


Communication

Complex Network Construction of Univariate Chaotic Time Series Based on Maximum Mean Discrepancy

Jiancheng Sun 

School of Software and Internet of Things Engineering, Jiangxi University of Finance and Economics, Nanchang 330013, China; sunjc@jeufe.edu.cn; Tel.: +86-0791-8384-5702

Received: 11 December 2019; Accepted: 23 January 2020; Published: 24 January 2020



Abstract: The analysis of chaotic time series is usually a challenging task due to its complexity. In this communication, a method of complex network construction is proposed for univariate chaotic time series, which provides a novel way to analyze time series. In the process of complex network construction, how to measure the similarity between the time series is a key problem to be solved. Due to the complexity of chaotic systems, the common metrics is hard to measure the similarity. Consequently, the proposed method first transforms univariate time series into high-dimensional phase space to increase its information, then uses Gaussian mixture model (GMM) to represent time series, and finally introduces maximum mean discrepancy (MMD) to measure the similarity between GMMs. The Lorenz system is used to validate the correctness and effectiveness of the proposed method for measuring the similarity.

Keywords: complex network; chaotic time series; Gaussian mixture model; maximum mean discrepancy

1. Introduction

Chaotic time series exist widely in many fields, such as economics, physics, hydrology and so on [1]. In chaotic systems, the “butterfly effect” is a typical phenomenon in which small causes can have large effects [2]. Therefore, a chaotic system usually has highly complex behaviours, and the relevant analysis is a challenging task.

In recent years, the application of complex network theory to time series analysis is increasing rapidly. Firstly, the time series was transformed into a network, and then, various complex network tools were used for analysis [3–7]. There are three kinds of network reconstruction methods: recurrence network based on phase space and visibility graphs and transition network based on Markov chain [6]. Regardless of the network construction method, a key problem to be solved is how to measure the similarity between nodes. For example, in order to measure the similarity between nodes, Euclidean distance, visual distance, and transition probability were applied to recurrence network [8], visibility graphs [9], and transition network [10], respectively.

In this communication, we focus on the construction of a complex network of univariate chaotic time series, which is an effective way to analyse the time series [11]. Similarly, in this task, a core problem to be solved is how to measure the similarity between time series. In the community of time series analysis, some commonly used metrics, such as Euclidean distance [12], correlation coefficient [13], and dynamic time warping distance (DTW) [14], were used to measure similarities between time series. Especially for DTW, its outstanding advantage lies in the ability to find the optimal nonlinear alignment between two given sequences. However, most of the metrics cannot effectively measure the similarity in the case of chaotic time series. For example, time series from the same chaotic system are completely different in the form of local sub-sequences, which makes the pair matching metric, such as Euclidean distance, unable to measure the similarity between them. Even with statistical metrics such

as the correlation coefficient, due to the rich structure of chaotic system, its effect is not as expected. In addition, the probability distribution of chaotic time series usually presents mixed distribution [15], which also brings challenges to measuring the similarity between sequences using statistical distances.

Considering the characteristics of chaotic time series mentioned above, in this communication, we improved the performance of similarity measurement from two aspects: (1) transform univariate time series (UTS) into a high-dimensional space to describe time series more accurately; (2) in the high-dimensional space, Gaussian mixture model (GMM) is used as the representation of time series, and distance metric is introduced to catch the similarity between GMMs.

2. Approach of Constructing Complex Networks

In the complex network constructed, nodes represent the time series themselves, and the edges between nodes are determined by the strength of similarity between the time series. The process of constructing a network can be divided into the following sections.

