

STUDYING CONNECTIVITY BETWEEN TIME-SERIES USING AN INTERACTIVE APPLICATION

Víctor J. López-Madrona¹, David Moratal², Christian G. Bénar¹

¹*Aix Marseille Univ, INSERM, INS, Inst Neurosci Syst, Marseille (FRANCE)*

²*Department of Electronic Engineering, Universitat Politècnica de València (SPAIN)*

Abstract

Connectivity is a complex concept whose meaning varies for each domain. In a broad sense, it is possible to divide the connectivity in two main groups: physical and statistical. The former represents whether there is a physical link connecting the elements, like the telephone line, while the latter indicates a relation between the dynamics of the elements, as the number of sales and the income of a company. In several scientific fields, including economics, physics and neuroscience, the statistical connectivity is computed based on the analysis of time-series. Several different methods can be applied on the data, using different features of the signals, like covariance and frequency, to estimate the connectivity, yielding different results and interpretations. The covariance indicates if two elements follow the same dynamics, i.e., if they increase or decrease at the same time. On the other hand, the frequency is related to the time required for the elements to change, which can be daily in the case of the tides, or milliseconds if referred to the computations in a microchip. A third factor is the directionality, with one element depending on the other, but not in the opposite case. This work proposes an interactive computer-based application to generate a three-node network and estimate its connectivity following different approaches. We describe how the application is structured, from the selection of the parameters to the interpretation of the output results, detailing the skills in connectivity that can be developed by the students.

Keywords: active learning, time-series, connectivity, network, interactive tools.

1 INTRODUCTION

Connectivity is a concept that might be treacherous for many graduate students [1]. The main idea of connectivity seems to be obvious, and anyone can tell if two elements are connected or not just looking at them. However, it is a broad term whose specific meaning depends on the field. We may consider that two students are connected if there is a physical link between them, for example if they are shaking their hands. It could also be an abstract connection, for example they are also connected if they play chess together every Thursday or if they are relatives. In several scientific fields, including economics, physics and neuroscience, the study of the connectivity shares a common background, which is the analysis of time series [2]. For each element, like the income of a company or the electric field of a neuron, we can acquire a signal that describes its activity along the time, with a different value for each observed time point [3].

There is an immense number of methods to analyse the connectivity between time series, each one based on a specific feature of the signals and with a different result. For example, we might be interested in the connectivity between the number of pizzas ordered and the profits of the pizza shop. These signals are correlated (they covary), as if people order more pizzas, the company will earn more money. The number of ordered pizzas depends on the day, with more sales during the weekend. That is the frequency of the signal, weekly in this case, and can be estimated with the coherence. Also, the connectivity has a certain delay: the company receives the money when the pizza is delivered, around 30 minutes after it is ordered. Thus, it is mandatory to understand the signal and the methodology before making any interpretation of the result. Using the same example, if we analyse the connectivity between the number of pizzas made (not ordered) and the profits, we may wrongly interpret that making more pizzas causes an increase of the benefits, regardless of whether people are ordering them or not. In this case, one method to differentiate correlations from potential cause-effect relationships is Granger Causality [4].

The objective of the present work is to develop an interactive computer-based application that generates time series based on variable parameters and analyse the connectivity between them with different methods, allowing a direct interpretation of the results in an environment controlled by the student. The software generates three oscillatory signals, where their specific frequency, phase and noise can be

tuned through a user-friendly interface, with an immediate representation of the time-series and their frequency information (power spectrum). The connectivity between the signals can be also parameterized, and the student can select the intensity of the link, the directionality (from A to B or from B to A) and the delay. Once the student has determined the desired parameters, several connectivity metrics can be applied on the system, with a visual representation of the results. Importantly, all the parameters can be changed at any moment, with an automatic update of all the figures. Therefore, it permits the comparison of dozens of different networks, i.e., set of elements with a specific configuration of connectivity between them, in a few minutes. Moreover, the software includes a help document, with the theoretical basis of the connectivity and examples of parameters to test certain situations of interest.

