Complex network analysis of time series

To cite this article: Zhong-Ke Gao et al. 2016 EPL 116 50001

View the article online for updates and enhancements.

Related content

- Multiscale complex network for analyzing experimental multivariate time series
  Zhong-Ke Gao, Yu-Xuan Yang, Peng-Cheng Fang et al.

- Reconstructing multi-mode networks from multivariate time series
  Zhong-Ke Gao, Yu-Xuan Yang, Wei-Dong Dang et al.

- Multifractality and Laplace spectrum of horizontal visibility graphs constructed from fractional Brownian motions
  Zu-Guo Yu, Huan Zhang, Da-Wen Huang et al.

Recent citations

- Information networks in the stock market based on the distance of the multi-attribute dimensions between listed companies
  Qian Liu et al.

- Testing for the Presence of Correlation Changes in a Multivariate Time Series: A Permutation Based Approach
  Jedelyn Cabrero et al.

- Dynamics and causalities of atmospheric and oceanic data identified by complex networks and Granger causality analysis
  A.K. Charakopoulos et al.
Perspective

Complex network analysis of time series

ZHONG-KE GAO\textsuperscript{1}, MICHAEL SMALL\textsuperscript{2} and JÜRGEN KURTHS\textsuperscript{3,4,5}

\textsuperscript{1} School of Electrical and Information Engineering, Tianjin University - Tianjin 300072, China
\textsuperscript{2} School of Mathematics and Statistics, University of Western Australia - Crawley, WA, 6009, Australia
\textsuperscript{3} Potsdam Institute for Climate Impact Research - Telegrafenberg A 31, 14473 Potsdam, Germany
\textsuperscript{4} Department of Physics, Humboldt University Berlin - 12489 Berlin, Germany
\textsuperscript{5} Institute for Complex Systems and Mathematical Biology, University of Aberdeen - Aberdeen AB24 3UE, UK

received 16 December 2016; accepted in final form 22 December 2016
published online 20 January 2017

PACS 05.45.Tp - Time series analysis
PACS 64.60.aq - Networks
PACS 89.75.-k - Complex systems

Abstract – Revealing complicated behaviors from time series constitutes a fundamental problem of continuing interest and it has attracted a great deal of attention from a wide variety of fields on account of its significant importance. The past decade has witnessed a rapid development of complex network studies, which allow to characterize many types of systems in nature and technology that contain a large number of components interacting with each other in a complicated manner. Recently, the complex network theory has been incorporated into the analysis of time series and fruitful achievements have been obtained. Complex network analysis of time series opens up new venues to address interdisciplinary challenges in climate dynamics, multiphase flow, brain functions, ECG dynamics, economics and traffic systems.

Introduction. – Characterizing dynamical processes in a time-dependent complex system from observed time series of just one or more variables is a fundamental problem of significant importance in many fields ranging from physics and chemistry to economy and social science. Different time series analysis methods have been developed to fulfill this challenging task, e.g., chaos analysis \cite{1}, fractal analysis \cite{2,3}, recurrence plot \cite{4}, complexity measure \cite{5}, multiscale entropy \cite{6}, and time-frequency representation \cite{7}. Time series analysis has been broadly adopted in scientific research and engineering applications. Many theoretical developments for time series analysis have significantly contributed to the understanding of complex systems. However, when system complexity increases, it becomes difficult to describe the dynamical behavior from time series and traditional time series analysis methods have difficulty in coping with the specific burdens of this increased complexity. During the last decade, a new multidisciplinary methodology using complex network has emerged for characterizing complex systems \cite{8–26}. Charting the interactions among system components, abstracted as nodes and edges, has allowed us to represent a complex system as a complex network and then assess the system in terms of network theory. Recently, several novel methodologies have been proposed to map a univariate/multivariate time series into a complex network. These methods have been applied to address interdisciplinary challenges and have already proven great potential for characterizing important properties of complex dynamical systems. The literature on complex network analysis of time series is growing at a very fast rate due to its wide applications in a large variety of research fields.