2.1. Representation of Univariate Chaotic Time Series

We can realize the representation of UTS by using the idea of phase space reconstruction. Although a complex system is usually described by multiple variables, in most cases we can only observe a univariate (scalar) time series $T = \{x_1, x_2, \dots, x_n\}$ from the system, where n is the length of the time series. Fortunately, using the embedded theorem [16], we can reconstruct the original space of the system by unfolding the scalar time series into higher dimensional phase space. With the help of phase space reconstruction, we can investigate the geometric and dynamic properties of the original phase space as well as unobserved variables. In other words, it provides a new approach, which can transform UTS into state vectors in higher dimensional space, so as to describe and understand the characteristics of the system more accurately. This is exactly the motivation for representation of univariate time series.

By choosing an appropriate embedding dimension m and a time delay τ , we can transform a UTS $T = \{x_1, x_2, \dots, x_n\}$ into a state vector in the phase space as

$$x_i = (x_i, x_{i+\tau}, x_{i+2\tau}, \dots, x_{i+(m-1)\tau})^T \quad (1)$$

where m can also be regarded as the number of variables in the original phase space. Therefore, the phase space can be described by a $m \times (n - \tau)$ matrix X , where each column represents the state point x_i at time i , and each row represents a subsequence of the UTS. On the other hand, each row in X can also be viewed as observation of a variable. Consequently, X is multivariable time series (MTS) with m variables, which is converted from the UTS.

To illustrate the phase space reconstruction clearly, the well-known Lorenz system is used as an example to illustrate the reconstruction process [2]. The Lorenz system is described by three ordinary differential equations:

$$\begin{aligned} \frac{dx}{dt} &= \sigma(y - x), \\ \frac{dy}{dt} &= x(\rho - z) - y, \\ \frac{dz}{dt} &= xy - \beta z \end{aligned} \quad (2)$$

where x , y , and z are system variables; t is the time, and σ , ρ , β are the system parameters. Two time series T_1 and T_2 (x component of the system) shown at the top of Figure 1a are generated using (2) with $\sigma = 8$, $\rho = 28$ and $\beta = 8/3$. All system parameters are kept the same here, except that the initial conditions for generating T_1 and T_2 are slightly different by 10^{-2} . The difference between the two-time series is shown at the bottom of Figure 1a. It can be seen from Figure 1a, in the beginning, the two-time series kept the same shape, but as time went on, the differences became larger and larger, and this phenomenon is known as the "butterfly effect". In other words, although the two time series are generated from the same system with slightly different initial values, they are very different in local

characteristics. Therefore, it is usually not feasible to measure the similarity between chaotic time series by general metric. In Figure 1b, pairs of time series values $x_i = (x_i, x_{i+\tau})^T$ are plotted with black dots; this is the state vector in the reconstructed phase space of the x component of T_1 using (1). In other words, a system with two variables is reconstructed from the observation of one variable (a scalar time series). As can be seen from Figure 1b, more abundant geometric structure of the time series can be seen by using the phase space reconstruction, thus it can provide more information about the time series.

