

This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at SciVerse ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Network analysis of time series under the constraint of fixed nearest neighbors



Yan Dong, Wenwen Huang, Zonghua Liu, Shuguang Guan *

Department of Physics, East China Normal University, Shanghai, 200241, PR China

ARTICLE INFO

Article history:

Received 20 July 2012

Received in revised form 25 September 2012

Available online 22 October 2012

Keywords:

Complex networks

Time series analysis

Phase space reconstruction

ABSTRACT

In this paper, we carried out network analysis for typical time series, such as periodic signals, chaotic maps, Gaussian white noise, and fractal Brownian motions. By reconstructing the phase space for a given time series, we can generate a network under the constraint of fixed nearest neighbors. The mapped networks are then analyzed from both the statistical properties, such as degree distribution, clustering coefficient, betweenness, etc, as well as the local topological structures, i.e., network motifs. It is shown that time series of different nature can be distinguished from these two aspects of the constructed networks.

© 2012 Elsevier B.V. All rights reserved.

Nonlinear time series analysis is an important branch in the field of nonlinear dynamics, which is of great theoretical and application significance in science, engineering, finance and economics, as well as life and medical science, etc. [1]. In the past decade, with the rapid development of complex network theory [2–4], a method, which investigates time series by constructing networks, has brought new ideas into this area. Basically, this method first maps a given time series into a network by certain operations. Then the generated network can be analyzed in terms of concepts and approaches developed in the area of complex networks. Hopefully, this essentially different description for time series would provide new perspectives to the traditional domain of time series analysis.

The key point to construct a network from a time series is to define nodes and edges properly. So far, it is still an open question to reasonably map a time series into a network so that the latter could keep sufficient information to effectively exhibit the characteristics of the former. Nevertheless, there are three types of tentative methods proposed already. Let us briefly review them in the following. (1) Constructing a network from pseudo-periodic time series. This method was proposed by Zhang et al. in Refs. [5,6]. In this method, the pseudo-periodic time series is approximately divided into many “cycles”, i.e., small intervals of approximate cycles. Each “cycle” then can be treated as a node and the connection between two nodes can be established if the distance between them, which is defined similar to Euclidean norm in vector space, is less than a given threshold. The approach is first designed for pseudo-periodic time series, such as signals from chaotic Rossler or human ECGs. Later, it has been extended to arbitrary time series by constructing networks from the correlation matrix in Ref. [7]. Employing this method, it has been shown that periodic series with noise can be easily distinguished from chaotic Rossler time series. (2) The method of visibility graph. This method was proposed by Lacasa et al. in Ref. [8]. In principle, this method is suitable for all time series. In this method, each point in the time series is naturally considered as a node in the network, and two nodes connect each other if a straight line between them does not intersect any other points between them. Geometrically, this means that all the points in between do not block the visibility of these two reference nodes. This is why the method is called a visibility graph. Later, a simplified version of this method is developed, namely, the horizontal visibility graph [9,10]. By constructing a visibility network, it has been shown that the dynamical properties of the time series can be represented in the transferred networks. For example, periodic time series lead to a regular network, random time

* Corresponding author.

E-mail addresses: sgguan@phy.ecnu.edu.cn, guanshuguang@hotmail.com (S. Guan).

series are mapped into a network with exponential degree distribution, and a fractional Brownian time series generates a network with power-law degree distribution [11–14]. (3) The method of reconstructing phase space. This method was proposed by Xu et al. in Ref. [15]. In this method, the phase space embedding is first applied for the time series. Then each vector in the phase space is considered as a node in the network. Whether there is a link between two nodes or not depends on the distance in phase space between them [16]. In this way, a network can be generated. In fact, the recurrence plots and the related analysis methods, which are developed before the booming of complex network science, are essentially the graph or network representation of nonlinear dynamics [17]. Given a dynamics, for example, a time series, its recurrence plot is a square binary matrix. If the distance between two state vectors in phase space is less than a certain recurrence threshold, the corresponding matrix element is 1, and 0 otherwise. Thus the recurrence plot characterizes the times when the trajectory of the dynamics visits roughly the same neighborhood in the phase space. If the state vectors in phase space are treated as nodes, a recurrence network is readily obtained based on the recurrence plot [18,19]. In this way, the recurrence network provides a generic method to map time series into a network, and analyze time series from the perspective of network topology.

