Complex Network from Pseudoperiodic Time Series: Topology versus Dynamics

J. Zhang and M. Small
Department of Electronic and Information Engineering, Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong
(Received 14 March 2006; published 14 June 2006)

We construct complex networks from pseudoperiodic time series, with each cycle represented by a single node in the network. We investigate the statistical properties of these networks for various time series and find that time series with different dynamics exhibit distinct topological structures. Specifically, noisy periodic signals correspond to random networks, and chaotic time series generate networks that exhibit small world and scale free features. We show that this distinction in topological structure results from the hierarchy of unstable periodic orbits embedded in the chaotic attractor. Standard measures of structure in complex networks can therefore be applied to distinguish different dynamic regimes in time series. Application to human electrocardiograms shows that such statistical properties are able to differentiate between the sinus rhythm cardiograms of healthy volunteers and those of coronary care patients.

DOI: 10.1103/PhysRevLett.96.238701 PACS numbers: 89.75.Hc, 05.45.Tp, 87.19.Hh

Characterizing complicated dynamics from experimental time series is a fundamental problem of continuing interest in a wide variety of fields. Different measures have been proposed to analyze these dynamics: for example, Lyapunov exponent, entropies, and correlation dimension [1]. In this Letter, we focus on time series with strong pseudoperiodic behavior such as the human electrocardiogram (ECG), human speech, laser output, and annual sunspot numbers. These time series have aroused great interest due to their close relation to some important physiological and natural systems [2,3].

Meanwhile, the past few years have witnessed dramatic advances in the field of complex networks [4] and complex networks have been observed to arise naturally in a vast range of physical phenomena. In this Letter, we show that pseudoperiodic time series can also be investigated from the complex network perspective: the nodes of the network correspond directly to cycles in the time series, and network connectivity is determined by the strength of temporal correlation between cycles. This representation encodes the underlying time series dynamics in the network topology, which may then be quantified via the usual statistical properties of the network. We study noisy periodic and chaotic time series within the above framework. In particular, we seek to characterize chaotic dynamics through the basic statistical properties of the network, such as the degree distribution, average path length, and clustering coefficient. These statistical properties actually reflect and quantify the hierarchy of unstable periodic orbits embedded in the chaotic attractor which leads to small world characteristics. Therefore, this approach provides information that is not available from classical nonlinear time series analysis.

The approach we present in this Letter is a transformation from time domain dynamics to complex network topology. We find that a complex network with small world and scale free characteristics can essentially be considered as the dual of a time series exhibiting chaotic dynamics. Chaotic dynamical systems therefore have a corresponding complex network topological signature.

We start from the construction of the network given a pseudoperiodic time series $\{x_i\}_1^n$ of n observations. First, the pseudoperiodic time series is divided into m disjoint cycles according to the local minimum (or maximum), denoted as $\{C_1, C_2, \ldots, C_m\}$. By considering each cycle as a basic node of a graph, a network representation of a given time series is achieved. Next we determine the connection between nodes. A natural idea is that two nodes are deemed to be connected if the phase space distance between the corresponding cycles is less than a predetermined value D. The phase space distance between C_i and C_j is defined as [5] $D_{ij} = \min_{l=0,1,\ldots,l_j-l_i} \frac{1}{l_i} \sum_{k=1}^{l_i} \|X_k - Y_{k+l}\|$, where X_k and Y_k is the kth point of C_i and C_j (with length l_i and l_j, respectively, and $l_i < l_j$) in the reconstructed phase space. Alternatively, we can also use the linear correlation coefficient ρ between two cycles [5]. The correlation coefficient characterizes the similarity between cycles, since two cycles with a larger temporal correlation will be close in phase space, these two measures are, in fact, inversely proportioned to each other and can be used equivalently. Because the phase space reconstruction is sometimes not reliable for noisy and nonstationary time series, we will use D for the toy models and ρ (which does not need phase space reconstruction) for the experimental time series.