2 METHODOLOGY

The proposed application allows (i) the creation of a three-node network and (ii) the estimation of the network connectivity using three different approaches. The application is composed by a single window divided into seven main panels (Figure 1), thus, all the parameters and results are directly accessible, without extra pathways:

- Intra-node parameters (Figure 1a): These values can be modified by the user and represent the properties of each node. Each node is defined by an oscillator with a specific frequency, phase and white noise.
- Network scheme (Figure 1b): Visual representation of the created network, where each arrow is the directionality in connectivity between two nodes and the width is the intensity of the link.
- Inter-node parameters (Figure 1c): Two values can be adjusted to define the connectivity between the nodes. The intensity represents the strength of the link and the delay is the time that takes the information to go from the sender to the receiver. If the delay is zero, then the connectivity is instantaneous; if it has a positive value, for example 10 seconds, that means that a change in the sender would not be seen in the receiver until 10 seconds after it.
- Signals (Figure 1d): This panel is a temporal representation of the three signals during one second, and its change in the parameters is immediately updated in the figure.
- Power spectrum (Figure 1e): Information about the frequencies associated to each node. If two nodes share the same frequency, it may indicate that there is a relationship between them.
- Connectivity metric (Figure 1f): Three different methods can be applied to estimate the connectivity: correlation, coherence and Granger Causality. The correlation measures whether the signals vary at the same time. The coherence includes the information of the frequency in case of covariance. Thus, the nodes may be correlated at one specific frequency but not at a different one. The Granger Causality estimates if there is a causal relationship between the nodes. Therefore, it determines if the changes in one signal strictly depends on the changes in another one.
- Connectivity results (Figure 1g): Visual representation of the connectivity results. These figures will vary for each associated metric.

Importantly, the proposed application permits a straightforward integration of new methodologies so it can be updated based on specific objectives for the students. The application has been implemented using Matlab 2018b (The Mathworks, Inc., Natick, MA).

3 RESULTS

The objective of the present application is the computation of different connectivity metrics and their interpretations. The results of the developed computer-based application are presented in two different subsections: the characterization of time-series (subsection 3.1), and the differences between correlation and causality (subsection 3.2).

3.1 Characterization of time-series

Before the analysis of the connectivity between elements, it is important to understand and interpret their time-series. The application allows the creation of three signals based on their frequency, phase and noise (Figure 1a). Most of the signals that we can measure have oscillatory patterns. The number of ice creams sold increases during the summer, decreases in winter, and increases again in the next

summer. Therefore, the profits of the ice cream company have a cycle of one year, which represents its frequency. Moreover, people buy ice creams during the day, when the shop is open, so there is also a daily frequency. The information of the different frequencies of one signal can be directly extracted by estimating the power spectrum (Figure 1e).

The second property is the phase of the signal. It indicates each specific moment of the cycle or oscillation, like winter or summer in the previous example. In that case, two signals may have the same frequency (years), but they may have different phases at one specific moment. Following the same example, two ice cream shops, one in the north and the other in the south pole, will have different phases as the seasons are different in each pole. Importantly, the phase between signals is not the same as the delay parameter in the present application (Figure 1c). The phase is an intrinsic property, while the delay implies a connectivity between signals.

The third property is the noise. It refers to all the unwanted or unknown activities that we measure within our activity or interest. The higher the number of sources of noise, the more difficult to identify the oscillatory patterns (Figure 1d, yellow signal). Therefore, the noise has a direct impact on the connectivity metrics. If it is too high, the methods would not be able to infer the connectivity as the noise will hide the activity of interest.