As we know, the properties of complex networks are usually described by their global or statistical properties, such as degree distribution, clustering coefficient, betweenness, etc. However, it has been shown that the statistical properties alone are not complete for characterizing complex networks. The local structures or connecting patterns also play significant roles in network dynamics. For example, if a network slightly changes its local topology, which hardly influences its statistical properties, its synchronizability, however, can be drastically changed [20]. Recently, the study on network motifs, i.e., the small subnetworks that typically consist of a few fixed numbers of nodes, aims at characterizing the local topological patterns in complex networks. In fact, networks have rich information at different scales, ranging from microscopic level (single node properties) to macroscopic level (global properties), with the mesoscopic level between [21,22]. In such a sense, network motifs can be classified as the properties of a network at the mesoscopic level. Definitely, the investigation on full scales of complex networks can broaden and deepen our understanding of network structures, organization and functions. It has been shown that network motifs play important roles in network structure and dynamics in many circumstances, such as in social networks, the ecological networks, the gene control networks and the protein interaction networks [23–25].

In Ref. [15], time series have been analyzed by investigating the properties of particular network motifs comprising four nearest nodes in the constructed network. Specifically, a network is first generated by embedding time series into a phase space. Then each phase space point is regarded as a node which connects its four nearest neighbors. It has been shown that the relative frequencies of these motifs occurring in a network can be used as a quantity to characterize different types of time series. This work was very insightful and enlightening, nevertheless, it has one disadvantage, that is, each node in the network only connects four nearest neighbors. In our opinion, such a network might not be adequate to represent the dynamical properties of the time series in phase space. This is because nonlinear dynamics is inherently determined by the local differential properties in the tangent space defined by the governing equations. After phase space embedding, the orbit actually becomes discrete, where the local properties are conserved and characterized by the relations among phase space points in the neighborhood. Under this circumstance, if the number of neighbors that each node connects is too small, the mapped network based on phase space embedding may not keep sufficient information to fully represent the dynamical properties of the original time series. Based on this consideration, in the present paper we extended the network-generating approach proposed in Ref. [15]. Mainly, we let each phase space node connect M nearest neighbors, where M is a moderately large number, and the choosing of M depends on a rule of thumb. By numerical experiments, it is found that M can be properly set as $10 \leq M \leq 30$. In Ref. [15], the number of connecting nearest neighbors $M = 4$, which is so small that the mapped network exhibits almost δ degree distribution (as shown in Fig. 3). As a consequence, the mapped networks are sparse and approximately homogeneous (as shown in Fig. 1). For such networks, it may not provide sufficient and reliable global information for the original time series. By extending the constraint number of nearest neighbors in phase space to appropriately larger values, the generated networks could be analyzed from both a global aspect, such as many statistical properties, and a local aspect, such as local topological patterns. Furthermore, choosing moderately large M also helps avoid possible redundant information in the generated network when M is very large. Therefore, networks mapped under this constraint could conserve essential and faithful dynamical information for analyzing the original time series.

To be specific, let us briefly describe our method to generate a network from time series. Totally, there are three steps. First, a time series of total length N is obtained from either dynamical models or practical observations. Then we choose an appropriate delay time and dimension to reconstruct a phase space for the time series [15,16]. In the present study, the dimension of phase space is usually chosen to be 5, and the results have been verified in a larger embedding dimension up to 10. Finally, we treat each phase space point (vector) as a node and connect them under the constraint of fixed nearest neighbors. In such a way, the time series can be mapped into a network. In Figs. 1 and 2, we schematically plotted the topologies of mapped networks for several typical time series, including periodic signals, chaotic maps, Gaussian white noise, and fractional Brownian motion (FBM).

We first report the results of network analysis for three typical time series, i.e., (1) periodic signals $x_n = \sin(2\pi\omega n)$; (2) chaotic logistic map $x_{n+1} = 4x_n(1 - x_n)$; and (3) Gaussian white noise $x_n \sim N_G(0, \sigma^2)$. Here, x is the variable, n is the discrete time step, and N_G denotes the Gaussian white noise with zero mean and σ standard deviation. Fig. 3 plots the degree distributions for the networks mapped from these three time series. From the figure, we found that the degree of a network from periodic signals follows a very steep exponential distribution; the degree of network from a chaotic logistic map is of typical exponential form; while the degree of a network from Gaussian white noise basically satisfies a Poisson distribution.