After constructing the complex network from the time series, we investigate its basic statistical properties [4] including: (1) the degree distribution $p(k)$—the of the degree of the nodes in a graph; (2) the average path length L—the average of the minimum number of links necessary to connect all pairs of nodes; (3) the clustering coefficient C—the fraction of connections between the topological neighbors of a vertex with respect to the maximum possible. We will use two typical pseudoperiodic time series with 4000 cycles for comparison and analysis. The first is a noisy periodic time series $y_n = \sin(2\pi\omega n) + b\eta_n$ ($b = 0.2836$), where η is identical and independently distributed noise following $\eta \sim N(0, \sigma^2)$. The second is from the x component of the well-known chaotic Rössler system given by: $x' = -(y + z)$, $y' = x + 0.398y$, $z' = 2 + z(x - 4)$. First, we will show that the complex network built from the noisy periodic time series is a random graph. For simplicity, the distance between cycles is defined as the average of the distance of all pairs of corresponding points in two cycles, i.e., $D_{ij} = \frac{1}{p}\sum_{n=1}^{p}[C_i(n) - C_j(n)]$ (it will be essentially the same if phase space distance is adopted after embedding), where p is the period. Because the periodic components are all the same in each cycle, we can simplify the distance as $D_{ij} = \frac{1}{p}\sum_{i,j=1}^{p}|\eta_i - \eta_j|$. According to the central limit theorem, D_{ij} follows a Gaussian distribution and we denote its as $f(x)$. For a given threshold D, the probability that two nodes are connected is then decided by $p = \int_0^D f(x)dx / \int_0^{+\infty} f(x)dx$, and the resulting network therefore can be deemed as a random graph, with the connection probability p.

Figure 1 plots $p(k)$ for the network extracted from the noisy periodic time series using different thresholds. A curve-fitting exercise indicates that these distributions are essentially Poisson with different $\lambda's$. We examine the clustering coefficient and average path length of the network constructed with $D > D_t = 0.306$. Here D_t indicates the threshold above which the subnetworks crosslink spontaneously to form a single giant network with corresponding ρ_t. We find that the clustering coefficient C scales linearly with p and L is less than 2 for $D > D_t$, i.e., small world behavior typical of a random network is exhibited.

We now look at the x component of the Rössler system. In building the network, we use threshold D for the cycles that are reconstructed from the time series. Figure 2 gives

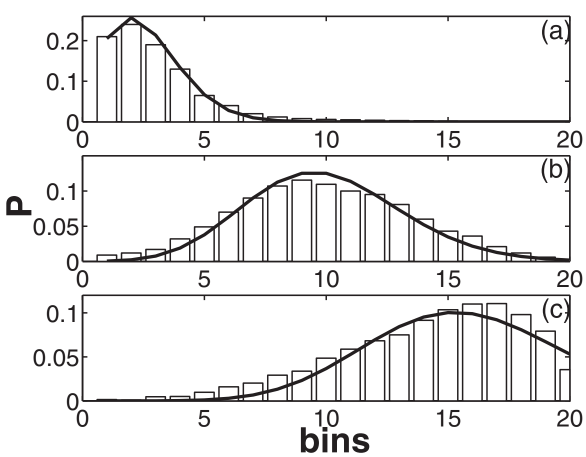

FIG. 1. Degree distribution for sinusoidal signal plus noise at threshold (a) $D = 0.3002$; (b) $D = 0.3190$; (c) $D = 0.3280$, with the mean degree being 424.2, 2004.3, 2894.3, respectively. The original degree [0, 4000] is rescaled to [0, 20] (i.e., 20 bins) in Poisson fitting, with $\lambda = 2.5$, 10, and 15.8 in the 3 cases.

the degree distribution at $D = 0.27$ above $D_t = 0.09$. In contrast to the noisy periodic time series, there are typically multiple peaks in the degree distribution curve. We now consider how these multiple peaks are related to the configurations of the unstable periodic orbits (UPOs) embedded in a chaotic attractor.

Unstable periodic orbits embedded in the chaotic attractor are fundamental to the understanding of the chaotic dynamics [6,7]. It is well known that chaotic attractors are closures of infinitely many UPOs, from which basic ergodic properties such as correlation dimension and Lyapunov exponent can be determined [8].