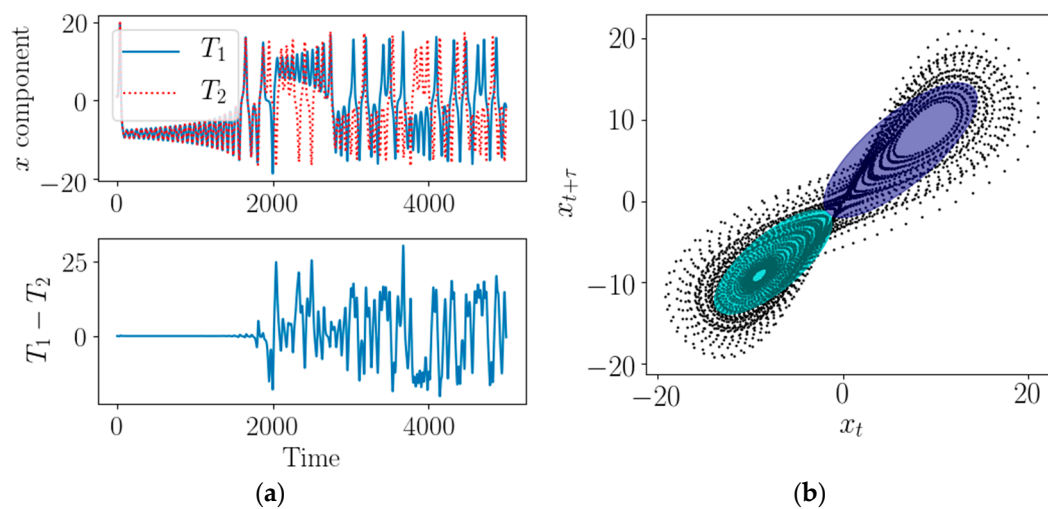


Figure 1. Chaotic time series of Lorenz system and its reconstructed phase space. (a) x component of Lorenz system. (b) Reconstructed phase space of T_1 .

When the phase space is constructed, the next problem to be solved is how to represent the time series in the space. As mentioned above, the general point-to-point metrics are hard to measure the similarity between chaotic time series due to the complexity. Therefore, it is a more reasonable choice to calculate the statistical characteristics of time series and then measure the similarity between them. An intuitive way is to estimate the covariance matrix of the multivariable time series (MTS) in phase space and then calculate geodesic distance between them [17]. However, considering the complex structure of MTS in phase space, a single covariance matrix usually cannot accurately describe the data distribution of a chaotic system. For example, in Figure 1b, at least two independent Gaussian distribution models are required to accurately describe the Lorenz systems. Thus, a natural way here is to adopt multivariate GMM to represent MTS generated by chaotic systems. The GMM is given by

$$G = \sum_{i=1}^n \alpha_i g_i = \sum_{i=1}^n \alpha_i \mathcal{N}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \quad (3)$$

where the i -th component is characterized by normal distributions $g_i = \mathcal{N}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$ with weights α_i , means $\boldsymbol{\mu}_i$, and covariance matrices $\boldsymbol{\Sigma}_i$. In other words, GMM is a linear combination of several Gaussian distribution functions. In theory, GMM can fit any type of distribution, which is usually used to solve the case that one data set contains many different distributions. As shown in Figure 1b, GMM with two components is used to model MTS, and it (denoting as the ellipses) can well describe the double scroll structure of Lorenz system.

2.2. Complex Network Construction with Similarity Metric

Once the time series is represented by GMM, the similarity between time series is converted into the similarity between GMMs. Kullback–Leibler divergence is a commonly used solution to measure the distance between two probability distributions. However, it has no closed form solution in the case

of GMM, and the implementation of Monte Carlo simulation becomes computationally expensive [18]. Therefore, we introduce maximum mean discrepancy (MMD) in reproducing kernel Hilbert spaces to quantify the similarity between GMMs [19]. Suppose we have two GMMs in \mathbb{R}^d :

$$\begin{aligned} P &= \sum_{i=1}^m \alpha_i p_i = \sum_{i=1}^m \alpha_i \mathcal{N}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \\ Q &= \sum_{j=1}^n \beta_j q_j = \sum_{j=1}^n \beta_j \mathcal{N}(\boldsymbol{\mu}'_j, \boldsymbol{\Sigma}'_j) \end{aligned} \quad (4)$$

where $p_i = \mathcal{N}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$; $q_j = \mathcal{N}(\boldsymbol{\mu}'_j, \boldsymbol{\Sigma}'_j)$; and m, n is the components number of P and Q , respectively. Given a kernel function $k(x, y) = \langle \varphi(x), \varphi(y) \rangle_{\mathcal{H}}$, the reproducing kernel Hilbert space (RKHS) \mathcal{H} corresponding to $k(x, y)$ can be defined, where $\varphi(x)$ is a feature mapping [20]. Given that we are in an RKHS, the mean map kernel can be defined as

$$K(P, Q) = \mathbb{E}_{x \sim P, y \sim Q} k(x, y) = \langle \mathbb{E}_{x \sim P}[\varphi(x)], \mathbb{E}_{y \sim Q}[\varphi(y)] \rangle \quad (5)$$