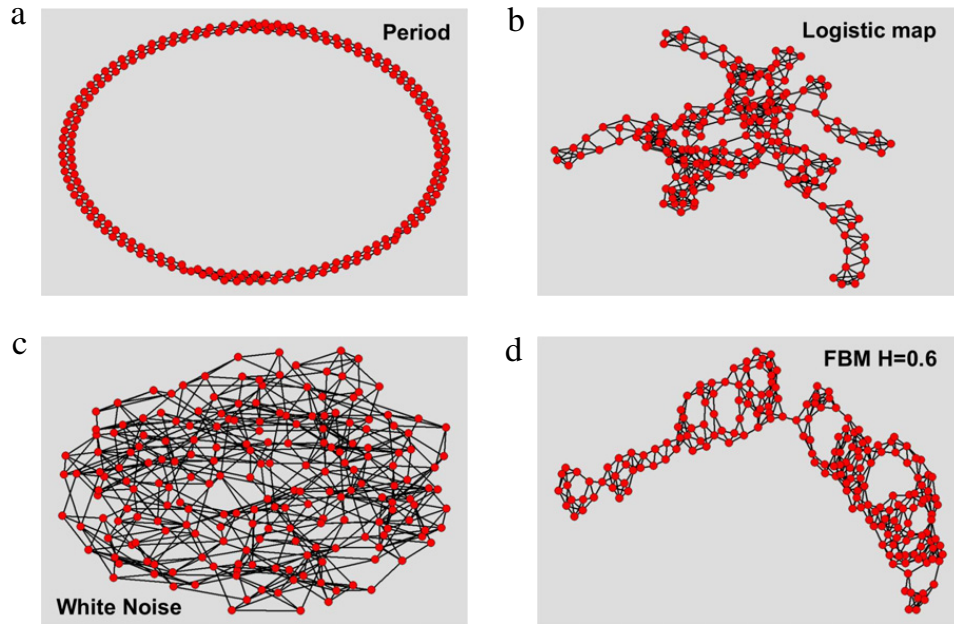


Fig. 1. (Color online) Networks mapped from (a) sine signals; (b) chaotic logistic map; (c) Gaussian white noise; and (d) fractal Brownian motion. The mapped networks are sparse and approximately homogeneous. $N = 200$, $M = 4$.

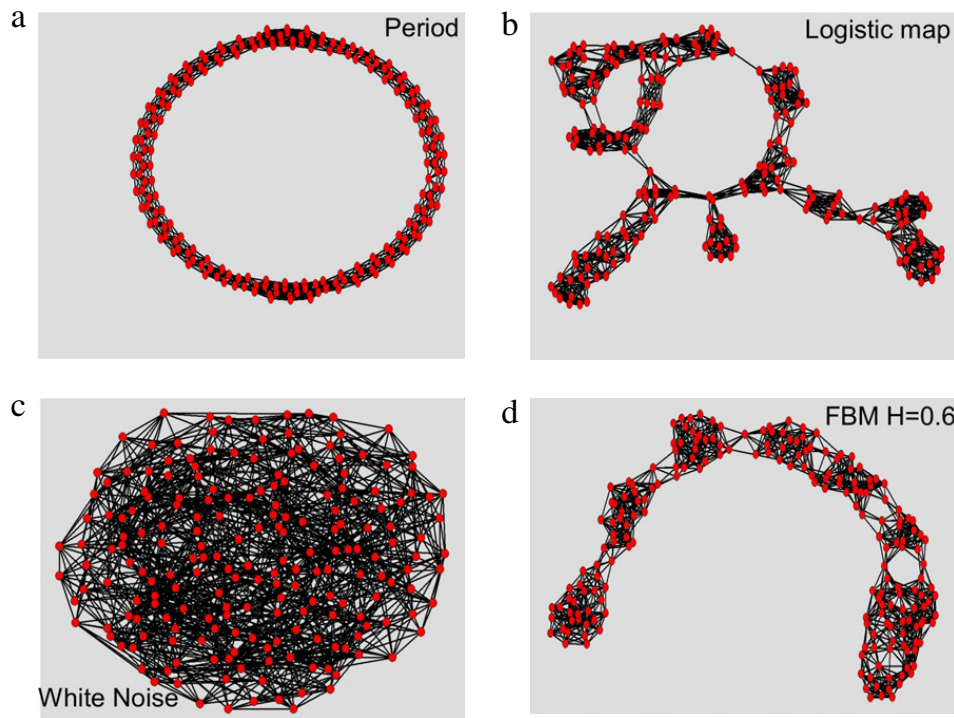


Fig. 2. (Color online) Comparison with Fig. 1; all the parameters are the same as in Fig. 1 except $M = 10$. (a) and (c) are basically a regular network and random network, respectively; while (b) and (d) exhibit typical modular structures. The network characteristics are more distinct compared with Fig. 1.