For a chaotic attractor, its trajectory will typically switch or hop among different UPOs. Specifically, the trajectory will approach an unstable periodic orbit along its stable manifold. This approach can last for several cycles during which the orbit remains close to the UPO. Eventually, the orbit is ejected along the unstable manifold and proceeds until it is captured by the stable manifold of another UPO. A UPO of order n contains n cycles (or n loops) lying in different locations in phase space. Therefore we will see n clusters of cycles distributed in phase space for this UPO-n, with the center of each cluster corresponding to a cycle of UPO-n. The density of cycles around a central cycle of UPO-n depends on the specific stability property at that cycle. Because the complex network construction is based exactly on the phase space distance between cycles, we will have the following conclusions: (1) spatially adjacent cycles in the phase space will also group into a cluster in the network; (2) cycles in each cluster will have approximately the same number of links to the remaining cycles since they are spatially adjacent. Since cycles in one cluster usually have a different number of links from another cluster due to the specific stability properties and phase space location of the central cycle associated with UPO-n, these clusters contribute different peaks to the degree distribution, and the UPO of order n

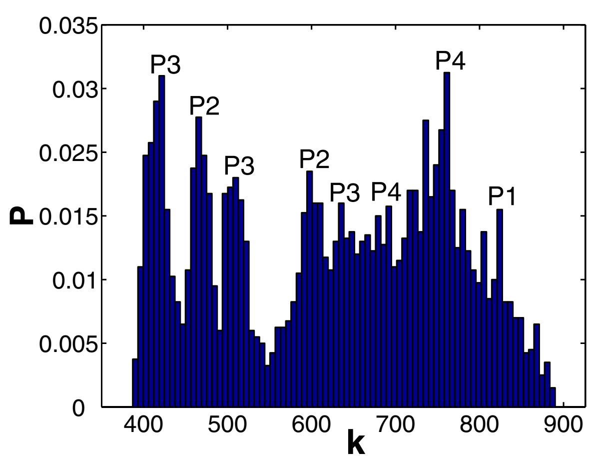

FIG. 2 (color online). Degree distribution for the Rössler system. The peak P_n denotes the accumulation of cycles near UPO of order n.

usually leads to n peaks; see Fig. 2. It should be noted that although a chaotic attractor has, in principle, infinitely many UPOs, we will not see infinitely many peaks due to the finite length of the time series. In fact, the peaks typically correspond to the low order UPOs known as dominant UPOs [9].

The degree distribution obviously relies on the choice of the threshold D. Small D will bring on many small clusters in which cycles surrounding a certain UPO are close enough to be covered by a D sphere, resulting in a large number of peaks in the degree distribution. As D gradually increases, nearby clusters merge to bigger clusters, and in turn to even bigger ones reflected by fewer and fewer peaks. Thus how the degree distribution changes with D (e.g., the 2D degree distribution shown in Fig. 4) represents the hierarchical structure of the UPOs. It quantifies the richness of the UPOs of the chaotic attractor and serves to characterize the chaotic dynamics in terms of the distribution and stability of the UPOs more adequately than a 1D degree distribution.

For a complex network constructed from a chaotic time series we find small world characteristics, i.e., the clustering coefficient obtained at $D_t = 0.09$ is 0.6959, with $L = 14.513$, and at $D = 0.26$ (one tenth of the maximum D between cycles), $C = 0.7849$, with $L = 5.532$. This indicates a strong trend to clustering in the network. Note that cycles in phase space are spatially clustered around the UPOs. Hence they also form clusters in the corresponding network.

We have so far only described binary networks, i.e., two nodes in the graph are either connected or not, and the importance of a node is reflected by its degree. Actually it is possible to generalize this and to consider a fully connected graph, where each pair of nodes are assigned a weight w_{ij}, and the importance of a node is measured by the vertex strength $S_i = \sum_{j \in G} W_{ij}$. Here we choose the weight between each pair of nodes as the distance between corresponding cycles in phase space. We find that for noisy periodic time series, S has a Gaussian distribution; for chaotic time series, S exhibits a power-law distribution. See Fig. 3 for details.

