series x^i, we can reconstruct a D-dimensional vector $\mathbf{x}_n^i = (x_n^i, ..., x_{n+(D-1)\tau}^i)$ with a time delay τ and a time index $n = 1, ..., N$. Now the conception of conditional neighbors is introduced. Selecting arbitrarily the delay vectors of two nodes from the whole network, namely $\mathbf{x}_n^i$ and $\mathbf{x}_n^j$, $i \neq j$. Let $r_{n,k}^i$ and $r_{n,k}^j$, $k = 1, ..., K$, denote the time indices of the K nearest neighbors of $\mathbf{x}_n^i$ and $\mathbf{x}_n^j$, respectively. Then the K nearest conditional neighbors of $\mathbf{x}_n^i$ mapping from $\mathbf{x}_n^j$ are denoted by $\mathbf{x}^i_{r^j_{n,k}}$. Analogously, the K nearest conditional neighbors of $\mathbf{x}_n^j$ mapping from $\mathbf{x}_n^i$ are $\mathbf{x}^j_{r^i_{n,k}}$.

In the present work, we define the multi-interdependency based upon the SVD. Applying the SVD to a $D \times K$ matrix A, one can decompose the matrix as $A = USV^T$, where U is a $D \times D$ orthogonal matrix, V is a $K \times K$ orthogonal matrix, and S is a $D \times K$ singular matrix whose off-diagonal elements are all zero. Replacing the matrix A by the true and conditional neighbor matrices of the vectors $\mathbf{x}_n^i$ and $\mathbf{x}_n^j$, we can get:

$$\left[\mathbf{x}^i_{r^i_{n,1}}, ..., \mathbf{x}^i_{r^i_{n,K}}\right] = U_{\mathbf{x}_n^i} S_{\mathbf{x}_n^i} V^T_{\mathbf{x}_n^i}, \tag{1}$$

$$\left[\mathbf{x}^i_{r^j_{n,1}}, ..., \mathbf{x}^i_{r^j_{n,K}}\right] = U_{\mathbf{x}_n^i|\mathbf{x}_n^j} S_{\mathbf{x}_n^i|\mathbf{x}_n^j} V^T_{\mathbf{x}_n^i|\mathbf{x}_n^j}, \tag{2}$$

$$\left[\mathbf{x}^j_{r^j_{n,1}}, ..., \mathbf{x}^j_{r^j_{n,K}}\right] = U_{\mathbf{x}_n^j} S_{\mathbf{x}_n^j} V^T_{\mathbf{x}_n^j}, \tag{3}$$

$$\left[\mathbf{x}^j_{r^i_{n,1}}, ..., \mathbf{x}^j_{r^i_{n,K}}\right] = U_{\mathbf{x}_n^j|\mathbf{x}_n^i} S_{\mathbf{x}_n^j|\mathbf{x}_n^i} V^T_{\mathbf{x}_n^j|\mathbf{x}_n^i}. \tag{4}$$

Then the dependency of x^i on x^j is given by

$$C(x^i|x^j) = \frac{1}{N} \sum_{n=1}^{N} \frac{tr\left(\left|U^T_{\mathbf{x}_n^i} U_{\mathbf{x}_n^i|\mathbf{x}_n^j}\right|, d\right)}{d}, \tag{5}$$

where $tr(|G|, d)$ is the summation of the absolute values of the first d diagonal elements of the matrix G [5]. The values of $C(x^i|x^j)$ are in the range of $[0, 1]$, and a great $C(x^i|x^j)$ indicates the synchronization between x^i and x^j. Now, the averaged interdependency of x^i among the other nodes in the network reads

$$\bar{C}(x^i) = \frac{1}{M-1} \sum_{j \neq i}^{M} C(x^i|x^j), \tag{6}$$

which defines the multi-interdependency of the node $\mathbf{X}^i$.

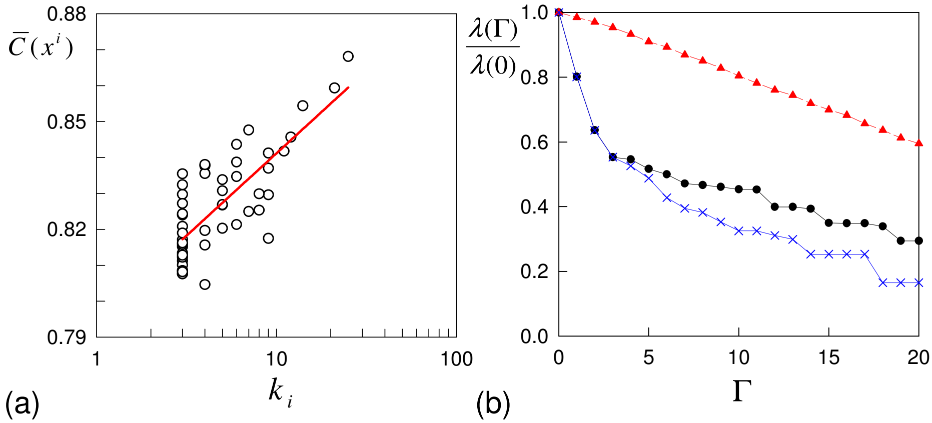

FIGURE 1. (a) Multi-interdependency $\bar{C}(x^i)$ as a function of the degree of a node k_i. (b) The largest eigenvalue $\lambda(\Gamma)$ of the adjacency matrix Θ resulting from removing Γ nodes in the order of descending degree (crosses), multi-interdependency (circles), and randomness (triangles). The time series are generated from the coupled Hénon maps network with $M = 50$, $\varepsilon = 0.48$, and $\sigma = 0.01$.