Complex network analysis of univariate time series. – We start from a time series $x_t$, $t = 1, 2, 3, \ldots, N$ where $N$ is the length of the measured record. Zhang and Small \cite{27} were the first to construct a complex network from a pseudoperiodic time series. They divided a pseudoperiodic time series into disjoint cycles according to the local minimum (or maximum), and then constructed a network by regarding each cycle as a node and determining the connection between nodes in terms of phase space distance between the corresponding cycles. Xu and Small \cite{28} developed an intriguing method for inferring complex network from $x_t$ and indicated that
the network motif distribution allows characterizing different types of \( x_t \). Yang et al. [29] proposed a procedure for constructing complex networks from the correlation matrix of \( x_t \). They first divided a time series into many segments and then constructed a complex network by regarding each segment as a point and determining the edge between nodes in terms of the Pearson correlation coefficient. Gao et al. [30,31] proposed methods to construct complex networks from experimental flow signals and employed network motifs to uncover the nonlinear flow behaviors underlying two-phase flows. In addition, Gao et al. [32] developed a directed weighted complex network to characterize chaotic dynamics from time series. The structures of complex networks constructed via phase space reconstruction from three chaotic systems are shown in fig. 1.

Another basic approach is viewing a time series as a landscape (fig. 2) leading to a visibility graph (VG) [33] which is an efficient and fast method for constructing a complex network from a time series. Figure 2 presents a scheme of the visibility algorithm. In the upper zone we plot the first 20 values of a periodic series by using vertical bars. Considering this as a landscape, one can link every bar (every point of the time series) with all those that can be seen from the top of the considered one (gray lines), obtaining the associated graph (shown in the lower part of the figure). In this graph, every node corresponds, in the same order, to series points in fig. 1.

Fig. 2: Example of a time series (20 data values) and the associated graph derived from the visibility algorithm. In the graph, every node corresponds, in the same order, to series data. The visibility rays between the data define the links connecting nodes in the graph. Figure from ref. [33].

limited penetrable distance is set to be \( L \), a connection between two nodes exists if the number of in-between nodes that block the horizontal line is no more than \( L \). The LPVG and MLPHVG not only inherit the merits of VG but also present a good anti-noise ability, which render LPVG and MLPHVG particularly useful for analyzing real signals polluted by unavoidable noise. The LPVG method has been successfully implemented to analyze many real signals from different fields, e.g., experimental flow signals [46], EEG signals [47–49], electromechanical signals [50]. Moreover, many developments of VG have been proposed, e.g., [51–53].

A third important network approach for analyzing time series is the recurrence network (RN), which has undergone an explosive growth in recent years [54–58]. In a RN, \( R_{i,j} = \Theta(\varepsilon - \| X(i) - X(j) \|) \), \( i \neq j \) where \( \Theta \) is the Heaviside function. \( X(i) \) is a phase space vector \( i = 1, 2, \ldots, N \), and \( N \) is the number of phase space vectors. The individual phase space vector \( X(i) \) serves as a node and the existence of an edge indicates the occurrence of a recurrence, i.e., the distance measure \( R_{i,j} \) between a pair of nodes in the phase space is smaller than a threshold \( \varepsilon \). RN has been successfully applied to many fields, e.g., EEG data [59], cardiovascular data [60], turbulent heated jets [61], multiphase flow system [62–65], climate system [66–68]. To give an application example, Marwan et al. [68] demonstrated that the RN allows identifying qualitative transitions in observational data, e.g., when analyzing paleoclimate regime transitions. They defined a network divergence \( \Delta S_i = S_i^n - S_i^o \) as the difference between in- and out-strength and found that negative values of network divergence \( \Delta S \) indicate the source regions of extreme events whereas positive values indicate sinks. They analyzed RN from extreme rainfall during Austral winter season, in fig. 3(a), and found negative network divergence values within the southeastern South America region, indicating that this region is a source region of extreme rainfall although it is one of the exit regions of the low-level moisture flow from the Amazon region. For the southeastern South America region, they found high values of impact not only in the direct vicinity of southeastern South America but also at the eastern slopes of the Central
Andes (fig. 3(b)), which suggests that extreme rainfall at the Central (in particular, Bolivian) Andes will precede after rainfall events in the southeastern South America region.