Thus from periodic series, to chaotic series, to noise, the degrees of the corresponding networks change from exponential form to Poisson distribution. The distribution width increases with M , i.e., the constraint number of nearest neighbors. Moreover, it is also found that for the former two degree distributions, the most probable degree exactly equals M , but for the last degree distribution corresponding to Gaussian white noise, the most probable degree is significantly greater than M . This is due to the different local differential properties of dynamics in phase space. If the phase space is locally homogeneous, the nearest neighbors are more likely to be mutual. However, if the phase space is locally very heterogeneous, the nearest neighbors mostly are not mutual. Apparently, in the latter case the average links (thus the average degree) can be larger than that in the former case. For comparison, in Fig. 3 we also plot the degree distributions for $M = 4$ cases, which are almost like a δ function due to the sparseness induced by the constraint of very small nearest neighbors. We further investigated

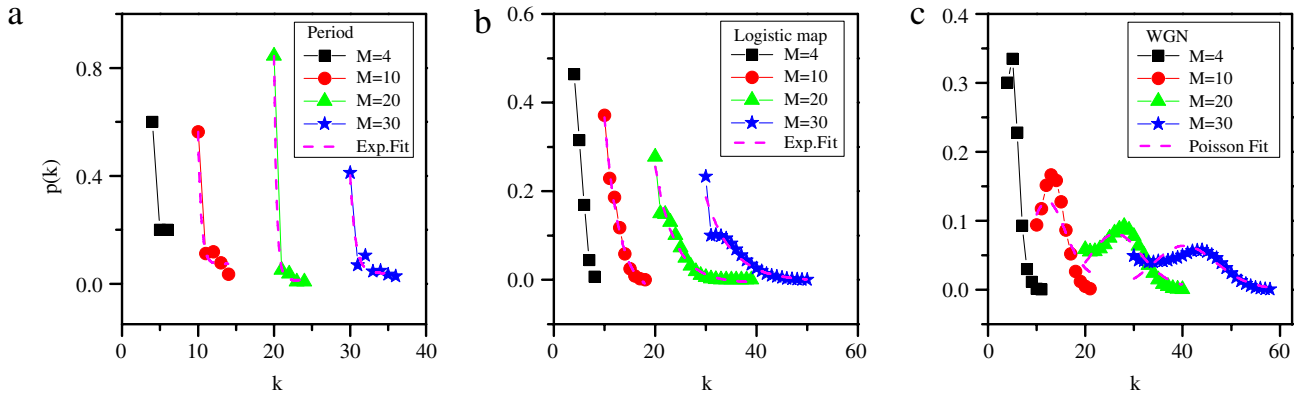


Fig. 3. (Color online) The degree distributions for networks mapped from typical time series. $N = 2000$.

