Physica A 396 (2014) 235–241

Contents lists available at ScienceDirect

Physica A

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

An analysis of the sectorial influence of CSI300 stocks within the directed network

Yong Mai *, Huan Chen, Lei Meng

School of Business, East China University of Science and Technology, 130 Meilong Road, Shanghai 200237, China

ARTICLE INFO

Article history:
Received 3 June 2013
Received in revised form 23 September 2013
Available online 22 November 2013

Keywords:
Directed weighted network
Scale-free
Hierarchical structure

ABSTRACT

This paper uses the Partial Correlation Planar maximally filtered Graph (PCPG) method to construct a directed network for the constituent stocks underlying the China Securities Index 300 (CSI300). We also analyse the impact of individual stocks. We find that the CSI300 market is a scale-free network with a relatively small power law exponent. The volatility of the stock prices has significant impact on other stocks. In the sectorial network, the industrial sector is the most influential one over other sectors, the financial sector only has a modest influence, while the telecommunication services sector's influence is marginal. In addition, such inter-sector influence displays quarterly stability.

1. Introduction

In recent years the studies on topological properties of the financial network have advanced both in depth and width. Many different models and analytical approaches have been applied to study the structure and dynamics of financial markets. Correlation-based networks are widely used methods, including the threshold network, the minimum spanning trees, and the Planar Maximally Filtered Graphs (PMFG) [1–4]. These methods usually use the Pearson correlation coefficient to investigate the relationship of different stocks. A hierarchical structure and a cluster formation have been investigated through the PMFG method in the New York Stock Exchange (NYSE) [4–6]. Studies on NYSE and the Shanghai Stock Exchange show that the stock markets demonstrate scale-free characteristics [7,8]. These studies show that the scale-free network is commonly observed in stock markets and with core nodes underpinning the network. The volatility of stocks represented by those nodes can cause profound impact on other stocks in the market.

However, the previous literature does not involve the directed issue of the inter-node influence. In particular, the mutual influence of individual stocks and sectors within the network is rarely studied amid the directed network of stock trading. The coarse-graining symbols method has been used to construct a directed weighted network for the Hang Seng index (HSI), and reveals some patterns of HSI fluctuation have remarkable statistical stability [9]. Partial correlation is applied to propose a directed network, i.e. the Partial Correlation Threshold Network (PCTN) and the Partial Correlation Planar maximally filtered Graph (PCPG), and leads to the finding that financial stocks are the most influential stock in the NYSE [10]. Similar research is rare in stock markets of developing economies. This motivates the analysis of this paper which examines the inter-node influence within the network of the CSI300 index basket, as well as the stability of such influence.

The organisation of this paper is as follows. Section 2 presents the data and methodology. In Section 3 the PMFG and PCPG are constructed using individual stocks underlying the CSI300 index. We also analyse the topological properties of the network. In Section 4 we construct a sectorial network structure and investigate the network affiliation of intra- and inter-network influence of ten sectors, and then check the stability of the results with moving windows. Section 5 concludes.

* Corresponding author. Tel.: +86 013817231962; fax: +86 02164253152.
E-mail addresses: maiyong1151@163.com, maiy@ecust.edu.cn (Y. Mai), carlan.chen@gmail.com (H. Chen), leimeng@ecust.edu.cn (L. Meng).

0378-4371/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.physa.2013.11.016

2. Methodology and data

2.1. Partial correlation and partial correlation planar maximally filtered graph

We define the logarithmic return of the stock price as:

$$R(t) = \ln p(t) - \ln p(t-1) \tag{1}$$

where t denotes the day, $p(t)$ and $p(t-1)$ are the prices on days t and $t-1$. Thus, the return indicates the magnitude of rise or fall during t and $t-1$. For the ease of comparison among different sorts of data, the return given above can be standardised to the following:

$$r(t) = \frac{R(t) - \langle R(t)\rangle}{\sqrt{\langle R(t)^2\rangle - \langle R(t)\rangle^2}} \tag{2}$$

where $\langle\cdots\rangle$ represents the averaging process. Then, the correlation between stock i and stock j is defined as:

$$\rho_{ij} = \frac{\langle r_i r_j\rangle - \langle r_i\rangle\langle r_j\rangle}{\sqrt{(\langle r_i^2\rangle - \langle r_i\rangle^2)(\langle r_j^2\rangle - \langle r_j\rangle^2)}}. \tag{3}$$

The Pearson correlation indicates the correlation between the returns of two stocks, which cannot be used to describe the impact of other stocks in the system on the relationship between two stocks [11–14]. Partial correlation statistics has been applied to investigate how a market index affects the relationships among stocks traded in the market [15].