Then MMD can be easily defined as

$$\begin{aligned} \text{MMD}(P, Q) &= \|\mathbb{E}_{x \sim P}[\varphi(x)] - \mathbb{E}_{y \sim Q}[\varphi(y)]\| \\ &= \sqrt{K(P, P) + K(Q, Q) - 2K(P, Q)} \end{aligned} \quad (6)$$

In the case of insufficient data, we can approximate the kernel function $K(P, Q)$ by empirical estimation [21]:

$$K_{\text{emp}}(P, Q) = \frac{1}{n_P n_Q} \sum_{i=1}^{n_P} \sum_{j=1}^{n_Q} k(x_i, y_j) \quad (7)$$

where $\{x_i\}_{i=1}^{n_P}$ and $\{y_j\}_{j=1}^{n_Q}$ are random samples. However, the approximation obtained with (7) introduces errors with high probability. Fortunately, when enough data is available, we can estimate the true distribution of the data; when GMM is used to approximate the distribution of the data, $K(P, Q)$ has a closed solution:

$$K(P, Q) = \sum_{i,j} \alpha_i \beta_j K(\mathcal{N}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i), \mathcal{N}(\boldsymbol{\mu}'_j, \boldsymbol{\Sigma}'_j)) \quad (8)$$

With (8), the form of $K(P, P)$ and $K(Q, Q)$ can be derived similarly. It turns out, introducing the Gaussian RBF kernel $k(x, y) = \exp(-\gamma \|x - y\|^2 / 2)$, the product kernel of the Gaussian distribution is derived as:

$$\left(\mathcal{N}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i), \mathcal{N}(\boldsymbol{\mu}'_j, \boldsymbol{\Sigma}'_j) \right) = 1 / |\gamma \boldsymbol{\Sigma}_i + \gamma \boldsymbol{\Sigma}'_j + \mathbf{I}|^{\frac{1}{2}} \exp\left(-\frac{1}{2}(\boldsymbol{\mu}_i - \boldsymbol{\mu}'_j)^T (\boldsymbol{\Sigma}_i + \boldsymbol{\Sigma}'_j + \gamma^{-1} \mathbf{I})^{-1} (\boldsymbol{\mu}_i - \boldsymbol{\mu}'_j)\right) \quad (9)$$

With (6), (8), and (9), we can obtain the analytic form of $\text{MMD}(P, Q)$ by introducing the Gaussian RBF kernel.

Once similarity measures are in place, the construction of complex networks is straightforward. First, each UTS is represented by a GMM, and then, MMD in (6) is used to calculate the distance between each pair of GMM to form a distance matrix $\mathbf{D} = (\text{MMD}(P_i, Q_j))$, where i and j denote different UTS. With a critical threshold r_c , \mathbf{D} can be converted into adjacent matrix whose elements indicate whether pairs of nodes are connected or not in the network. An adjacent matrix $\mathbf{A} = (a(P_i, Q_j))$, here $a(P_i, Q_j) = 1$ if $\text{MMD}(P_i, Q_j) \leq r_c$ and $a(P_i, Q_j) = 0$ if $\text{MMD}(P_i, Q_j) > r_c$.

3. Experiments and Results

The Lorenz system in (2) has highly complex behaviors with the variation of the system parameters. With the change of system parameters, Lorenz system presents highly complex behavior. We randomly

generate 150 time series of x components by keeping $\sigma = 10.0$ and $\beta = 8/3$ while varying $\rho \in [28, 45]$. The reason is that (σ, ρ, β) form a vast three-dimensional parameter space. Considering the complexity of the Lorenz system, its characteristics have not been fully studied when σ and β take other values [22]. To simplify the problem, many researchers fix σ and β , while changing ρ . That is, each set of (σ, ρ, β) corresponds to a UTS and different parameter ρ corresponds to different class of time series. The length of each time series is 6000 data points, and the first 1500 points are removed to reduce the initialization effect of the system. The differential equation (2) is solved by `scipy.integrate.odeint()` in Python package SciPy [23], and the time point step is 0.01.

