Physica A xx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Physica A

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

Phase synchronization based minimum spanning trees for analysis of financial time series with nonlinear correlations

Q1 Srinivasan Radhakrishnan [c], Arjun Duvvuru [b], Sivarit Sultornsanee [a], Sagar Kamarthi [c,*]

[a] School of Business University of the Thai Chamber of Commerce, Bangkok, Thailand
[b] Quality Engineer at JDA Software Pvt Ltd, India
[c] Department of Mechanical and Industrial Engineering, Northeastern University, Boston, USA

HIGHLIGHTS

- We analysed financial time series using a phase synchronization (PS) measure.
- We built a network where times series are nodes and PS measure is a link strength.
- We analysed characteristics of the minimum spanning tree of the time series network.
- We compared the performance of PS measure with that of cross correlation measure.
- We found PS measure is more suitable for time series with time-lagged correlations.

ARTICLE INFO

Article history:
Received 7 March 2015
Received in revised form 31 August 2015
Available online xxxx

Keywords:
Phase synchronization
Time series
Financial systems
Stock correlation network
Exchange rate networks
Complex networks

ABSTRACT

The cross correlation coefficient has been widely applied in financial time series analysis, in specific, for understanding chaotic behaviour in terms of stock price and index movements during crisis periods. To better understand time series correlation dynamics, the cross correlation matrices are represented as networks, in which a node stands for an individual time series and a link indicates cross correlation between a pair of nodes. These networks are converted into simpler trees using different schemes. In this context, Minimum Spanning Trees (MST) are the most favoured tree structures because of their ability to preserve all the nodes and thereby retain essential information imbued in the network. Although cross correlations underlying MSTs capture essential information, they do not faithfully capture dynamic behaviour embedded in the time series data of financial systems because cross correlation is a reliable measure only if the relationship between the time series is linear. To address the issue, this work investigates a new measure called phase synchronization (PS) for establishing correlations among different time series which relate to one another, linearly or nonlinearly. In this approach the strength of a link between a pair of time series (nodes) is determined by the level of phase synchronization between them. We compare the performance of phase synchronization based MST with cross correlation based MST along selected network measures across temporal frame that includes economically good and crisis periods. We observe agreement in the directionality of the results across these two methods. They show similar trends, upward or downward, when comparing selected network measures. Though both the methods give similar trends, the phase synchronization based MST is a more reliable representation of the dynamic behaviour of financial systems than the cross correlation based MST because of the former's

* Corresponding author. Tel.: +1 617 373 3070.
E-mail address: sagar@coe.neu.edu (S. Kamarthi).

http://dx.doi.org/10.1016/j.physa.2015.09.070

ability to quantify nonlinear relationships among time series or relations among phase shifted time series.

1. Cross correlation based networks

Segmentation and segment clustering techniques give a macro view of a dynamical process embedded in time series pertaining to economics and finance [1]. However the application of these techniques to high frequency time series creates a significant exclusion of important information. In order to retain richer information a cross-correlation technique is preferred in which the cross correlation value is computed between two time series i and j (without lag) using Eq. (1).

$$C_{i,j} = \frac{\sum_{t=1}^{T}(x_{it} - \bar{x}_i)(x_{jt} - \bar{x}_j)}{\sqrt{\sum_{t=1}^{T}(x_{it} - \bar{x}_i)^2 \sum_{t=1}^{T}(x_{jt} - \bar{x}_j)^2}}. \quad (1)$$

The cross correlation coefficient has been widely applied to analysing stock price and index movement time series, in specific, for understanding chaotic behaviour during crisis periods [2–8]. Several studies have focused on understanding non-trivial cross correlation using random matrix theory in which the cross correlation is computed over either entire data or certain pre-set time windows [1,9–16]. The information dispersed in independent cross correlation matrices is difficult for human understanding and interpretation. In an attempt to better understand correlation dynamics, the cross correlation matrices are given graphical representations. Individual time series is denoted as a node and individual nodes are connected to each other in accordance with pre-formed rules. These rules determine the outcome of the representation, which could be either a tree or a graph. In case of a tree, all the nodes are retained with acyclic connections (cyclic connections are inhibited by connection criteria). In contrast, graphs by virtue of its connection criteria allow formation of isolated nodes and loops. The resultant graphical representation in case of graphs may vary significantly compared to that of a tree. In this work we limit our focus to trees only.

1.1. Graphical representation: tree approach

The minimum spanning trees (MST) are weighted graphical representations, the construction of which can involve one of the two widely used algorithms, namely, Kruskal's algorithm [17] and Prim's algorithm [18]. MSTs found few applications in the field of economics until Mantegna [19] indicated robust patterns in the underlying correlations [20,21]. Since then MSTs have been widely used for statistical analysis of financial market data [22–28].

1.2. Cross correlation based MST

The following steps are executed in order to create MSTs from computed cross correlation between two time series.

1. Individual time series are considered as nodes in the network.
2. The cross correlation denoted by C_{ij} is computed for each pair of time series, i and j, using Eq. (1).
3. The correlation coefficients forming an $N \times N$ correlation matrix with $-1 \le C_{i,j} \le 1$ are transformed into a $N \times N$ distance matrix with elements $d_{ij} = \sqrt{2(1 - C_{i,j})}$, such that $0 \le d_{ij} \le 2$. The symmetric property of the distance formula ensures that $d_{ij} = d_{ji}$. The triangular property reveals the relationship between the distance value and the correlation coefficient (smaller distance values indicate higher correlation values).
4. The distance matrix is essentially an adjacency matrix representing the correlation network. The distance matrix is then used to determine the minimum spanning tree (MST), which is simply a connected graph that connects all the N nodes of the network with $(N - 1)$ edges such that the sum of all distances is minimum.
5. Kruskal algorithm[1] is applied to the distance matrix to form the desired MST [17].

In this work we investigate a new approach that has not been explored earlier for establishing reliable correlations among financial time series (e.g. stocks). The advantage of this new method is that it can characterize the synchronized variations among different time series better than the measures, such as cross correlation, that capture only the linear relationships.

[1] Kruskal's algorithm and Prims algorithm are two widely adopted methods to from MST. From space aspect, Kruskal's algorithm is relatively superior to Prims algorithm when the number of nodes are less than 100, however when the number of nodes are greater than 100, from the time complexity's aspect, Prim algorithm is superior [29]. In this work we adopt Kruskal's algorithm for converting distance matrix into MST [17].

Please cite this article in press as: S. Radhakrishnan, et al., Phase synchronization based minimum spanning trees for analysis of financial time series with nonlinear correlations, Physica A (2015), http://dx.doi.org/10.1016/j.physa.2015.09.070