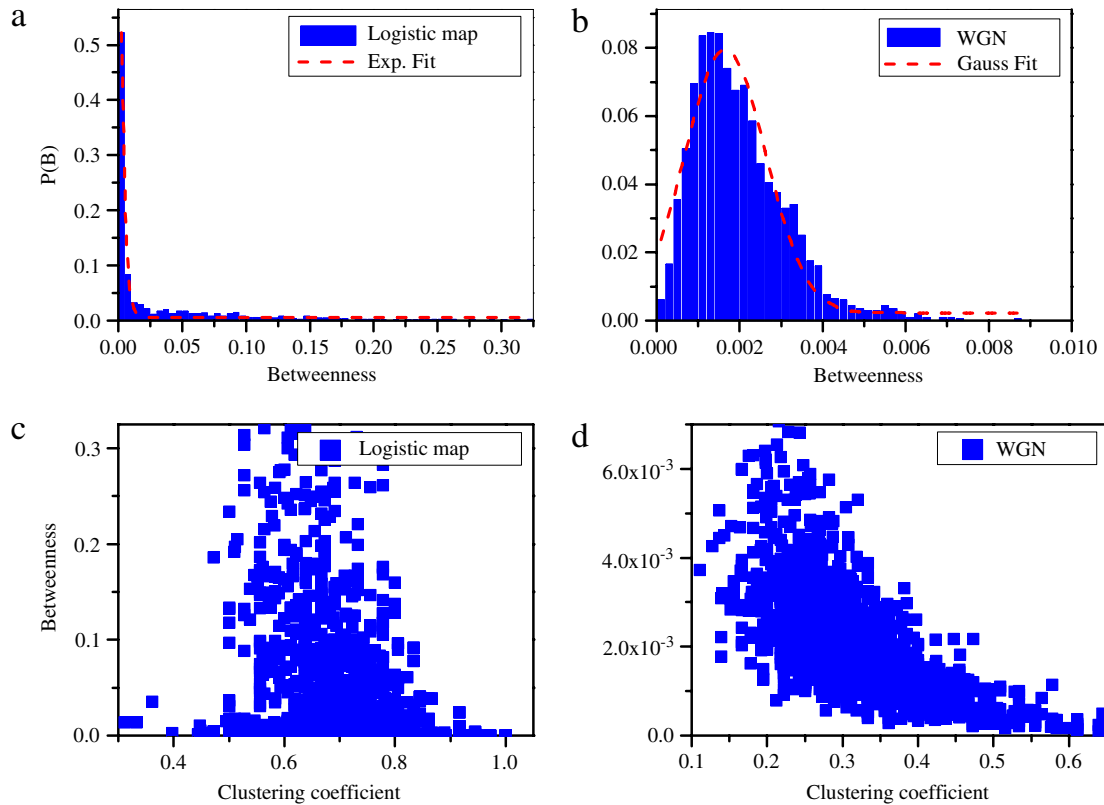


Fig. 4. (Color online) Comparing the betweenness distribution (upper panel) and the clustering coefficient-betweenness correlation (bottom panel) for networks mapped from chaotic logistic map and Gaussian white noise. $N = 2000$, $M = 10$.

other statistical properties for the mapped networks, such as cluster coefficients, betweenness, etc. Fig. 4 compares these quantities for mapped networks from chaotic map and Gaussian white noise. It is seen that betweenness for the two mapped networks follow different distributions. The former approximately satisfies exponential distribution (Fig. 4(a)), whereas the latter approximately follows a Gaussian form (Fig. 4(b)). Furthermore, it is also found that for a chaotic map, in the generated network the betweenness vs. clustering coefficient gives a bell-shaped function (Fig. 4(c)); while for the noise, the betweenness is negatively correlated with the clustering coefficient (Fig. 4(d)). As we know, due to the deterministic stochasticity of chaotic motions, the behaviors of chaotic time series and noise are very similar in coordinate space. Here, as we show by mapping time series into networks, some statistical quantities of the mapped network can be conveniently used to distinguish chaotic series from noise, which might be of practical applications in related engineering fields.

In the above, we have characterized the mapped networks in terms of their global statistical properties, such as distributions of degree, clustering coefficients and betweenness. In fact, the local topological patterns in complex networks are equally important in affecting the dynamics on it [20]. Recently, Xu et al. considered the subnetwork patterns or motifs of size four in networks mapped from time series, focusing on the local properties of networks [15]. Mainly, they investigated the occurrence of subnetworks of size four in their work. As shown in Fig. 5, in total, there are six different motifs consisting of four nodes. In the present work, we also analyze these network motifs in terms of their relative frequencies in the networks

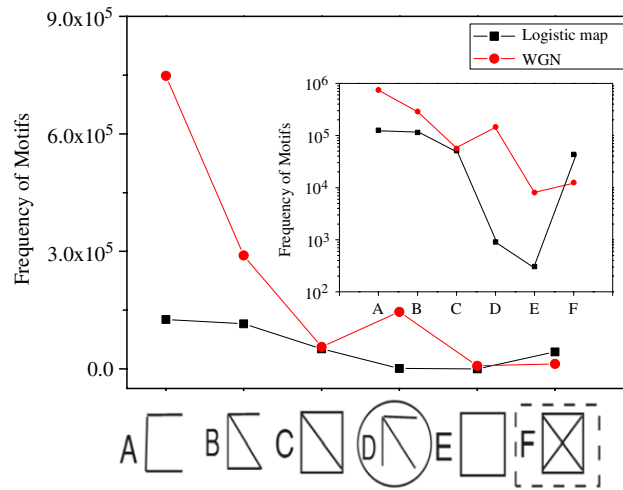


Fig. 5. (Color online) Frequencies of motifs D and F in the networks mapped from chaotic logistic map and Gaussian white noise. $N = 2000$, $M = 10$.