Firstly, with $m = 3$, $\tau = 8$, each UTS is transformed into MTS in phase space by (1). Then, the GMM corresponding to each MTS is estimated, where the components number is 3. Finally, the MMD between the GMMs is calculated and eventually converted to the adjacency matrix. In addition, to evaluate the proposed method, three other metrics (geodesic distance, DTW and correlation coefficient) are also used to construct the network for comparison. By estimating the covariance matrix of MTS in phase space, the geodesic distance can be obtained and then a network formed [5]. For DTW and correlation coefficient, the metrics can be calculated directly between UTS.

The spring layout method in NetworkX package [24] was used to plot the network, and the results were shown in Figure 2. In the network, each node corresponds to a UTS, and the connection between nodes is determined by the adjacency matrix. The selected threshold r_c enables 20% of the edges to be preserved to highlight the geometric structure of the network. The validity of network construction can be evaluated from two aspects: one is to see whether the similarity between nodes can be effectively captured; the other is to see whether the geometry of the network is conducive to the analysis of time series. In the first aspect, geodesic distance (Figure 2a) and MMD (Figure 2b) are better metrics of similarity because nodes with similar ρ are clustered together. In contrast, in Figure 2c,d, nodes with different ρ are mixed together, especially in Figure 2d, the nodes are completely confused and indistinguishable, like a random network, indicating that the metric used cannot effectively measure similarity. From the second aspect, the MMD is superior to the geodesic distance in the geometry of the network because the nodes in Figure 2a are squeezed together to make it difficult to distinguish. This phenomenon also indicates that MMD is more sensitive to measure similarity, which results in a looser network. In the following description, we will explain why a loose network structure is better than a tight one.

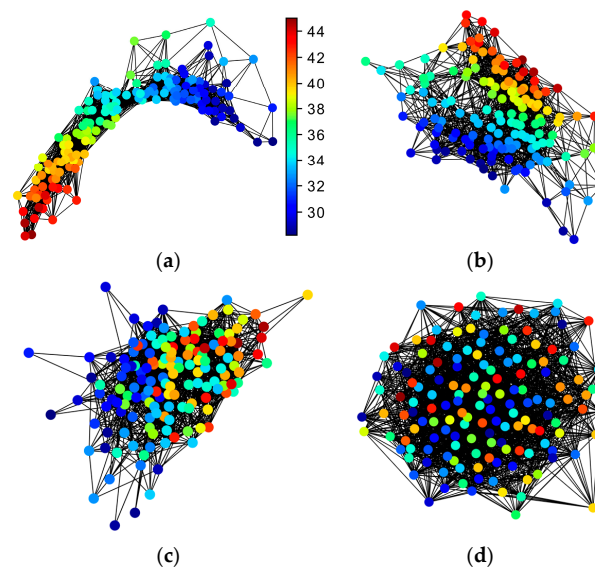


Figure 2. Construction of complex network based on different metric (colour bar denotes the value of ρ). (a) Network construction based on geodesic distance; (b) network construction based on maximum mean discrepancy (MMD); (c) network construction based on dynamic time warping distance (DTW); (d) network construction based on correlation coefficient.

To analyse the characteristics of MMD and geodesic distance in detail, we show the heat map of the related distance matrix in Figure 3, which corresponds to the network structure. The size of the heat map is 150×150 , corresponding to 150 nodes, and each pixel denotes the distance between a pair of nodes (time series). The nodes are arranged in ascending order according to the value of ρ . From Figure 3a,b, it can be seen that the distance near the diagonal is small (high similarity), otherwise the distance is large, which means that the node pairs with similar ρ have high similarity. To investigate this point more clearly, we set a certain threshold r_c and retained the 20% of the edge (mentioned in Section 2.2), as shown in Figure 3c,d. As you can see, the reserved edges are centered diagonally and gradually spread to both sides. Therefore, if more edges are retained by increasing r_c , the topology of the network (the relationship between adjacent nodes) can still remain stable to some extent.