NUMERICAL EXPERIMENTS

The coupled Hénon maps network is adopted as our testing model :

$$\begin{aligned} x_i(t+1) &= 1.4 - (\varepsilon x_i(t)F_i(t) + (1-\varepsilon)x_i^2(t)) + b_i y_i(t) + \sigma\xi(t), \\ y_i(t+1) &= x_i(t), \end{aligned} \tag{7}$$

where the subscript $i = 1, 2, ..., M$ denotes the node index, $\varepsilon \in [0, 1]$ is the coupling strength, and the chaotic parameter $b_i = 0.1$, $\forall i$. $F_i(t) \equiv (1/k_i)\Sigma_{j=1}^{N}\Theta_{ij}x_j(t)$ denotes the local mean filed of the node i, where Θ_{ij} is an element of the $M \times M$ adjacency matrix Θ that characterizes the topology of such a network. Elements Θ_{ij} take the value one when there is a connection existing between the node i and j with $i \neq j$ and zero otherwise ($\Theta_{ii} = 0$). The connections are bi-directional; i.e., $\Theta_{ij} = \Theta_{ji}$ and Θ is a symmetric matrix. $k_i = \Sigma_j \Theta_{ij}$ is the degree of the node i. $\xi(t)$ is a white noise with an according intensity σ. All nodes evolve from different and random initial conditions.

The adjacency matrix Θ is generated by using the Barabási-Albert scale-free model [7]. Starting with a small number m_0 of fully interconnected nodes, a new node is introduced to connect with m ($m \leq m_0$) previous node at every time step. The connection probability depends on the degree of already-existing nodes (preferential attachment). As a general feature, the connectivity distribution follows a power law with an exponent constant $\lambda = 3$, regardless of m_0 and m. The conditions $m_0 = m = 3$ and the network size $M = 50$ are used throughout this article.

After sufficient transitions, $N = 5000$ iterations of $x_i(t)$ are recorded as the measured time series to calculate the multi-interdependency. With respect to the coupled Hénon maps, delay vectors are constructed with a time delay $\tau = 1$ and a dimension $D = 4$. The number of the nearest neighbors $K = 30$ is selected. Fig.(1a) presents the multi-interdependency $\bar{C}(x^i)$ as a function of the degree k_i of the node with $\varepsilon = 0.48$ and $\sigma = 0.01$. As one can observe, *the highly connected node is provided with a relatively larger* $\bar{C}(x^i)$, regardless of the effect of intrinsic noise. To quantify our results further, we explore the relative change in the largest eigenvalue of the network adjacency matrix against the removal of nodes [3]. Three different removal strategies are applied for the purpose of comparison, i.e., in the order of descending degree, descending multi-interdependency, and randomness. Fig.(1b) shows the normalized largest eigenvalue $\lambda(\Gamma)/\lambda(0)$ as a function of the number of the removed nodes Γ. We find that the removal of nodes by using interdependency extremely alters the value of $\lambda(\Gamma)$ (circles), which is similar to the results by using the degree (crosses). That is, the essential nodes are unveiled successfully by such a measure. These two methods are, undoubtedly, much more efficient than removing nodes randomly (triangles). The multi-interdependency method still works well even though it is evaluated from the series veiled by stronger fluctuation (the largest noise intensity we have investigated is $\sigma = 0.2$). It has also been successfully applied to other oscillators forming the networks [8].

DISCUSSION AND CONCLUSION

The conception of interdependency used in this paper was originally developed to characterize the generalized synchronization (GS) between two coupled oscillators. Recently, the GS behaviors have also been found emerging in complex networks of dynamical units [9], which then make it profitable to redefine the measure to the scale of the whole network. Since the method is on the grounds of synchronization, it would become invalid once the coupling strength is too weak to induce GS. In our case, the working region of the multi-interdependency is behind $\varepsilon_c = 0.32$.

The method shows possible applicability in the real-world complex systems, such as the stock market and the human brain. A study of more realistic network is under way.

ACKNOWLEDGMENTS

This work was partially supported by National Science Council of the Republic of China (Taiwan) under Grant No. NSC 95-2112-M 001-008 and National Center for Theoretical Sciences in Taiwan.

REFERENCES

1. H. Yu, D. Greenbaum, H. X. Lu, X. Zhu and M. Gerstein, *Trends Genet* **20**, 227 (2004).
2. S. Abe and N. Suzuki, *Europhys. Lett.* **65**, 581 (2004).
3. J. G. Restrepo, E. Ott, and B. R. Hunt, *Phys. Rev. Lett.* **97**, 094102 (2006).
4. J. Fan, and X. F. Wang, *Physica A* **349**, 443 (2005).
5. X. Hu and V. Nenov, *Phys. Rev. E* **69**, 026206 (2004).
6. M. Henon, Commun. Math. Phys. **50**, 69 (1976).
7. A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
8. Y.-C. Hung and C.-K. Hu, preprint.
9. Y.-C. Hung, Y.-T. Huang, M.-C. Ho and C.-K. Hu, unpublished.