2015 International Symposium on Nonlinear Theory and its Applications
NOLTA2015, Kowloon, Hong Kong, China, December 1-4, 2015

Complex Network Transformations of Nonlinear Time Series: the Ordinal Partitions Method

Konstantinos Sakellariou†, Michael Small†, Thomas Stemler† and Kevin Judd†

†School of Mathematics and Statistics, University of Western Australia
35 Stirling Highway, Crawley, WA 6009, Australia
Email: konstantinos.sakellariou@research.uwa.edu.au, michael.small@uwa.edu.au, thomas.stemler@uwa.edu.au, kevin.judd@uwa.edu.au

Abstract—Recently several methods which utilise network theory as a tool for nonlinear time series analysis have been developed. Given data from a single measurement process alone, network transformations of time series produce an associated network space. Studying its topology and properties can provide insight regarding the underlying dynamical system. We focus on the so-called *ordinal partitions* method, a technique where nodes encode for symbolic orderings of discrete dynamical states and network connectivity is defined by temporal succession. In contrast to alternative approaches which mainly concentrate on topological aspects, dynamical information is directly encoded into the resulting *ordinal networks*. By applying this technique to data generated by numerical simulation of well-studied systems, we explore how network properties reflect characteristics of the underlying dynamics governing the original time series. Additionally, we perform a parameter investigation.

1. Introduction

Natural systems often exhibit very high complexity which may be (partially) captured by high-dimensional models. These models are formed by continuous refinement, through the juxtaposition of *observations* and an associated *theoretical framework*. Observations are usually drawn from a single measurement process which essentially projects a high-dimensional dynamical state onto a single number. Therefore, a data set - and in particular *time series*, a type of data collected at successive points in time - is a form of compressed information about a dynamical system. Extracting and interpreting this information can provide us with additional insight and enhance our knowledge about the dynamical behaviour of the underlying system.

The aim of time series analysis is to develop techniques which can capture characteristics of the dynamics from the data. Linear processes have traditionally been explored using methods from stochastic analysis, for instance autoregressive (AR) or moving average (MA) models. However, nature is inherently nonlinear. Our research concentrates on *nonlinear time series analysis*, which can be very effective when the dynamics are generated by a deterministic process [3]. Approaches such as phase-space reconstruction through delay-coordinate embedding [4], dimension estimates, Lyapunov exponents, complexity measures, information theory, surrogate data methods, noise filtering and recurrence plots have proved very successful in uncovering dynamical aspects about a system from raw data, especially if the state-space dimension is low.

A novel approach was first introduced by [5] who developed a transformation of time series to complex networks. Recognising the immense potential provided by the abstraction of a network representation, the authors used it as a means of encoding information rather than a structure tied to a physical/virtual interpretation. Soon after various other techniques followed [6, 7, 8]. Complex network transforms often do not require a delay reconstruction and, in addition, are often robust to measurement noise. This idea became popular very rapidly and led to a surge in the interest for complex networks among the dynamical theory community. Recently, a transform which explicitly encodes dynamical information into a network was formulated by [2], the *ordinal partitions network transform*. In this paper, we explore some of the properties of this technique and show aspects of its applicability.

2. Methodology

2.1. The Ordinal Partitions Method

Let x_t denote a scalar time series of N measurements. We segment it into partitions (windows) of fixed size w, which may be overlapping by a certain amount τ. In this paper we focus on the *sliding variant*, i.e. successive partitions have a lag of one point ($\tau = w - 1$). This almost fully overlapping variant captures more dynamical information in contrast to the non-overlapping variant which is formed in a static manner. Denote each such partition by $y_i = (x_i, x_{i+1}, ..., x_{i+w-1})$ for $i = 1, ..., N - w + 1$. The corresponding *ordinal partition* is defined by the permutation $o^{(i)} = (\pi_1, \pi_2, ..., \pi_w)$ where $\pi_j \in \{1, 2, ..., w\}$, $\pi_j \neq \pi_k$ if $j \neq k$ such that x_{i+j-1} is the π_j-th largest element of the w-vector y_i. In the event that two elements of y_i are equal, we arbitrarily pick the one that occurs first as the smallest. In other words, each partition is mapped onto a symbolic ordering of the natural numbers depending on the relative magnitude of the points $\{x_i, x_{i+1}, ..., x_{i+w-1}\}$. For instance, consider a partition of 6 time series points x_1, x_2,

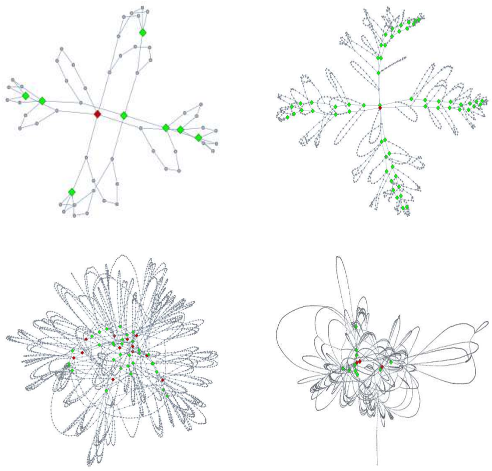

Figure 1: Ordinal networks for the Lorenz system for $w = 8$ (top left), 24 (top right), 42 (bottom left) and 50 (bottom right). The red and green diamonds depict the most highly and second most-highly connected nodes respectively.