There are also many other methods for mapping a time series into a complex network, such as the methods based on stochastic processes [69], coarse geometry theory [70], nonlinear mutual information [71], or event synchronization [72]. Recently, McCullough and co-workers [73] have used ordinal codings to construct networks from time series. The time series is first replaced by a sequence of permutations of integers, where each permutation encodes the shape of the time series waveform in a short window. These permutations then map to nodes of a complex network and nodes are connected if the corresponding permutations occur in succession. Effectively these permutation methods provide a robust partitioning of state space.

**Wide applications of complex network in real data analysis.** – Complex network analysis of time series has been widely used to solve challenging problems in different research fields. Liu et al. [74] employed VG to infer complex networks from time series of energy dissipation rates in three-dimensional fully developed turbulence. Zhuang et al. [75] used VG to analyze financial time series and identified important historical incidents that influenced market integration coincide with variations in the measured graphical node degree. Yang et al. [76] proposed a novel VG-based method and applied this method to the analysis of empirical records for stock markets in USA, in which series segments are mapped to VG as descriptions of the corresponding states and the successively occurring states are linked. Qian et al. [77] employed VG to investigate 30 world stock market indices and a universal allometric scaling law was uncovered in the minimal spanning trees, whose scaling exponent is independent of the stock market and the length of the stock index. Ahmadlou et al. [52] improved the VG to analyze EEG signals associated with autistic children. Gao et al. [53] proposed an adaptive optimal kernel VG, which combines the advantages of VG and adaptive optimal-kernel time-frequency representation, and applied it to characterize the EEG recordings associated with epileptic seizures. Subramaniyam et al. [59] proposed the application of randomness and nonlinear independence test based on RN measures to distinguish between the dynamics of focal and nonfocal EEG signals. Ramírez et al. [60] used RN to distinguish pregnancies already in the second trimester, using the cardiovascular time series including the variability of heart rate and systolic and diastolic blood pressures. Charakopoulos et al. [61] used a phase space reconstruction and VG to analyze experimental temperature time series from a vertical turbulent heated jet and suggested that the complex network approach allows distinguishing, identifying, and exploring in detail various dynamical regions of the jet flow. Telesca et al. [78] used VG to investigate the seismicity of Italy between April 16, 2005 and December 31, 2010. Jiang et al. [79] used VG to construct complex networks from heartbeat interval time series and investigated the statistical properties of the network before and during chi and yoga meditation. Meenatchidevi et al. [80] investigated the scale invariance of combustion noise generated from turbulent reacting flows in a confined environment using complex networks. Zou et al. [81] employed VG to analyze the daily and monthly sunspot series and got some new insights. Tang et al. [82] constructed and analyzed complex networks from traffic time series and indicated that complex network is a practical tool for exploring dynamics in traffic time series. Banerjee et al. [83] presented a fresh and broad yet simple approach towards information retrieval in general and diagnostics in particular by applying the theory of complex networks on multidimensional, dynamic images. Doza et al. [84] constructed and analyzed global climate complex networks. Gao et al. [85] proposed an approach combining econometrics and complex network theory to explore the transmission mechanism of forex burden fluctuant patterns. Ghaffari et al. [86,87] analyzed the dynamics of static friction, i.e., nucleation processes, with respect to friction networks, and indicated that complex networks can successfully capture the crack-like shear ruptures and possible corresponding acoustic features.

