INPE – National Institute for Space Research
São José dos Campos – SP – Brazil – May 16-20, 2016

RELATION BETWEEN AUTOCORRELATION SEQUENCE AND AVERAGE SHORTEST-PATH LENGTH IN A TIME SERIE TO NETWORK MAPPING

Amanda Leite de Camargo[1] *and Marcio Eisencraft*[2]

[1]Universidade Federal do ABC, Santo André, Brazil, amanda.leite@ufabc.edu.br
[2]Escola Politécnica, University of São Paulo, São Paulo, Brazil, marcio@lcs.poli.usp.br

Abstract: An invertible mapping between time series and networks was recently proposed. It can be used as a tool to figure out properties of the mapped time series. In the present work we use controlled artificial signals to numerically investigate how correlation properties of time series are mapped in the topological measures of associated networks. More specifically, we employ filtered uniform white noise and analyse how the autocorrelation sequence influences the average shortest-path length.

keywords: Time Series Analysis, Nonlinear Dynamics and Complex Systems, Discrete Dynamical Systems, Noise, Mapping.

1. INTRODUCTION

Many works have proposed different procedures to characterize time series using properties of complex networks. For instance, in [1] a network is obtained by partitioning a time serie and mapping each of the resulting subsets in its nodes. In [2] the time serie is divided in cycles and the correlation between these cycles is used to construct an associated network. The mapping between time series and networks is done using the *Visibility Algorithm* in [3] and the *Horizontal Visibility Graph* in [4].

Although these works use topological measures like average shortest-path length or clustering coefficient to classify time series, it is not evident how conventional and classical time series properties, like correlation coefficients, are related to these measures. In this work we use controlled artificial random time series to evaluate how their correlation properties are mapped in the average shortest-path length of networks obtained using the invertible mapping procedure proposed in [1].

This extended abstract is organized as follows: in Section 2 we briefly describe the employed topological measures and mapping. In section 3 the mapping is applied to filtered white noise time series and we analyse the average shortest-path length of obtained networks. Finally, in Section 4 the conclusions are drawn.

2. NETWORKS AND MAPPING

In this section, we set the complex networks notation and definitions [5] used in this extended abstract.

Consider a graph composed by N nodes and M links connecting them. The link between nodes i and j in this direction is represented by l_{ij} . Two nodes i and j of a network are neighbours whether the link l_{ij} exists.

A path between two nodes i and j is a set of links starting in i and ending in j. A path with the smaller number of links, L_{ij}, is called a shortest-path. In a connected network, i.e., a network where there is a path between any two nodes, the average shortest-path length $\langle L\rangle$ is given by

$$\langle L\rangle = \frac{1}{N(N-1)}\sum_{i=1}^{N}\sum_{j=1}^{N} L_{ij}. \quad (1)$$

The degree K_i of a node i is the number of links leaving or arriving in it. The average degree $\langle K \rangle$ is

$$\langle K \rangle = \frac{1}{N} \sum_{i=1}^{N} K_i. \tag{2}$$

The clustering coefficient C_i of a node i is given by the proportion of links set up between the neighbours of i and the total number of links which could be set up between them. The average clustering coefficient $\langle C \rangle$ of a directed network is given by

$$\langle C \rangle = \frac{1}{N} \sum_{i=1}^{N} C_i. \tag{3}$$

Recently it was proposed by [1] an invertible mapping between time series and networks using quantiles. Using the direct mapping it is possible to convert a time serie in a directed network.

The m-order quantile $q(m)$, $0 < m < 1$, is the sample so that $100m\%$ of all the samples is lower than $q(m)$.

Given a discrete time serie $x(n)$ with N_p samples, the mapping proposed in [1] can be described by the following steps:

i. the samples are sorted in ascending order;

ii. they are divided in Q subsets with N_p/Q samples. The last sample of each of this subsets is the quantile $q_1, q_2, \ldots, q_Q$ of the serie $x(n)$, where $q_i = q(i/Q)$;

iii. each subset is assigned to a node in the corresponding network. The transition matrix $\mathbf{T}_{Q \times Q}$ element t_{ij} is given by the number of times a sample $x(n)$ belonging to the subset defined by the quantile q_i is followed by a sample $x(n+1)$ belonging to the subset defined by q_j.

iv. in order to produce an unweighed network, links whose weight is smaller than a threshold μ are eliminated from the network.

3. RESULTS AND DISCUSSION

In the present work we consider uniform white noise filtered by finite impulsive response filters. More specifically, it is considered Moving Average filters given by

$$r_f(n) = \frac{1}{\beta + 1} \sum_{k=0}^{\beta} r(n-k), \tag{4}$$

where $r(n)$ is a N_p samples realization of an uniform white noise distributed in the interval $[-1, 1]$ and β is the filter order. These signals were normalized in order to have their higher sample value equal to one. The $\beta+1$ first samples are discarded in order to remove transient effects.