$x_3,\ x_4, x_5, x_6$ which have relative magnitude of the form $x_1 < x_4 < x_6 < x_2 < x_3 < x_5$. This partition is assigned the ordering $(6, 3, 2, 5, 1, 4)$ since x_1 is the smallest member of the group, x_2 is the third largest, x_3 is the second largest etc.

We now proceed to the *ordinal network* formation. Each *distinct* ordering $o^{(i)}$ is mapped to a node with connectivity determined by temporal succession, i.e. only the nodes corresponding to neighbouring partitions in the time series are connected. Thus, the size of the network will be at most $w!$ (in practice this number is significantly smaller due to important role forbidden patterns play in continuous dynamical systems [9]. If we wish to incorporate additional dynamical information, we can further consider the directed and weighted constructions where a transition from $o^{(i)}$ to $o^{(j)}$ is manifested on the network by the inclusion of edge (i, j) - but not (j, i) - with a weight given by the relative frequency of transition (i, j),

$$f_{ij} = \frac{\#(o^{(i)} \rightarrow o^{(j)})}{\sum\limits_i \sum\limits_j \#(o^{(i)} \rightarrow o^{(j)})}. \tag{1}$$

Selecting a value for the main parameter w is of critical significance. As portrayed in Fig. 1, completely different structures emerge for different partition sizes. Various means for finding an optimal value for w may be explored, most notably the *graphical inflexion* and *maximum network link entropy* possibilities proposed in the original paper [2]. However, numerical evidence suggests that useful information may be obtained by considering various partition sizes, in a sense looking at the dynamics from a different perspective. In Section 3.4 we elaborate further using the results of the parameter investigation.

2.2. Node Centrality Properties

The generated networks' structure is explored using a variety of measures, both local and global in nature. One of the main tasks of network analysis is the identification of node importance. The simplest such characterisation is provided by the notion of the *node degree*, the number of links adjacent to a node. Other common computational measures to this end are given by the three most used *centrality* definitions. *Closeness* centrality represents the ease of reaching other nodes (relative 'closeness' or 'farness' of a node) and is a natural distance metric between pairs of nodes. It is defined as

$$CC_i = \frac{1}{\sum\limits_j d(i, j)}, \tag{2}$$

where $d(i, j))$ is the number of edges between nodes i and j. *Betweenness* centrality, on the other hand, is more representative of a node's importance in information flow through the network. It quantifies the number of times a node acts as a bridge along the shortest path between two other nodes and is defined by

$$BC_i = \sum_{i \neq j \neq k} \frac{\sigma_{jk}(i)}{\sigma_{jk}}, \tag{3}$$

where σ_{jk} denotes the number of shortest paths between nodes j and k and $\sigma_{jk}(i)$ is the number of those paths that pass through i. *Eigenvector* centrality measures the influence of nodes in information flow by looking at its neighbours, i.e. influence is estimated by the number of strong 'friends' you have in a social context. It is defined by looking at the spectral properties of the adjacency matrix.

2.3. Network Entropy Measures

In addition to all of the above, which comprise standard computational tools in complex network theory, we make use of entropy measures applied to specific network properties. *Entropy* is a measure of the information content in a signal and may be interpreted also as the uncertainty associated with random behaviour (amount of unpredictability). We apply this notion to the degree distribution of our networks to obtain a global measure, termed *degree entropy*, and to the distribution of links of each node to obtain a local measure, termed *node-link entropy*. In this case, define $p(i, j)$ as the probability of traversing to node j if we are currently residing on node i. Then, the node-link entropy of node i is given by

$$H_i = -\sum_{j,\, a_{ij} \neq 0} p(i, j) \cdot log[p(i, j)], \quad p(i, j) = \frac{a_{ij}}{\sum\limits_j a_{ij}}, \tag{4}$$

where the term a_{ij} denotes the corresponding value in the adjacency matrix (1 if unweighted). The median nodal entropy of the entire network, a quantity known as *network link entropy*, may also be computed.