The power-law distribution is reminiscent of scale free networks, for which Barabási and Albert [10] have proposed a model that emphasizes growth and preferential attachment. They show that most networks continuously grow by the addition of new nodes, which are preferentially attached to the existing nodes to a high degree. Interestingly, we find similar phenomena in the networks for chaotic time series. That is, by gradually increasing the length of the time series, the resulting network can also be considered as a growing network. Moreover, the new nodes are found to make preferential attachment to existing nodes of different S, i.e., for nodes with small S, the new nodes will also attach little weight to it. We find that cycles with small S are more stable, and therefore a new cycle is more likely to reside near them. Cycles with small S are also common in the "middle" of the attractor, which makes the distance of new cycles to them generally shorter than the distance to those outlying cycles. Hence we find that new nodes make preferential attachment to existing nodes.

This complex network view of time series reveals interesting connections and yields measures that enhances our understanding of the chaotic dynamics. We apply this analysis to time series of human ECG to demonstrate the method's practicality. We study sinus rhythm electrocardiogram recordings of the coronary care unit patients [11] (patients admitted to coronary intensive care unit, mean age 60, denoted by "P") and healthy volunteers (students, mean age 21, denoted by "H"). Although the subjects are different, the time series corresponds to the same physiological state and is morphologically similar.

In constructing the network, we use the correlation coefficient ρ to determine the connection between nodes, which is more robust to noise and avoids phase space reconstruction. A series of degree distribution curves are obtained by using different threshold ρ's, this is plotted in order in Fig. 4. We call this the two-dimensional degree distribution. We can see that the 2D degree distribution for a coronary care unit patient demonstrates more prominent fluctuations, in comparison to that of the healthy volunteer, which varies rather smoothly. We quantify the level of fluctuations by computing the variance of the normalized derivative of the 2D degree distribution VND [$\mathrm{VND} = \mathrm{var}(DD')$, $DD'(i,j) = [DD(i,j+1) - DD(i,j)]/DD(i,j)$, where DD is the 2D degree distribution, with $DD(i,:)$

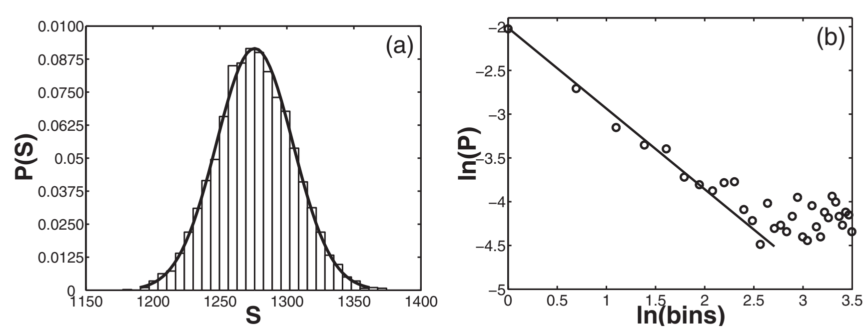

FIG. 3. Vertex strength distribution for (a) Sinusoidal signal plus noise. Mean S is 1276, standard deviation is 30.12. (b) x component of Rössler system. S is rescaled from [3097, 6034] to [0, 50] (i.e., 50 bins), and the slope of the power-law fit is −0.9259.

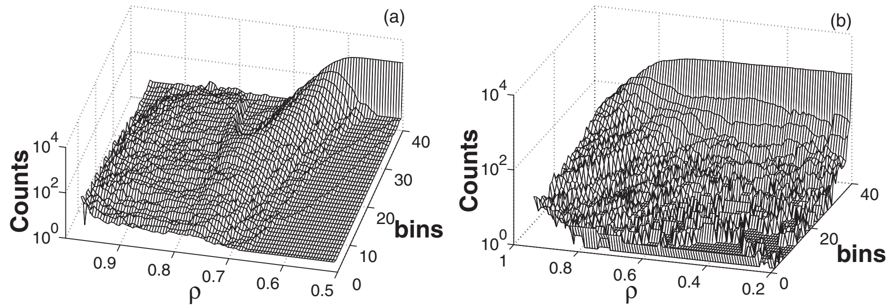

FIG. 4. 2D degree distributions for sinus rhythm cardiograms from (a) a healthy volunteer (H), and (b) a patient with acute risk of sudden cardiac death (P). Both ECGs contain 976 cycles.

