Contents lists available at ScienceDirect

### Physica A

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

# Effective network inference through multivariate information transfer estimation



PHYSICA

Carl-Henrik Dahlqvist<sup>a,b,c,\*</sup>, Jean-Yves Gnabo<sup>b,c</sup>

<sup>a</sup> LFIN, Louvain Finance, Université catholique de Louvain, Place des Doyens, 1, 1348 Louvain-la-Neuve, Belgium

<sup>b</sup> CeReFiM, DeFiPP, Department of Business Administration, University of Namur, Rempart de la Vierge 8, 5000 Namur, Belgium

<sup>c</sup> NaXys, Department of Mathematics, University of Namur, Rempart de la Vierge 8, 5000 Namur, Belgium

#### HIGHLIGHTS

- Development of a new network inference algorithm to treat the issue of indirect links.
- Indirect links caused by both intra and inter-lag causal relationships are treated.
- Minimization of the number of conditions at each step of the algorithm.
- Frequency analysis of the information content of US financial network.
- Relationships between topological properties and risk investigated.

#### ARTICLE INFO

Article history: Received 2 May 2017 Received in revised form 10 November 2017 Available online 16 February 2018

*Keywords:* Effective network Indirect link Systemic risk Bank network

#### ABSTRACT

Network representation has steadily gained in popularity over the past decades. In many disciplines such as finance, genetics, neuroscience or human travel to cite a few, the network may not directly be observable and needs to be inferred from time-series data, leading to the issue of separating direct interactions between two entities forming the network from indirect interactions coming through its remaining part. Drawing on recent contributions proposing strategies to deal with this problem such as the so-called "global silencing" approach of Barzel and Barabasi or "network deconvolution" of Feizi et al. (2013), we propose a novel methodology to infer an effective network structure from multivariate conditional information transfers. Its core principal is to test the information transfer between two nodes through a step-wise approach by conditioning the transfer for each pair on a specific set of relevant nodes as identified by our algorithm from the rest of the network. The methodology is model free and can be applied to high-dimensional networks with both inter-lag and intra-lag relationships. It outperforms state-of-the-art approaches for eliminating the redundancies and more generally retrieving simulated artificial networks in our Monte-Carlo experiments. We apply the method to stock market data at different frequencies (15 min, 1 h, 1 day) to retrieve the network of US largest financial institutions and then document how bank's centrality measurements relate to bank's systemic vulnerability.

© 2018 Elsevier B.V. All rights reserved.

\* Correspondence to: University of Namur, Rempart de la Vierge, 8, 5000 Namur, Belgium. *E-mail addresses:* carl-henrik.dahlqvist@unamur.be (C.-H. Dahlqvist), jean-yves.gnabo@unamur.be (J.-Y. Gnabo).

https://doi.org/10.1016/j.physa.2018.02.053 0378-4371/© 2018 Elsevier B.V. All rights reserved.



#### 1. Introduction

Over the past decades, complex network research has received increasing attention from the literature analyzing the developments in experimental science as well as in human and social sciences. Such developments have become necessary in order to describe systems in which interactions are governed by stochastic phenomena. Along these lines, several tools have been developed in or borrowed from areas such as statistical physics, information theory and computer sciences to explore and understand further complex systems.

An important problem faced by researchers when analyzing network structure from the real world lies in the lack of reliable and comprehensive datasets on the physical relationships between each components of the system. To deal with this issue, a common approach in the literature consists in using auxiliary information such as time-series of observable features pertaining to the nodes for reconstructing the network. Areas such as epidemiology [1,2], genetics [3,4], neuroscience [5] or human travel [6] are for instance considering interacting systems that can be modeled in terms of signals propagating over underlying networks [7]. Recently, the question of network reconstruction has drawn marked interest in the field of finance amid the adoption of network perspective to analysis systemic risk. Examples of data exploited to perform such analysis are fMRI or EEG data in neuroscience, gene expression data in genetics or asset prices in finance [5,8,9].