Brain network represents an important application of complex network analysis of time series. The human brain, as one of the most complicated and complex systems in nature, is an open, dissipative, and adaptive dynamical system with immense functionality, which can be regarded as a network with lots of interacting subsystems. There are so many related works in brain network analysis, and we here cite a few as examples as follows: Chavez et al. [88] analyzed the connectivity structure of weighted brain networks extracted from spontaneous magnetoencephalographic signals of healthy subjects and epileptic patients (suffering from absence seizures) recorded at rest and suggested that modularity plays a key role in the functional organization of brain areas during normal and pathological

---

Fig. 3: (Colour online) (a) RN divergence of extreme rainfall network during Austral winter season. Negative values indicate source and positive values indicate sink regions of extreme rainfall; (b) impact of a region (here the southeastern South America, marked by the box) in form of contributing propagated extreme rainfall. Figure from ref. [68].
networks. Liu et al. [90] explored abnormal functional magnetic resonance imaging (fMRI) resting-state dynamics, functional connectivity, and weighted functional networks, in a sample of patients with severe Alzheimer’s disease and age-matched healthy volunteers. Using network analysis of DTI data from healthy volunteers, and meta-analyses of published MRI studies in 26 brain disorders, Crossley et al. [91] indicated that lesions across disorders tend to be concentrated at hubs. Gong et al. [92] summarized the current findings and historical understanding of structural and functional connectomes in depression, focusing on graph analyses of depressive brain networks. Dai et al. [93] employed resting-state fMRI data and voxel-based graph-theory analysis to systematically investigate intrinsic functional connectivity patterns of whole-brain networks in Alzheimer’s disease patients and healthy controls. Using simultaneous EEG/functional MRI and functional MRI/DTI data, Zhang et al. [94] identified the neural and anatomical basis of variability in regional functional architecture, and revealed diseasespecific changes. Their findings shed light on the dynamic organization of normal and disordered brain networks. Liang et al. [95] employed graph-based modularity analysis to identify the default-mode network, executive-control network, and salience network during an N-back working memory (WM) task and further investigated the modulation of intra- and inter-network interactions at different cognitive loads. Lehnertz et al. [96] well reviewed analysis techniques for human epileptic brain networks and summarized recent findings derived from studies investigating these networks. More examples and methods for inferring brain networks can be seen in ref. [96]. It is believed that the brain network analysis will advance the understanding of dynamical diseases and may guide new developments for diagnosis, treatment, and control.

**Multi-information fusion in complex network.**—Besides the above network inference methods and their applications, more recently, the multi-information fusion in complex network gradually has become a hot topic, especially in the nowadays Big Data time. The complex network methods dedicated to multivariate time series analysis are drawing increasing attention on account of their significant importance. Kramer et al. [97] proposed a method for inferring complex networks from multivariate time series, which yielded as output both the inferred network and a quantification of uncertainty in the number of edges. Jachan et al. [98] introduced the nonparametric partial directed coherence that allows disentanglement of direct and indirect connections and their directions in a network. Gao et al. developed the multivariate RN [62,63] and multivariate weighted RN [99] to analyze multivariate time series, and then applied them to characterize nonlinear flow behaviors underlying multiphase flow. Lucas et al. [100] presented a non-parametric method to analyze multivariate time series, based on the mapping of a multidimensional time series into a multilayer network, which allows to extract information on a high dimensional dynamical system through the analysis of the structure of the associated multiplex network. Nakamura et al. [101] described a method for constructing networks from multivariate nonlinear time series, in which each time series is regarded as a node and the connection between nodes is determined in terms of small-shuffle surrogate method.