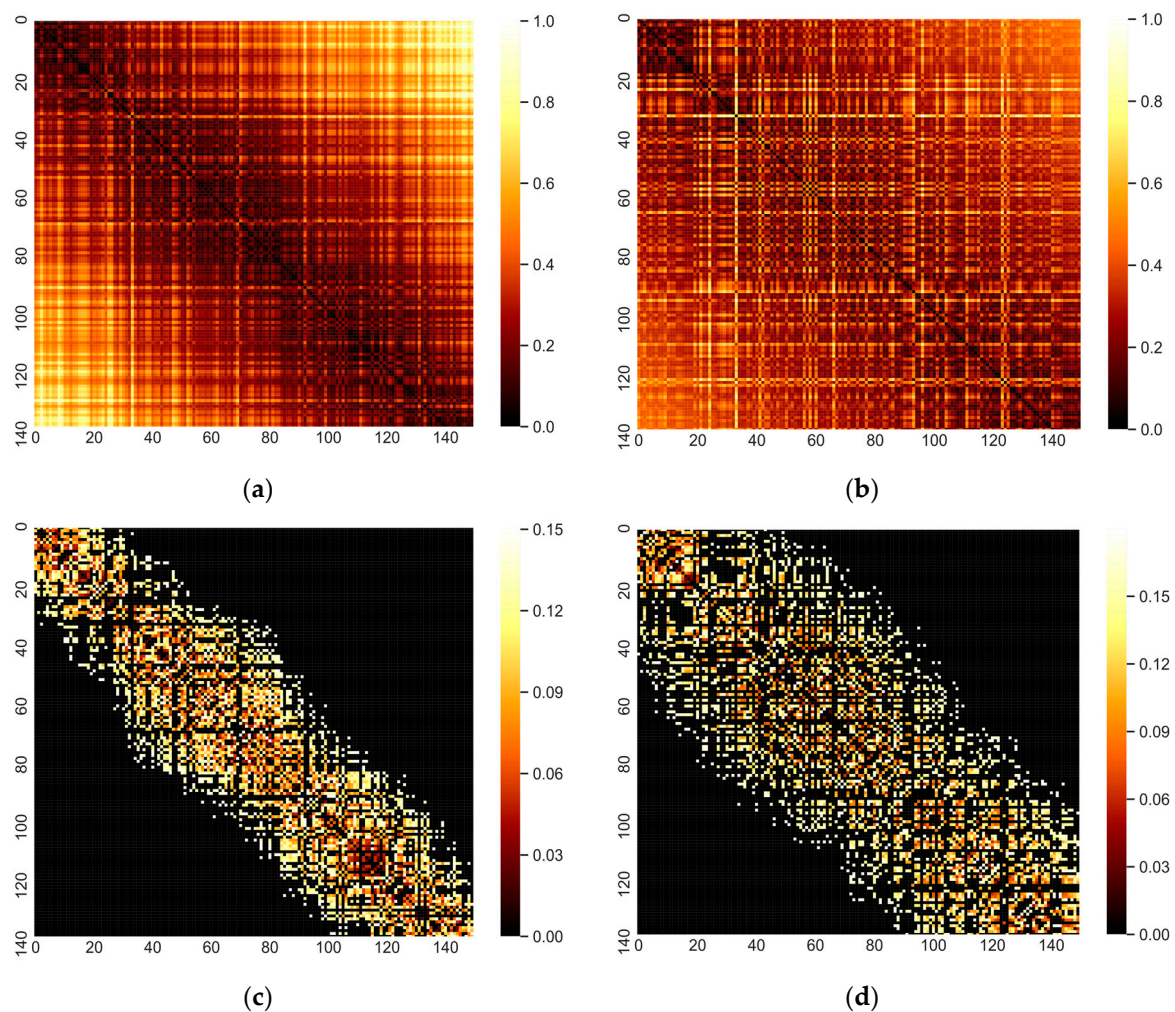


Figure 3. Heat map of distance matrix based on MMD and geodesic distance (coordinate label indicates the number of nodes and colour bar denote the value of distance between two nodes). (a) Heat map based on geodesic distance; (b) heat map based on MMD; (c) heat map in (a) with 20% of the edges to be preserved; (d) heat map in (b) with 20% of the edges to be preserved.

By comparing Figure 3c,d (similar to Figure 2a,b), we can find that the network structure based on MMD is looser. From the characteristics of Lorenz system, the loose network structure is more reasonable. This is because small changes in ρ do not exactly correspond to smooth changes in the properties of time series. Although time series with similar ρ usually have similar properties, time series with different ρ sometimes have similar behaviors [25]. That is, a node should be similar to a node with a similar ρ , but it may also be similar to a node with a different ρ , which results in a loose network structure. Compared with geodesic method, MMD can capture the two similarities more

effectively, and this results in a looser network. This is because the geodesic method is a special case of MMD in some ways. The deeper reason is that geodesic method uses only ONE covariance matrix (Gaussian distribution with a zero mean vector) to represent the data, while the MMD method uses GMM (linear combination of multiple Gaussian distributions) to fit the data. In contrast, MMD can capture more detailed information to find more neighbor nodes.

4. Conclusions

In this communication, a method was proposed for constructing a complex network of univariate chaotic time series. Compared with the commonly used metric, the introduced MMD can capture the similarity between GMMs more effectively, which is the key problem of constructing complex networks of the chaotic time series. Although the proposed method is specific to chaotic time series, it can also be applied to time series in other fields. In addition, it can be directly generalized to the case of multivariate time series by omitting phase space reconstruction.