Equipped with time series for each nodes, the reconstruction consists in making use of the statistical measurement (be it correlation, transfer entropy or Granger causality measures for instance) of temporal dependencies between the series. While convenient and intuitive, the time-series approach to network reconstruction is not immune to specific problems. One of them stems from the inability to accurately separate a direct dependence between a pair of entities from indirect effects coming through the remaining part of the network. We propose to address this issue in this contribution.

In a regression setting, this problem corresponds to the omitted-variable bias. Such a bias, in a network context, could end-up to wrongly associate two nodes that are not directly connected (false positive) or fail to detect a connection that exists (false negative) because the parameters capturing their relationship are not accurately estimated. Traditionally, this is dealt with by extending the set of independent variables in the regression to control for the auxiliary effects. However, when the network is large, the solution is not trivial, as one potentially needs to add a very large set of regressors, the so-called control variables, especially if the dependence stems from lagged observations. Without a specific treatment of the problem, statistical tests, whether in the absence of controls or conversely in the presence of an excessive number of controls, are likely to perform poorly.

Different strategies are proposed in the econometric literature to deal with or at least mitigate the aforementioned problem. Among them we can notably quote the least absolute shrinkage and selection operator (LASSO) technique, developed two decades ago [10], and recently used for financial networks reconstruction [11] or the so-called general-to-specific simplification algorithm developed by David Hendry and Hans-Martin Krolzig [12–14]. Alternative strategies have been proposed outside the field of econometrics to eliminate spurious relationships from time series data such as the silencing approach [15] used to separate the direct and indirect links in biological networks or the decomposed transfer entropy proposed by [16]. Other approaches based on the concept of step algorithm developed in the general-to-specific model have proven capable to infer effective networks (see, among others, [3,4,17–19]). The main idea is to consider the environment around individuals transmitting and receiving the information. Conditional measures are used to identify the true path through which information travels and to remove possible redundancies in the information transfer. Two types of dependencies are considered in the literature, the instantaneous dependence estimated using correlation or mutual information and then studied using methodologies such as the directed acyclic graph (DAG) technique [17,20], and the dynamic dependence focusing on lagged relationships and estimated using transfer entropy or Granger causality [4,16,21]. This second approach, which we aim to further extend, provides more comprehensive information on the network, as it enables to feature the direction of edges between two nodes.

In this study, we propose to follow the literature on step algorithms using both the pre-search and pruning steps and develop a methodology that combines conditional information transfer and greedy algorithm to infer effective networks. The effective network approach [22] attempts to infer a minimal topological structure, regrouping only the directed relationships that can properly describe the evolution of the system. We further extend this approach and consider the possible spurious connections due to both instantaneous and lagged causal relationships by applying an extended set of possible conditions. As described above, the conditional information transfer approach faces the problem of dimensionality stemming from both the large number of nodes in the system and the fact that the relationship can exist for different lags. We pay particular attention in the algorithm to minimize the number of conditions for each information transfer by considering only the most relevant ones. Overall, this approach enables us to detect redundancies (i.e., lower the false positive rate) as well as to mitigate the loss of efficiency stemming from high dimensionality (i.e., lower the false negative rate).

In a series of simulation exercises using artificial networks, we confirm that our methodology performs well compared to a set of state-of-the-art methodologies such as the global silencing approach. Equipped with our approach, we investigate in the application part, the link between the banks' topology and systemic fragility. The past decade shows an increasing interest among the academic community and financial regulators to adopt a network perspective to describe and analyze financial markets. From the regulators' perspective, such an approach aims to identify critical relationships in the market as well as to improve regulatory policies. Following the recent studies of [8,23] or [11], who used asset prices to retrieve financial networks and analyze their properties, we use data on large US banks to assess the explanatory power of topological properties such as in-degree and out-degree to describe financial institution fragility. As described in Section 5, our approach is well designed to deal with several important aspects of network reconstruction in finance.

Download English Version:

## https://daneshyari.com/en/article/7375798

Download Persian Version:

https://daneshyari.com/article/7375798

Daneshyari.com