More recently, three novel methodologies [102–104] have been proposed for realizing multi-information fusion in a complex network. In particular, Gao et al. [102] proposed a modality transition-based network for mapping the experimental multivariate measurements into a directed weighted complex network. The basic idea is the following: For a multivariate time series, e.g., four experimental signals of equal length of $L$, we use a moving window that contains four sub-time series to partition the four time series. In particular, the moving window slides with the time from left to right by a step of $S$ and the length of each window is $W$. Thus we can partition the four time series into $(L - W + S)$ windows. Within each window, we first calculate the Pearson correlation coefficient for each pair of sub-time series. We thus obtain the correlation coefficient $r_{12}, r_{13}, r_{14}, r_{23}, r_{24}, r_{34}$, respectively, and denote each correlation coefficient as an element of the modality as follows: $r_{12} \rightarrow A, r_{13} \rightarrow B, r_{14} \rightarrow C, r_{23} \rightarrow D, r_{24} \rightarrow E, r_{34} \rightarrow F$. Then we rank the six correlation coefficients in ascending order to get a permutation of $ABCDEFG$, named as a modality, for each window. The largest number of modality is the permutation of $ABCDEFG$, i.e., $6! = 720$. We represent each modality as a node and determine the edge according to the transition of modality. In particular, node $i$ (e.g., modality $FBCEAD$) and node $j$ (e.g., modality $FEBCAD$) are connected if the modality $FBCEAD$ changes to modality $FEBCAD$ in one time step and the direction is from node $i$ to node $j$. A weight $w_{ij}$ of an edge is the number of times for the transition from modality $i$ to modality $j$.

![Fig. 4](image-url)
Note that self-connections of a node are excluded. Repeated transitions between modalities lead to edges with large weights. Figure 4 demonstrates the construction of a modality transition-based network from multivariate time series.

Gao et al. [103] also developed a multiscale complex network and clustering coefficient entropy for analyzing multivariate time series. The basic procedure for constructing a multiscale complex network from multivariate time series is as follows: For a multivariate signals that contains $p$ sub-signals of equal length $L = \{x_{k,i}\}_{k=1}^{L-1}$, $k = 1, 2, \ldots, p$, we first perform a coarse-grain process to define temporal scales and further obtain multivariate coarse-grained signals as follows:

$$y^s_{k,j} = \frac{1}{s} \sum_{i=(j-1)s+1}^{js} x_{k,i}$$

where $s$ is the scale factor and $1 \leq j \leq \lfloor \frac{L}{s} \rfloor$, $k = 1, 2, \ldots, p$. Then we use multivariate embedding theory to construct a complex network from each obtained $y^s_{k,j}$, i.e., constructing a complex network at a different scale factor $s$. Specifically, we perform the multivariate phase space reconstruction on the $\{x_{k,i}\}_{k=1}^{L-1}$, $k = 1, 2, \ldots, p$ as follows:

$$X_m(i) = [x_{1,i}, \ldots, x_{1,i+(m-1)\tau}, x_{2,i}, \ldots, x_{2,i+(m-1)\tau}, \ldots, x_{p,i}, \ldots, x_{p,i+(m-1)\tau}],$$

where $\tau = [\tau_1, \tau_2, \ldots, \tau_p]$ and $M = [m_1, m_2, \ldots, m_p] \in \mathbb{R}^p$ is the vector of time delay and vector of embedding dimension, respectively, and $X_m(i) \in \mathbb{R}^m$ ($m = \sum_{k=1}^{p} m_k$).

Then we can infer multiscale complex networks from multivariate signals $\{x_{k,i}\}_{k=1}^{L-1}$, $k = 1, 2, \ldots, p$ by the following steps: a) we produce $(L-n)$ composite delay vectors $X_m(i) \in \mathbb{R}^m$, where $n = \max\{M\} \times \max\{\tau\}$ and $i = 1, 2, \ldots, L-n$. b) We define the phase space distance between any two vectors $X_m(i)$ and $X_m(j)$, $j \neq i$ in terms of maximum norm