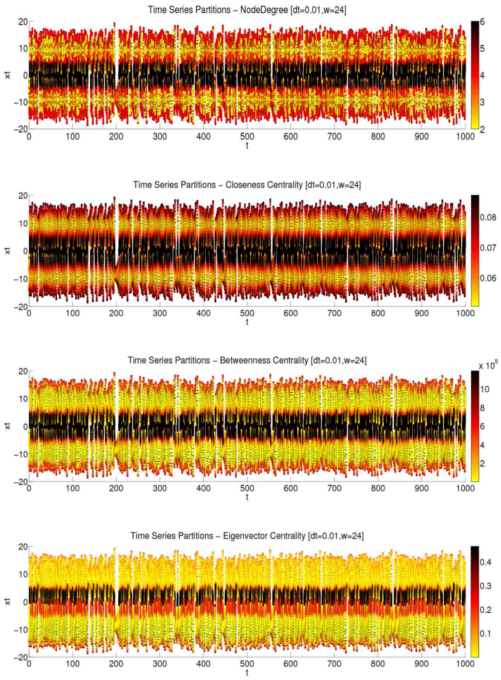

Figure 2: Time series of the x-component of the Lorenz system. Network properties (generated with $w = 24$) are mapped back onto the partitions by colouring the first point. From top to bottom the node degree, closeness centrality, betweenness centrality and eigenvector centrality are shown.

3. Results

3.1. Lorenz System

We compute sample time series from the x component of the *Lorenz* system, developed in 1963 as a simplified model for atmospheric convection [1]. It is one of the most famous and well-studied systems exhibiting chaotic behaviour under certain parameter regimes. It is given by the following differential equations

$$\begin{aligned} \frac{dx}{dt} &= \sigma(y - x) \\ \frac{dy}{dt} &= x(\rho - z) - y \\ \frac{dz}{dt} &= xy - \beta z, \end{aligned} \tag{5}$$

where $\sigma = 10, \rho = 28$ and $\beta = \frac{8}{3}$. Using MATLAB routines based on the fourth Runge-Kutta method, we integrated the system numerically for $t \in [0, 1000]$ with a sampling rate equal to 0.01 ($N = 100,000$ points, period of oscillations approximately equal to 1 unit of time).

3.2. Ordinal Networks

The networks (unweighted, undirected variant) generated by applying the ordinal partitions method on the Lorenz system are shown in Fig. 1 for four different values of w. As is evident, the four networks have a significantly different topology. The first thing to note is that the size of four networks grows as the partition size is increased (see Section 3.4). Additionally, the first two networks possess a structure which clearly exhibits four legs, potentially a sign of community structure. The number of 'loops' (loosely defined as paths traversed from one neighbouring node to the other through a number of other degree-2 nodes) also increases with the partition size, as is the overall connectivity of the network. Finally, note that for the Lorenz networks in general, the most highly connected node (depicted in red diamonds) is usually found at the heart of the network (and often coincides with the node with the highest betweenness and closeness centrality) and serves as the main connector for information flow. It corresponds to the most frequent ordinal pattern occurring in the time series. The second most highy connected nodes (depicted in green) serve as the main structural components holding the 'leg' structures together. However, in the case of $w = 42$ (bottom left graph) there are several highly connected nodes and the number of 'legs' has multiplied significantly. These observations indicate that depending on the partition size, the generated ordinal network captures different amounts of information about the dynamics of the underlying system.

3.3. Mapping Network Properties onto the Time Series

We applied the local network measures defined above to the generated networks ($w = 24$ shown here) and mapped these nodal properties back onto the time series (Fig. 2) and onto the phase-space attractor (Fig. 3) of the Lorenz system. Since the parameter values we have chosen lead to chaotic behaviour, this data set serves as a good example for testing the amount of complexity captured by the ordinal networks. Fig. 2 clearly shows that the middle region where trajectories transition from one regime to the other (corresponding to the separatrix on the attractor) is identified merely by looking at network properties. As expected, this region is crucial both in terms of the dynamics (connects the two wings; high frequency of trajectories etc.) and in terms of the networks for the transmission of information. Betweenness centrality especially is a great indicator which is sensible intuitively; the most influential patterns in the time series determine the transitions between the two different regimes. Another noteworthy observation is that the stationary points are also clearly traced by the node-wise properties with no exception, with betweenness again being the most representative. Extrema points also hold major importance for the dynamics of the system and the network representation seems to capture this hierarchy in a natural fashion: the separatrix region is mapped onto nodes of very high betweenness (the most influential people in a social context); stationary points are not as influential but still play a very central role in that they determine when a trajectory will move into the separatrix region and consequently jump to the other wing or continue the rotation around the same unstable fixed point. Finally, other

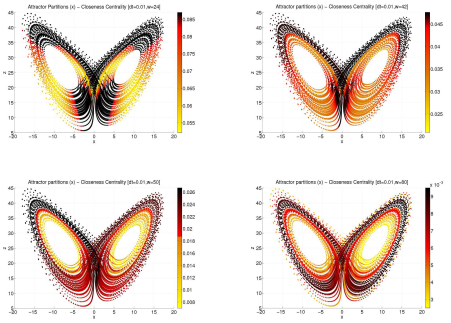

Figure 3: Closeness centrality mapped back onto the attractor for $w = 24, 42, 50$ ad 80 going from top to bottom row-wise.

regions have much less influence in the dynamics of the system and this is reflected by the low centrality values of those nodes. Fig. 3, which shows the Lorenz attractor with closeness centrality mapped back onto it, depicts this distinction accurately for ordinal networks generated with $w = 24$. Therefore, local dynamical behaviour is certainly incorporated into the complex network representations.