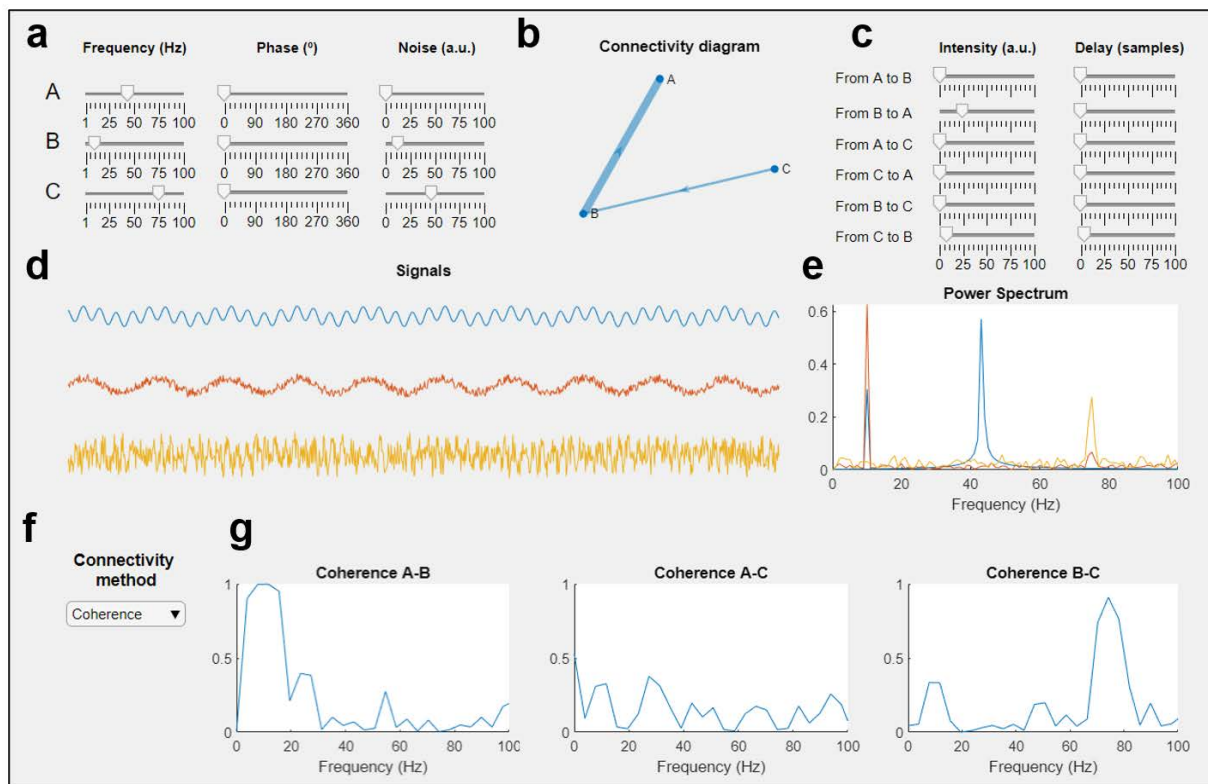


Figure 1. Application interface. (a) Intra-node parameters; (b) network scheme; (c) inter-node parameters; (d) signals; (e) power spectrum; (f) connectivity metric, (g) connectivity results.

3.2 Differentiating correlation from causality

Different networks with correlated or causal links are implemented based on the time delay between the time-series. Two variables are correlated if they share the same frequency or if there is a link between them, independently of the time delay. Thus, it is possible to obtain correlated signals by imposing the same parameter “frequency” to both, even if it is not reflected in connectivity diagram. For example, the routines of two employees are correlated if they work during the morning and rest during the evening, even if they have never seen each other. In addition, every connection between signals would imply a certain degree of correlation based on the parameter “intensity”. The higher the value, the higher the similarities between the signals and, therefore, their correlation value.

A causal link between signals requires two elements: a correlation between them and a certain time-delay in their connectivity. The effect of time delay indicates that, if two variables are causally related, a

change in one variable would influence the other variable in the future. For example, if the price of the diesel rises for one year, the number of sells of gasoline vehicles would probably increase during the next year, but there is no immediate effect. This effect is taken into account when computing Granger Causality [3], which determines if the future of one signal can be predicted based on the past values of the other variables. If it is possible, then there it can be established a causal link between the variables. This is set with the parameter “delay”, which adds a determined time that requires one signal to have an effect on the other. If set to zero, the connectivity would be instantaneous [5], while the directionality could not be determined.

4 CONCLUSIONS

Through the testing of different network configurations, students will acquire the basic knowledge of connectivity that it is applied in several fields as econometrics, physics or neuroscience. The students will develop their skills in network analysis, including the characterization of variables as time-series; their intrinsic properties, as the frequency of the oscillation, the phase or the noise; and the inter-node features, with a special focus on the time delay. Moreover, the application allows the students to understand and interpret the results of several connectivity metrics. Importantly, the proposed software can be easily updated with the integration of new methodologies and parameters based on specific objectives of the students.

REFERENCES

- [1] Kenny D. *Correlation and Causality*. Vol. 1, New York: Wiley, 1979. 1979.
- [2] Goyal S. *Connections*. Princeton University Press; 2009.
- [3] Lütkepohl H. Estimation of Vector Autoregressive Processes. In: Lütkepohl H, editor. *New Introduction to Multiple Time Series Analysis*. Berlin, Heidelberg: Springer; 2005. p. 69–133. Available from: https://doi.org/10.1007/978-3-540-27752-1_3
- [4] Kirchgässner G, Wolters J, Hassler U. Granger Causality. In: Kirchgässner G, Wolters J, Hassler U, editors. *Introduction to Modern Time Series Analysis*. Berlin, Heidelberg: Springer; 2013. p. 95–125. (Springer Texts in Business and Economics). Available from: https://doi.org/10.1007/978-3-642-33436-8_3
- [5] Baccalá LA, Sameshima K. Assessing Connectivity in the Presence of Instantaneous Causality. In: *Methods in Brain Connectivity Inference through Multivariate Time Series Analysis*. CRC Press; 2014.