Funding: This research was funded by the National Natural Science Foundation of China, grant number 61362024, 61761020.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Kantz, H.; Thomas, S. *Non-Linear Time Series Analysis*, 2nd ed.; Cambridge University Press: London, UK, 2004; ISBN 978-0521529020.
2. Lorenz, E.N. Deterministic Nonperiodic Flow. *J. Atmos. Sci.* **1963**, *20*, 130–141. [[CrossRef](#)]
3. Bezsudnov, I.V.; Snarskii, A.A. From the time series to the complex networks: The parametric natural visibility graph. *Phys. A Stat. Mech. Its Appl.* **2014**, *414*, 53–60. [[CrossRef](#)]
4. Zhao, Y.; Weng, T.; Ye, S. Geometrical invariability of transformation between a time series and a complex network. *Phys. Rev. E* **2014**, *90*, 012804. [[CrossRef](#)] [[PubMed](#)]
5. Sun, J.; Yang, Y.; Xiong, N.N.; Dai, L.; Peng, X.; Luo, J. Complex Network Construction of Multivariate Time Series Using Information Geometry. *IEEE Trans. Syst. Man, Cybern. Syst.* **2019**, *49*, 107–122. [[CrossRef](#)]
6. Zou, Y.; Donner, R.V.; Marwan, N.; Donges, J.F.; Kurths, J. Complex network approaches to nonlinear time series analysis. *Phys. Rep.* **2019**, *787*, 1–97. [[CrossRef](#)]
7. Gao, Z.-K.; Small, M.; Kurths, J. Complex network analysis of time series. *Europhysics Lett.* **2016**, *116*, 50001. [[CrossRef](#)]
8. Donner, R.V.; Zou, Y.; Donges, J.F.; Marwan, N.; Kurths, J. Recurrence networks—A novel paradigm for nonlinear time series analysis. *New J. Phys.* **2010**, *12*, 033025. [[CrossRef](#)]
9. Lacasa, L.; Luque, B.; Ballesteros, F.; Luque, J.; Nuño, J.C. From time series to complex networks: The visibility graph. *Proc. Natl. Acad. Sci.* **2008**, *105*, 4972–4975. [[CrossRef](#)] [[PubMed](#)]
10. Ruan, Y.; Donner, R.V.; Guan, S.; Zou, Y. Ordinal partition transition network based complexity measures for inferring coupling direction and delay from time series. *Chaos An Interdiscip. J. Nonlinear Sci.* **2019**, *29*, 043111. [[CrossRef](#)] [[PubMed](#)]
11. Zhang, J.; Small, M. Complex Network from Pseudoperiodic Time Series: Topology versus Dynamics. *Phys. Rev. Lett.* **2006**, *96*, 238701. [[CrossRef](#)] [[PubMed](#)]
12. Berthold, M.R.; Höppner, F. On Clustering Time Series Using Euclidean Distance and Pearson Correlation. *arXiv* **2016**, arXiv:1601.02213.