What about more global behaviour? Fig. 3 shows the Lorenz attractor for four different ordinal networks (from top and traversing rows, $w = 24, 42, 50$ and 80). Clearly, node-wise closeness centrality differs significantly as expected. For each partition size, different aspects about the underlying dynamics are uncovered by the network and, hence, different importance is assigned to each node (and by extrapolation each w-pattern). For the larger partition sizes ($w = 50, 80$) we have a completely different picture where local dynamical behaviour is not captured any more. Rather, the attractor seems to be partitioned in different bands with boundaries possibly defined by unstable periodic orbits (UPOs) of low order (i.e. the most dominant, least unstable ones). Therefore, as we increase the size of partitions in which we segment our time series, the generated ordinal networks reflect more global, macroscopic properties of the underlying dynamics.

3.4. Parameter Investigation

Here we present results regarding global network properties and how they vary with the partition size w. Fig. 4 shows plots of the number of nodes (V) and the network's link entropy (top and bottom respectively) as we vary w. The graphical inflection in the w-V graph corresponds to $w = 42$. For small partition sizes, $w!$ is not very large and, coupled to the fact that there are many forbidden patterns, the generated networks are also small and carry no particular significance regarding the underlying dynamics. As we increase w further, however, local dynamical behaviour starts being incorporated into the network until we reach the inflexion point, above which global dynamical properties start being encapsulated. Network link entropy shows the amount of information carried by the average distribution of links and can also act as a strong indicator of the usefulness of networks generated using a particular w. The weighted variant will be more informative in this case as information about transitions between patterns needs to be incorporated.

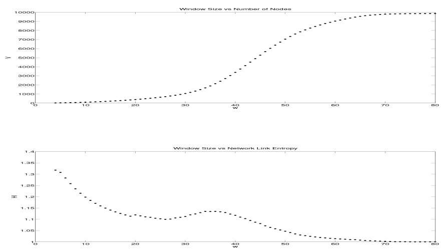

Figure 4: Number of nodes in the network (top) and network link entropy (bottom) as the partition size w is varied, $w \in [4, 80]$.

4. Conclusions and Discussion

The ordinal partitions network transform suggests a new approach to nonlinear time series analysis and constitutes a field which is ripe for exploration. It incorporates dynamical information explicitly into the network representation, it does not require embedding of the data, is very easy to implement and seems to reflect subtle properties about the underlying dynamics.

Acknowledgments

MS is supported by an Australian Research Council Future Fellowship (FT110100896).

References

[1] E. N. Lorenz, "Deterministic non-periodic flow", *J. Atoms. Sci.*, vol.20, pp.130–141, 1963.

[2] M. Small, "Complex networks from time series: Capturing dynamics", *2013 IEEE International Symposium on Circuits and Systems (ISCAS)*, pp.2509–2512, 2013.

[3] E. Bradley and H. Kantz, "Nonlinear time-series analysis revisited", *Chaos*, vol.25, pp.097610, 2015.

[4] F. Takens, "Detecting strange attractors in fluid turbulence", *Dynamical Systems and Turbulence*, pp.366–381, 1981.

[5] J. Zhang and M. Small, "Complex Network from Pseudoperiodic Time Series: Topology versus Dynamics", *Phys. Rev. Lett.*, vol.96, pp.238701, 2006.

[6] X. Xu, J. Zhang and M. Small, "Superfamily phenomena and motifs of networks induced from time series", *Proc. Nat. Acad. Sci.*, vol.105, pp.19601–19605, 2008.

[7] R. V. Donner, Y. Zou, J. F. Donges, N. Marwan, and J. Kurths, "Recurrence networks: a novel paradigm for nonlinear time series analysis", *New Journal of Physics*, vol.12, pp.033025, 2010.

[8] R. V. Donner, M. Small, J. F. Donges, N. Marwan, Y. Zou, R. Xiang and J. Kurths, "Recurrence-based time series analysis by means of complex network methods", *International Journal of Bifurcation and Chaos*, vol.21, pp.1019–1046, 2011.

[9] J. M. Amigo, S. Zambrano and M. A. F. Sanjuan, "True and false forbidden patterns in deterministic and random dynamics", *Europhys. Lett.*, vol.79, pp.50001, 2007.