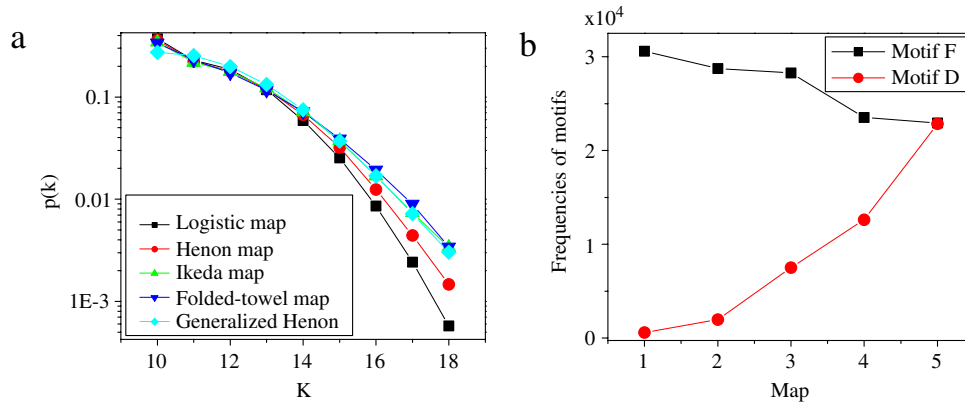


Fig. 6. (Color online) The degree distributions (a) and the frequencies of motifs D and F; (b) for networks generated from 5 chaotic maps. $N = 2000$, $M = 10$, and the average degree is about 11.

mapped from time series, i.e., for each constructed network, we count the number of occurrence of the target motifs. For example, Fig. 5 plots the occurrence rank of network motifs for two networks mapped from a chaotic logistic map and Gaussian white noise, respectively. It is found that the two networks can be distinguished by the relative frequencies of two particular network motifs, i.e., motifs D and F. For Gaussian white noise, the frequency of motif F is obviously smaller than that of motif D; while for a chaotic map, the frequency of motif F (and C) is significantly larger than that of motif D as shown in the inset of Fig. 5. This result is reasonable because motif D has more chance to occur in high-dimensional dynamics where the distribution of phase space points is more heterogeneous. On the contrary, motif F occurs more frequently in low-dimensional dynamics where the phase space points are relatively homogeneous. This finding is consistent with that in Ref. [15], showing that by properly allowing more nearest neighbors of connection, the mapped networks can still keep essential local properties of the time series. Furthermore, this extension in our work greatly increases the total occurrence of motifs. As pointed out in Ref. [15], the latter is a macroscopic measure similar to entropy, so a larger M , thus a larger total frequency, implies a relatively more reliable result.

Analysis of chaotic time series is of special theoretical and application importance. Therefore, in this work we further investigated the networks mapped from various chaotic maps. Typically, we choose the following chaotic maps:

1. Logistic map $x_{n+1} = 4x_n(1 - x_n)$;
2. Henon map $x_{n+1} = y_n + 1 - 1.4x_n$, $y_{n+1} = 0.3x_n$;
3. Ikeda map $x_{n+1} = 1 + 0.7(x_n \cos t_n - y_n \sin t_n)$, $y_{n+1} = 0.7(x_n \sin t_n + y_n \cos t_n)$, $t_n = 0.4 - 6/(1 + x_n^2 + y_n^2)$;
4. Folded-towel map $x_{n+1} = 3.8x_n(1 - x_n) - 0.05(y_n + 0.35)(1 - 2z_n)$, $y_{n+1} = 0.1[(y_n + 0.35)(1 + 2z_n) - 1](1 - 1.9x_n)$, $z_{n+1} = 3.78z_n(1 - z_n) + 0.2y_n$;
5. Generalized Henon map $x_{n+1} = 1.9 - y_n^2 - 0.03z_n$, $y_{n+1} = x_n$, $z_{n+1} = y_n$.

By constructing networks under the constraint of fixed nearest neighbors for the above time series, we are able to analyze the networks from both statistical and local perspectives. For example, Fig. 6(a) plots the degree distributions for the five mapped networks. They are qualitatively similar to each other, approximately following an exponential form. Furthermore, as shown in Fig. 6(b), the examination of the two key network motifs reveals that the frequencies of motif D and motif F