being the 1D degree distribution at the ith threshold]. As expected, the VND usually assumes a much higher value for coronary care patients than for volunteers. For the cases shown in Fig. 4, the VND are 2.7119 and 0.4089 for the coronary care patient and the volunteer, respectively.

This result agrees with that of Narayanan *et al.* [12], who have discovered quantitative differences between healthy and pathological groups in terms of UPOs. The authors found that the normal cardiac system is characterized by three to four UPOs, while various pathological conditions contain significantly more UPOs of higher periods. According to our analysis, the time series with more UPOs will lead to more peaks in the degree distribution, and thus a larger VND. In addition, the clustering coefficient and the average path length also show significant difference between the healthy and the coronary care patients. For the cases shown in Fig. 4, we have $C_H = 0.67$, $C_P = 0.98$, and $L_H = 1.5589$, $L_P = 1.0287$ at $\rho = 0.72$ (near ρ_t). This is because the patients have UPOs of large number and higher order. They fill in the chaotic attractor, resulting in more small clusters and therefore a high level of clustering, which is reflected by a higher C and a lower L.

In fact, these statistics reflect the topological properties associated with the number of UPOs as well as their density in phase space. Compared with the classic metric properties, these topological indices are independent of coordinate-system changes and control-parameter variation. This can be of special benefit in ECG analysis, in which the time series are always noisy and nonstationary.

In summary, we have introduced a transformation between temporal dynamics of pseudoperiodic time series and the topological structure of a corresponding complex network. By studying the basic statistical properties of the network, we find useful and interesting connections: the network built from noisy periodic time series corresponds to a random graph, while the chaotic time series typically exhibits small world and scale free features. We attribute this distinction to the unstable periodic orbits that form the skeleton of the chaotic attractor. The basic statistical properties, which reflect the number and distribution of the unstable periodic orbits, have further been used to characterize and quantify the different dynamics of the cardiac behavior associated with UPOs. The results show that these complex network statistics combine to be a powerful tool in differentiating healthy from pathological groups.

We wish to thank D. Walker, G. Chen , K. Zhang, X. Luo, J. Sun, and T. Nakamura for their valuable suggestions. This research was funded by a Hong Kong University Grants Council Grant Competitive Earmarked Research Grant (CERG) number PolyU 5235/03E.

[1] H. D. I. Abarbanel, R. Brown, J. J. Sidorowich, and L. Sh. Tsimring, Rev. Mod. Phys. **65**, 1331 (1993).
[2] M. Small, D. J. Yu, and R. G. Harrison, Phys. Rev. Lett. **87**, 188101 (2001).
[3] M. Small and C. K. Tse, Physica (Amsterdam) **164D**, 187 (2002).
[4] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).
[5] J. Zhang, X. Luo, and M. Small, Phys. Rev. E **73**, 016216 (2006).
[6] D. Auerbach *et al.*, Phys. Rev. Lett. **58**, 2387 (1987).
[7] B. Hunt and E. Ott, Phys. Rev. Lett. **76**, 2254 (1996).
[8] C. Grebogi, E. Ott, and J. A. Yorke, Phys. Rev. A **37**, 1711 (1988).
[9] K. Pawelzik and H. G. Schuster, Phys. Rev. A **43**, 1808 (1991).
[10] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
[11] M. Small, D. J. Yu, N. Grubb, J. Simonotto, K. A. A. Fox, and R. G. Harrison, in *Proceedings of Computers in Cardiology, Boston, 2000* (IEEE, New York, 2000), Vol. 27, p. 355.
[12] K. Narayanan, R. B. Govindan, and M. S. Gopinathan, Phys. Rev. E **57**, 4594 (1998).