Commun Nonlinear Sci Numer Simulat 42 (2017) 396–403

Contents lists available at ScienceDirect

Commun Nonlinear Sci Numer Simulat

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

Research paper

Visibility graph analysis for re-sampled time series from auto-regressive stochastic processes

Rong Zhang[a], Yong Zou[a,*], Jie Zhou[a], Zhong-Ke Gao[b], Shuguang Guan[a]

[a] Department of Physics, East China Normal University, 200062 Shanghai, China
[b] School of Electrical Engineering and Automation, Tianjin University, 300072 Tianjin, China

ARTICLE INFO

Article history:
Received 8 January 2016
Accepted 25 April 2016
Available online 8 June 2016

Keywords:
Visibility graph
AR processes
Nonlinear time series analysis

ABSTRACT

Visibility graph (VG) and horizontal visibility graph (HVG) play a crucial role in modern complex network approaches to nonlinear time series analysis. However, depending on the underlying dynamic processes, it remains to characterize the exponents of presumably exponential degree distributions. It has been recently conjectured that there is a critical value of exponent $\lambda_c = \ln 3/2$, which separates chaotic from correlated stochastic processes. Here, we systematically apply (H)VG analysis to time series from autoregressive (AR) models, which confirms the hypothesis that an increased correlation length results in larger values of $\lambda > \lambda_c$. On the other hand, we numerically find a regime of negatively correlated process increments where $\lambda < \lambda_c$, which is in contrast to this hypothesis. Furthermore, by constructing graphs based on re-sampled time series, we find that network measures show non-trivial dependencies on the autocorrelation functions of the processes. We propose to choose the decorrelation time as the maximal re-sampling delay for the algorithm. Our results are detailed for time series from AR(1) and AR(2) processes.

1. Introduction

In the last years, complex network representations of time series have been proposed to characterize the underlying system [1–6], covering a great variety of fields of applications. For example, recurrence network approaches have been applied to climate data analysis [4,7], chaotic electro-chemical oscillators [8], or two-phase flow data [9]. Some basic network motif structures have be identified in music data [10]. Visibility graph (VG) and horizontal visibility graph (HVG) algorithms have been successfully applied to hurricane data in the US [11], financial market [12], turbulence [13,14], and sunspot time series [15,16], providing novel insights from a complex systems perspective. Several other methods have been discussed in [10]. In this work, we construct (H)VGs [3,17,18] for time series generated by auto-regressive models.

To apply (H)VG approaches, a proper transformation of the time series to a network representation is required. The VG algorithm maps a time-ordered set of N real numbers to a graph G with N nodes, which is completely described by the binary $N \times N$ adjacency matrix $\mathbf{A}$. More specifically, let us consider a univariate time series $\mathbf{x} = [x_1, x_2, \ldots, x_i, \ldots, x_N]$, where N denotes the time series length and the subscript i represents the discrete sampling time. In terms of the associated VG, individual observations are interpreted as vertices of a complex network, and edges are placed between pairs of vertices

* Corresponding author.
E-mail address: yzou@phy.ecnu.edu.cn (Y. Zou).

http://dx.doi.org/10.1016/j.cnsns.2016.04.031
1007-5704/

that exhibit some visibility condition, namely,

$$\frac{x_i - x_k}{k - i} > \frac{x_i - x_j}{j - i}, \tag{1}$$

which has to be fulfilled for all time points k with $i < k < j$ [3].

As a notable modification of the VG algorithm, the simplified HVG criterion has been proposed in [17]. Specifically, two observations made at times i and j are considered to be connected in a HVG if and only if

$$x_k < \min\left(x_i, x_j\right) \tag{2}$$

for all k with $i < k < j$. It can be easily seen that the edge set of the HVG associated with a given time series is a subset of the edge set of the associated VG. One advantage of the HVG is that for certain simple stochastic processes, some basic graph properties can be calculated analytically [18].