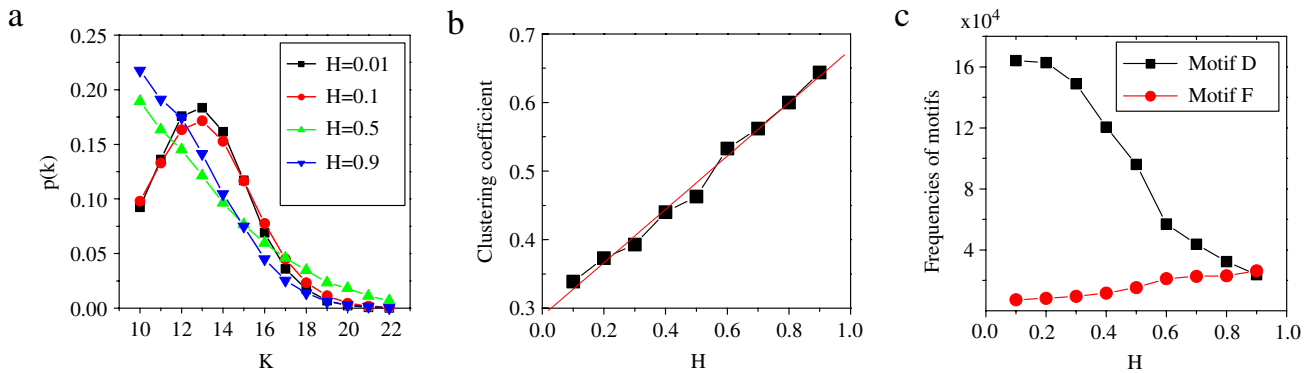


Fig. 7. (Color online) Characterizing the networks mapped from FBM with different Hurst exponents. (a) The degree distributions. (b) The clustering coefficient versus the Hurst exponent. (c) Frequencies of motifs D and F. $N = 2000$, $M = 10$ for all networks.

follow different trends, namely, from maps 1–5, motif D increases while motif F decreases. This suggests that when the dimension of chaotic maps increase, the corresponding networks become more and more heterogeneous.

Finally we apply network analysis for fractional Brownian motion, which is frequently used to model the fractal dynamics such as turbulence, economic and financial series, as well as physiological signals in life systems. FBM is a nonstationary random process characterized by the Hurst exponent H . $H = 0.5$ corresponds to one-step memory Brownian motion, and the time series with $H > 0.5$ and $H < 0.5$ show persistence and antipersistence, respectively. In Ref. [26], the FBM time series is mapped into a network by the phase space reconstruction method, and it is found that the relative frequencies of occurrence of the tetrad motifs can be used to classify time series into superfamilies. In the present work, we characterize the FBM time series by converting them into networks with appropriate numbers of nearest neighbors, i.e., M is of the order of 10. In this way, the network properties can be analyzed from both statistical quantities and local motifs. The results are illustrated in Fig. 7. It is found that FBM time series with different Hurst exponents H have different network characteristics. As the Hurst exponent H increases, the degree distributions of the mapped networks gradually change from Poisson form to exponential form, as shown in Fig. 7(a). In particular, we found a linear dependence between the clustering coefficients c of the mapped networks and the Hurst exponent H of the FBM time series. As shown in Fig. 7(b), the relation is numerically fitted as

$$c \sim 0.29 + 0.39H. \quad (1)$$

Eq. (1) shows that the clustering coefficient, which is one of the most important statistical properties of networks, has a linear relation with the Hurst exponent employing the network-generating method in this paper. On the other hand, the motif analysis shows that with the increase of the Hurst exponent, in the mapped networks motif F gradually increases while motif D drastically decreases, as shown in Fig. 7(c). This implies that networks mapped from FBM time series with larger Hurst exponents turn out to be more homogeneous locally. Combining the results in Fig. 7(b) and (c), we conclude that with the increase of the Hurst exponent, the transferred networks under the constraint of fixed nearest neighbors become more connected, and simultaneously more homogeneous.

To summarize, in this paper we have applied network analysis to time series by extending the network-generating method in Ref. [15] that is based on the properties of phase space embedding. We first reconstruct the phase space for a given time series. Then we treat the vectors in phase space as nodes, and allow each node to connect a fixed number of its nearest neighbors. By properly choosing a moderately large number of connecting neighbors, we can analyze the mapped networks from both global and local perspectives. We carried out extensive numerical simulations to typical time series, including periodic signals, chaotic maps, Gaussian white noise, and fractional Brownian motion. It is found that the network generated by the present method can characterize the dynamics of time series through both the statistical properties and local motifs in the associated networks. Thus time series of different dynamical nature can be distinguished from both aspects of the mapped networks. Our results are helpful for theoretical analysis as well as potential applications in the field of time series analysis.

Acknowledgments

SGG is sponsored by: the Science and Technology Commission of Shanghai Municipality grant no. 10PJ1403300; the Innovation Program of Shanghai Municipal Education Commission grant no. 12ZZ043; and the NSFC grant nos. 11075056 and 11135001. ZHL is sponsored by the NSFC grant nos. 10975053 and 11135001.