13. Zhang, J.; Luo, X.; Small, M. Detecting chaos in pseudoperiodic time series without embedding. *Phys. Rev. E* **2006**, *73*, 016216. [[CrossRef](#)] [[PubMed](#)]
14. Jeong, Y.-S.; Jeong, M.K.; Omitaomu, O.A. Weighted dynamic time warping for time series classification. *Pattern Recognit.* **2011**, *44*, 2231–2240. [[CrossRef](#)]
15. Shekofteh, Y.; Jafari, S.; Sprott, J.C.; Hashemi Golpayegani, S.M.R.; Almasganj, F. A Gaussian mixture model based cost function for parameter estimation of chaotic biological systems. *Commun. Nonlinear Sci. Numer. Simul.* **2015**, *20*, 469–481. [[CrossRef](#)]
16. Takens, F. Detecting strange attractors in turbulence. In *Dynamical Systems and Turbulence, Warwick 1980*; Springer: Berlin/Heidelberg, Germany, 1981; pp. 366–381.

17. Pennec, X.; Fillard, P.; Ayache, N. A Riemannian Framework for Tensor Computing. *Int. J. Comput. Vis.* **2006**, *66*, 41–66. [[CrossRef](#)]
18. Hershey, J.R.; Olsen, P.A. Approximating the Kullback Leibler Divergence Between Gaussian Mixture Models. In Proceedings of the 2007 IEEE International Conference on Acoustics, Speech and Signal Processing—ICASSP '07, Honolulu, HI, USA, 15–20 April 2007; pp. IV-317–IV-320.
19. Gretton, A.; Borgwardt, K.M.; Rasch, M.J.; Schölkopf, B.; Smola, A. A kernel two-sample test. *J. Mach. Learn. Res.* **2012**, *13*, 723–773.
20. Reed, M.; Simon, B. *Methods of Modern Mathematical Physics*; Academic Press: San Diego, CA, USA, 1980; ISBN 9780125850506.
21. Muandet, K.; Fukumizu, K.; Dinuzzo, F.; Schölkopf, B. Learning from distributions via support measure machines. In Proceedings of the Advances in Neural Information Processing Systems, Lake Tahoe, NV, USA, 3–6 December 2012; pp. 10–18.
22. Strogatz, S.H. *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*; Westview Press: Cambridge, MA, USA, 1994.
23. Scipy.integrate.odeint. Available online: <https://docs.scipy.org/doc/scipy/reference/generated/scipy.integrate.odeint.html> (accessed on 24 January 2020).
24. Software for Complex Networks. Available online: <http://networkx.github.io/> (accessed on 24 January 2020).
25. Sparrow, C. *The Lorenz Equations: Bifurcations, Chaos, and Strange Attractors*; Springer: New York, NY, USA, 1982.



© 2020 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).