Figure 1 shows the first 100 samples of a realization and the autocorrelation sequences numerically obtained considering $\beta = 0, 5, 20$ and 50. The $Q = 10$ quantiles are also

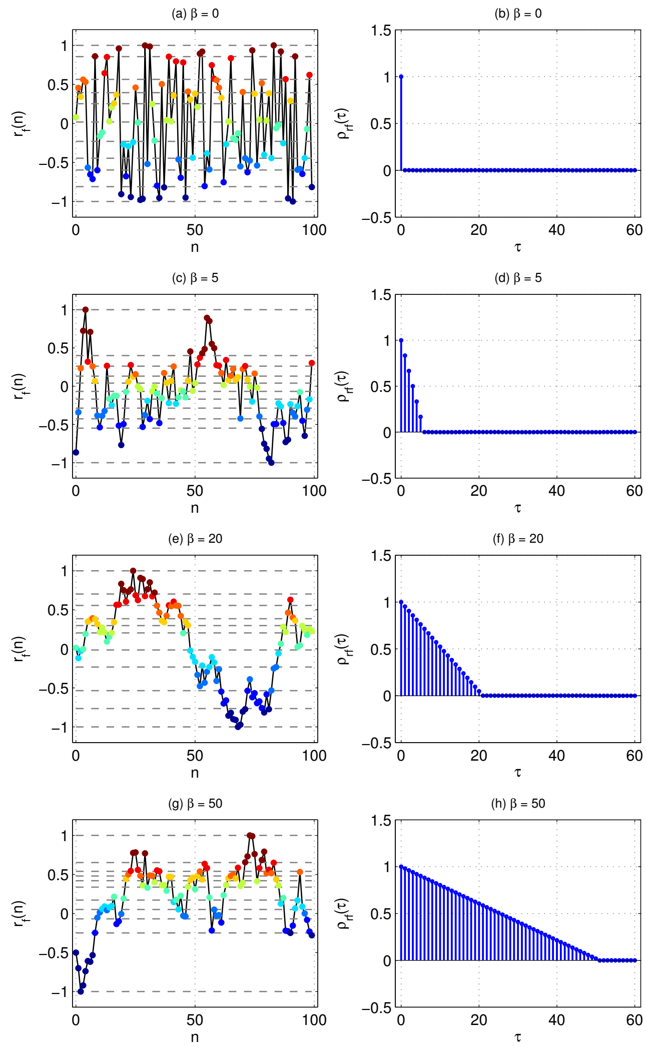

Figure 1 – Instances of filtered white noise and their autocorrelation sequences for: (a) and (b) $\beta = 0$; (c) and (d) $\beta = 5$; (e) and (f) $\beta = 20$; (g) and (h) $\beta = 50$.

shown. Figure 2 shows the networks obtained for these values of β, considering as threshold

$$\mu = \frac{\sum_{i=1}^{Q} \sum_{j=1}^{Q} t_{ij}}{2Q^2} \tag{5}$$

and 2000 realizations with 10000 samples each.

For $\beta = 0$, $r_f(n)$ is a white noise with impulsive autocorrelation sequence, as shown in Figure 1(b). This way, the subset where a sample belongs is independent of the subset where the previous sample belongs. As the number of possible transitions between subsets is Q^2, the number of transitions between any two subsets tends to N_p/Q^2. Therefore, the transition matrix $\mathbf{T}$ will converge to

$$\mathbf{T} = \frac{N_p}{Q^2} \mathbf{1}, \tag{6}$$

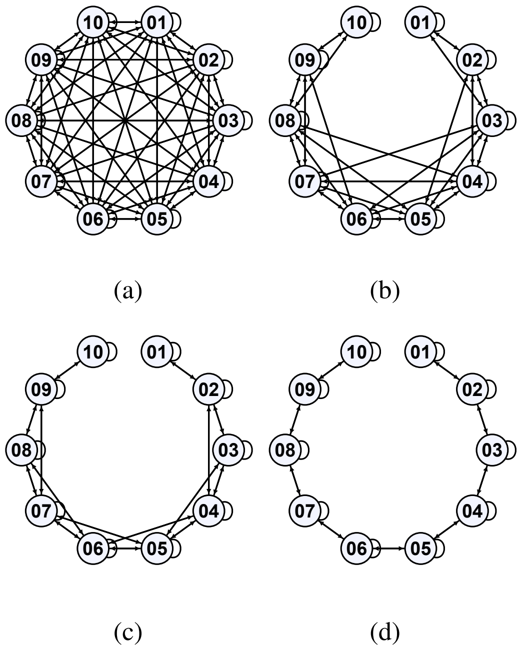

Figure 2 – Networks obtained from the filtered white noise of Figure 1: (a) $\beta = 0$, (b) $\beta = 5$, (c) $\beta = 20$, (d) $\beta = 50$ and $Q = 10$.

where $\mathbf{1}$ is a matrix of unitary elements. So, using a threshold lower than $\mu = N_p/Q^2$ the mapping results in a totally connected network, as shown in Figure2(a). For this network $\langle L\rangle = 1$, $\langle K\rangle = 2(Q-1)$ and $\langle C\rangle = 1$.