References

- [1] H. Kantz, T. Schreiber, *Nonlinear Time Series Analysis*, Cambridge University Press, Cambridge, 2004.
- [2] R. Albert, A.-L. Barabási, Statistical mechanics of complex networks, *Rev. Modern Phys.* 74 (2002) 47–97.

- [3] M.E.J. Newman, The structure and function of complex networks, *SIAM Rev.* 45 (2003) 167–256.
- [4] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D.-U. Hwang, Complex networks: structure and dynamics, *Phys. Rep.* 424 (2006) 175–308.
- [5] J. Zhang, M. Small, Complex network from pseudoperiodic time series: topology versus dynamics, *Phys. Rev. Lett.* 96 (2006) 238701.
- [6] J. Zhang, J. Sun, X. Luo, K. Zhang, T. Nakamura, M. Small, Characterizing pseudoperiodic time series through the complex network approach, *Physica D* 237 (2008) 2856–2865.
- [7] Y. Yang, H. Yang, Complex network-based time series analysis, *Physica A* 387 (2007) 1381–1386.
- [8] L. Lacasa, B. Luque, F. Ballesteros, J. Luque, J. Nuño, From time series to complex networks: the visibility graph, *PNAS* 105 (2008) 4972–4975.
- [9] L. Lacasa, B. Luque, J. Luque, J.C. Nuño, The visibility graph: a new method for estimating the Hurst exponent of fractional Brownian motion, *Europhys. Lett.* 86 (2009) 30001.
- [10] L. Lacasa, R. Toral, Description of stochastic and chaotic series using visibility graphs, *Phys. Rev. E* 82 (2010) 036120.
- [11] Y. Yang, J. Wang, H. Yang, J. Mang, Visibility graph approach to exchange rate series, *Physica A* 388 (2009) 4431–4437.
- [12] X.H. Ni, Z.-Q. Jiang, W.-X. Zhou, Degree distributions of the visibility graphs mapped from fractional Brownian motions and multifractal random walks, *Phys. Lett. A* 373 (2009) 3822–3896.
- [13] W.-J. Xie, W. Zhou, Horizontal visibility graphs transformed from fractional Brownian motions: topological properties versus Hurst index, *Physica A* 390 (2011) 3592–3601.
- [14] G. Gutin, T. Mansour, S. Severini, A characterization of horizontal visibility graphs and combinatorics on words, *Physica A* 390 (2011) 2421–2428.
- [15] X. Xu, J. Zhang, M. Small, Superfamily phenomena and motifs of networks induced from time series, *PNAS* 105 (2008) 19601–19605.
- [16] Z. Gao, N. Jin, Complex network from time series based on phase space reconstruction, *Chaos* 19 (2009) 033137.
- [17] N. Marwan, M.C. Romano, M. Thiel, J. Kurths, Recurrence plots for the analysis of complex systems, *Phys. Rep.* 438 (2007) 237–329.
- [18] R.V. Donner, Y. Zou, J.F. Donges, N. Marwan, J. Kurths, Recurrence networks—a novel paradigm for nonlinear time series analysis, *New J. Phys.* 12 (2010) 033025.
- [19] Y. Zou, et al., Power-laws in recurrence networks from dynamical systems, *Europhys. Lett.* 98 (2012) 48001.
- [20] S. Guan, X. Wang, K. Li, B.-H. Wang, C.-H. Lai, Synchronizability of network ensembles with prescribed statistical properties, *Chaos* 18 (2008) 013120.
- [21] J.A. Almendral, R. Criado, I. Leyva, J.M. Buldú, I. Sendiña-Nadal, Introduction to focus issue: mesoscales in complex networks, *Chaos* 21 (2011) 016101.
- [22] J. Reichardt, R. Alamino, D. Saad, The interplay between microscopic and mesoscopic structure in complex network, *PLoS One* 6 (2011) e21282.
- [23] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, U. Alon, Network motifs: simple building blocks of complex networks, *Science* 298 (2002) 824–827.
- [24] A. Vázquez, R. Dobrin, D. Sergi, J.-P. Eckmann, Z. Oltvai, A.-L. Barabási, The topological relationship between the large-scale attributes and local interaction patterns of complex networks, *PNAS* 101 (2004) 17940–17945.
- [25] R. Sinatra, D. Condorelli, V. Latora, Networks of motifs from sequences of symbols, *Phys. Rev. Lett.* 105 (2010) 178–702.
- [26] C. Liu, W.-X. Zhou, Superfamily classification of nonstationary time series based on DFA scaling exponents, *J. Phys. A* 43 (2010) 495005.