$$d[X_m(i), X_m(j)] = \max_{l=1,\ldots,m} |x_{(i+l-1)} - x_{(j+l-1)}|.$$ 

c) We then regard each phase space vector as a node and use the phase space distance to determine the edges for constructing a complex network. By choosing a threshold, we obtain the adjacency matrix $A$ of the complex network: An edge between node $i$ and $j$ exists ($A_{ij} = 1$) if the phase space distance between them is smaller than the threshold; while node $i$ and $j$ are not connected ($A_{ij} = 0$) otherwise. The topology of the derived complex network at different scales is determined entirely by the adjacency matrix $A$. d) Finally, we obtained the multiscale complex networks by performing steps a)–c) on each coarse-grained multivariate signals. The threshold can be determined by the percentage (i.e., 15%) of total variation $T_2(S)$, where $S$ is the covariance matrix of the multivariate signals. Gao et al. applied their proposed multiscale complex network to analyze multi-channel measurements from gas-liquid flows and suggested that multiscale complex networks allow quantitatively revealing the nonlinear flow behavior governing the transitions of flow patterns from the perspective of multiscale analysis and complex network analysis. Moreover, Gao et al. proposed a multi-frequency complex network for analyzing multivariate time series and implemented it to uncover the flow structures undergoing horizontal oil-water flows. The detected community structures of the multi-frequency complex networks for five horizontal oil-water flow patterns are shown in fig. 5. The results allow us to recognize different horizontal oil-water flow patterns and further pave a way for realizing network visualization of complex flow patterns from a community structure perspective.

Multiphase flow is commonly observed in many industrial applications and its behaviors under a wide range of flow conditions and inclination angles constitute an outstanding interdisciplinary problem [105,106]. Due to the interplay among many complex factors such as fluid turbulence, phase interfacial interaction, local relative movements between phases, multiphase flow exhibits highly irregular, random, and unsteady flow structures and thus represents a typical complex system. The above results render the complex network as a powerful tool for characterizing the complicated flow behaviors underlying different multiphase flow systems from multi-channel measurements. These results also deepen the understanding about the flow mechanism governing the formation and evolution of different multiphase flow patterns.

Fig. 5: (Colour online) Community structures of the multi-frequency complex networks for five horizontal oil-water flow patterns. (a) Stratified flow pattern (ST); (b) stratified flow with mixing at an interface pattern (SKMIF); (c) dispersion of oil in water and water flow pattern (D O/W&W); (d) dispersion of water in oil and oil in water flow pattern (D W/OkD O/W); (e) dispersion of oil in water flow pattern (D O/W). The networks from different flow patterns exhibit distinct community structures and the community structures faithfully represent the structural features of different flow patterns. Figure from ref. [104].
Conclusions. — In summary, we have reviewed the current status in complex network analysis of time series, and presented some often used but also newly developed methodologies. Broader applicability of these novel methods have been demonstrated and articulated. Complex network analysis that originates from graph theory has undergone a brilliant development in the past decade, and it has contributed significantly to the understanding of complex systems. Complex network analysis of time series brings us a new analytical framework for characterizing complicated behavior from observational data. This will have a strong impact on a number of applications because abundant observational data exist in many research fields. Furthermore, new challenges for complex network analysis of time series appear in different fields. For example, in the brain network analysis, how to efficiently define the brain functional connectivity and then infer brain network from EEG recordings or fMRI data to probe the underlying complicated dynamics still represents a significant challenge. In the study of multiphase flow, new complex network methods are required to realize the multi-information fusion for coping with the increased complexity underlying the new multiphase flow system. These challenges will facilitate the generations of new methods for mapping time series into a complex network and correspondingly the characterization of various complex networks from different real data will bring out new network topological measures and then promote the topological analysis of network structure. Complex network analysis of time series could open up new venues for cross-fertilization between network theory and time series analysis, and also provide a fascinating analytical framework for addressing interdisciplinary challenges.

***

This work was supported by National Natural Science Foundation of China under Grant No. 61473203, and the Natural Science Foundation of Tianjin, China under Grant No. 16JCYBJC18200.

REFERENCES

[38] Ravetti M. G. et al., PLoS ONE, 9 (2014) e108004.