As β increases, due to the rising of the correlation between samples, consecutive samples tend to belong to adjacent subsets. Therefore, transitions between adjacent subsets occur more frequently than those between distant subsets. Consequently, obtained networks present many connections between nearby nodes and only eventual connections between distant nodes, as shown in Figures 2(b) and 2(c).

For sufficiently large β, we obtain signals whose all consecutive samples belong to the same or to adjacent subsets. As a result, in the networks obtained, only neighbour nodes are connected, as shown in Figure 2(d). For these regular networks it is expected $\langle L\rangle = (Q+1)/3$, $\langle K\rangle = 4(Q-1)/Q$ e $\langle C\rangle = 0$ [9]. For $Q = 10$, we obtain $\langle L\rangle = 11/3$, $\langle K\rangle = 3.6$ and $\langle C\rangle = 0$.

Table 1 shows values of $\langle L\rangle$, $\langle K\rangle$ e $\langle C\rangle$ numerically obtained for the networks of Figure 2. The obtained values confirm the transition from a completely connected network to a regular one.

Figure 3 shows how $\langle L\rangle$ increases with β for different values of Q. It can be seen a direct relation between β and $\langle L\rangle$. Specifically, for large values of Q this relationship tends to be one-to-one. This way, given a time series modeled as filtered white noise, the $\langle L\rangle$ of the associated network can be used to estimate the filter order of the model or the autocorrelation sequence length.

Table 1 – Properties of networks of the Figure 2.

β	$\langle L\rangle$	$\langle K\rangle$	$\langle C\rangle$
0	1.0000	18.0000	1.0000
5	1.6222	9.6000	0.7467
20	2.3111	6.0000	0.4000
50	3.6667	3.6000	0

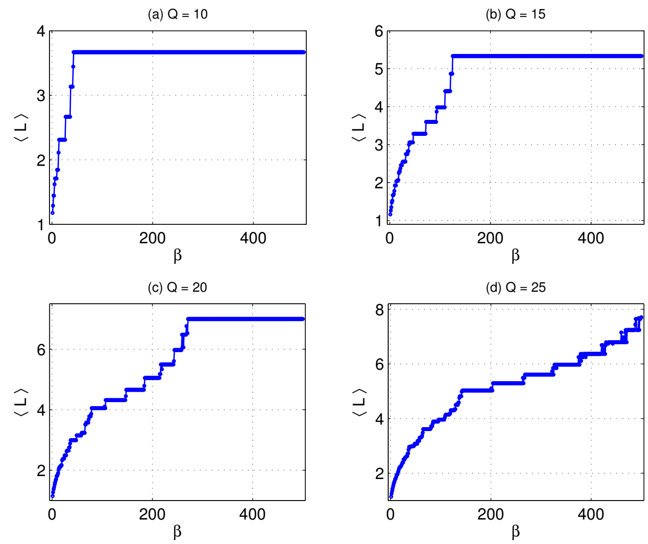

Figure 3 – Relationship between average shortest-path length and β for networks with (a) $Q = 10$, (b) $Q = 15$, (c) $Q = 20$ and (d) $Q = 25$ nodes.

4. CONCLUSION

In the present work we investigated how correlation properties of a time serie are mapped in the average shortest-path length of the network associated to it. We are currently applying these results in the modeling of financial time series.

ACKNOWLEDGMENTS

ALC thanks to the financial support from CNPq and UFABC. ME thanks to the financial support from CNPq, processes 479901/2013-9, 311575/2013-7 and from FAPESP, process 2014/04864-2.

References

[1] A. Campanharo et al., “Duality between time series and networks”, PloS one, v. 6, n. 8, p. e23378, 2011.

[2] J. Zhang; M. Small, “Complex network from pseudoperiodic time series: Topology versus dynamics”, Physical Review Letters, APS, v. 96, n. 23, p. 238701, 2006.

[3] L. Lacasa et al., “From time series to complex networks: The visibility graph”, Proceedings of the National Academy of Sciences, National Acad Sciences,

v. 105, n. 13, p. 4972-4975, 2008.

[4] R. Baggio; R. Sainaghi, “Mapping time series into networks as a tool to assess the complex dynamics of tourism systems”, Tourism Management, v. 54, p. 23-33, 2016.

[5] B. D. West et al., “Introduction to graph theory”, Prentice hall Upper Saddle River, v.2, 2001.

[6] P. Erdos, A. Rényi, “On the evolution of random graphs”, Publ. Math. Inst. Hungar. Acad. Sci, v.5, p.17-61, 1960.

[7] D. J. Watts, S. H. Strogatz, “Collective dynamics of small-world’ networks”, Nature, v.393, n.6684, p.440-442,1998.

[8] A. L. Barabási, R. Albert, “Emergence of scaling in random networks”, Science, v. 286, n. 5439, p. 509-512, 1999.

[9] A. L. Camargo, “Mapeamento de séries financeiras em redes complexas”, Masters Thesis, Universidade Federal do ABC - Mestrado em Engenharia da Informação, 2016.