Recent work on (H)VGs has mainly concentrated on the properties of the degree distribution $p(k)$. Concerning the HVG, exponential functional forms have been obtained for many random processes, namely, $p(k) \sim e^{-\lambda k}$. A scaling factor of $\lambda_c = \ln(3/2)$ has been found in the case of uncorrelated noise (white noise), which has been further proposed to separate stochastic from chaotic dynamics in the following senses [18–20]: (i) correlated stochastic series are characterized by $\lambda > \lambda_c$, slowly tending to an asymptotic value of $\ln(3/2)$ for very weak correlations, whereas (ii) chaotic series are often characterized by $\lambda_{chaos} < \lambda_c$ for decreasing correlations or increasing chaos dimensionality, respectively [18]. In this work, we provide some further examples supporting argument (i). Meanwhile, we show some peculiar results indicating that λ_c should not be interpreted as a general critical value separating chaos from noise.

In fractal processes, numerical results suggest that $p(k)$ exhibits a power law [3], $p(k) \sim k^{-\gamma}$. For instance, VG analysis has been suggested to characterize fractional Brownian motions, finding some heuristic relationship between γ and the process' Hurst exponent [21,22]. Depending on the fractality properties of the underlying process, recently an algorithm for constructing VGs from segmented time series has been proposed [23], which estimates power-law exponents reflecting scale-free properties quite well. The idea hinges on a proper choice of the time delay τ, which re-samples the original time series resulting in a number of segmentations. As we will discuss below, the number of segments to be used is often not known in the original algorithm. In other words, we do not know the upper bound of τ_{max} to terminate the computation. So far, only a heuristic choice of τ_{max} has been suggested when network characteristics of segmented time series show more or less convergent behavior [23].

In this work, we will focus on applying (H)VG analysis to auto-regressive (AR) stochastic processes, which often describe certain time-varying processes in nature, economics, etc. AR models also have wide applications to climate research. The AR model specifies that the output variable depends linearly on its own previous values and on a stochastic term. More specifically, $\mathbf{x} = [x_1, x_2, \ldots, x_i, \ldots, i \in \mathbb{Z}]$ is an AR model of order p denoted as AR(p) if

$$x_t = \sum_{j=1}^{p} \varphi_j x_{t-j} + \varepsilon_t, \tag{3}$$

where φ_j, $j \in [1, p]$, are real-valued coefficients of the model, and ε_t is white noise. We further assume that the error terms ε_t follow a Gaussian distribution with zero mean and unit variance. Specifically, we will perform both VG and HVG analysis for AR(1) and AR(2) processes in their stationary regimes, namely, (i) $|\varphi_1| < 1$ for the AR(1) model, and (ii) $\varphi_1 + \varphi_2 < 1, \varphi_2 - \varphi_1 < 1, |\varphi_2| < 1$ for the AR(2) model. Time series of AR models exhibit serial correlations conveniently captured by the autocorrelation function (ACF). For AR(p) processes, the ACFs can be computed analytically [24]. In general, the correlation length of a stochastic process increases if the corresponding ACF shows rather slow decays to zero [25].

The objectives of this work are the following. First, we test the hypothesis of the universality of the critical value of λ_c when anti-correlated increments are present in AR processes. Second, we apply the re-sampling algorithm to time series from AR(p) models and extend the existing (H)VG analysis from disclosing degree distribution properties to characterizing global network properties. Third, we suggest to choose the decorrelation time of the given time series as the maximal delay τ_{max} for the re-sampling algorithm, after which network measures converge to some asymptotic values that are expected for uncorrelated stochastic processes with the given probability distribution function.

This paper is organized as follows: In Section 2, we review the main network measures that will be used in this work. We present the re-sampling algorithm to obtain segmented time series in Section 3. The results will be shown in Section 4 and some conclusions are drawn in Section 5.

2. Structural properties of (H)VG

We characterize the structural properties of VG and HVG based on the following two aspects: (i) degree distributions $p(k)$ and (ii) global network measures.

2.1. Degree distributions

The degree k_i of a node i simply counts the number of direct connections associated with i, $k_i = \sum_j A_{i,j}$, where A is the adjacency matrix. The degree distribution $p(k)$ gives the fraction of nodes in the network with degree k. Exponential