The partial correlation method calculates the correlation coefficient between two stocks while eliminating the impact of other variables. Assume that there are p random variables $X_1, X_2, \ldots, X_p$. Eliminate the impact of $X_3, \ldots, X_p$ from X_1 and X_2, and denote the remainder X_1' and X_2'. In fact, X_1' and X_2' are the optimal linear fitting based on $X_3, \ldots, X_p$ against X_1' and X_2'. The simple correlation coefficient is the partial correlation coefficient for X_1' and X_2' against $(X_3, X_4, \ldots, X_p)$, denoted as $\rho(X_1, X_2 : X_3, X_4, \ldots, X_p)$. In fact, there are three methods to compute partial correlation coefficient. The first method, mentioned above, is to use the linear regression. The second is to use the correlation matrix inversion. The third one involves the iterative method, i.e. when taking simple correlation coefficients as zero-order partial correlation coefficients, any n-order partial correlation coefficient can be calculated through three $n-1$ order partial correlation coefficients. The iterative method has been used to calculate the first-order partial correlation coefficients and we propose a method to construct a directed network [10]. The PCPG is an adaptation of the PMFG to deal with asymmetric interactions among the elements of a system. This paper applies the PCPG method to construct directed networks. With variables X_1, X_2, X_3, the first-order partial correlation coefficient $\rho(X_1, X_2 : X_3)$ can be computed through three zero-order partial correlation coefficients $\rho(X_1, X_2)$, $\rho(X_1, X_3)$ and $\rho(X_2, X_3)$:

$$\rho(X_1, X_2 : X_3) = \frac{\rho(X_1, X_2) - \rho(X_1, X_3)\rho(X_2, X_3)}{\sqrt{[1-\rho^2(X_1, X_3)][1-\rho^2(X_2, X_3)]}}. \tag{4}$$

To measure the magnitude of impact from X_3 to X_1 and X_2, we define the following correlation influence index of X_3 on the pair of elements X_1 and X_2:

$$d(X_1, X_2 : X_3) = \rho(X_1, X_2) - \rho(X_1, X_2 : X_3). \tag{5}$$

Obviously, the bigger the value of $\rho(X_1, X_2 : X_3)$, the stronger the impact X_3 has on X_1 and X_2. Next we define the average correlated influence index as $d(X_1 : X_3)$:

$$d(X_1 : X_3) = \langle d(X_1, X_2 : X_3)\rangle_{X_2 \neq X_1, X_3}. \tag{6}$$

The relationship of $d(X_1 : X_3) \neq d(X_3 : X_1)$ can be used to treat the asymmetric correlation among elements in a system. In addition, the directed weighted network, i.e. PCPG, can be constructed through the PMFG method.

The procedures for building PCPG are as follows: (1) Calculate the mean correlation influence indices between all nodes using Eq. (6), i.e. there are $N(N-1)$ influence indices (the influence of each node upon themselves is not considered, i.e. N influence indices have the value of 0), which are listed in descending order. (2) Then we introduce an empty network with N nodes, and add pairs according to the sequence of the previous list when and only when the new network is still plain. For example, to judge when a pair is qualified, let it be $d(I : J)$. If the network is still plain after adding $J \to I$ (if a sphere can be drawn without inter-crossing links) then the pair is added, otherwise the pair is discarded. In addition, to avoid overlapping links and to keep only the essential information, if $d(I : J) > d(J : I)$, only $J > I$ is considered for the PCPG. (3) The above procedures repeat until the number of nodes reaches $3(N-2)$ since that is the maximum number of links that can be allowed under the constraints for a plain graph.

2.2. Data

The CSI300 index is normally rebalanced every six months except for temporary adjustments. In order to obtain a reasonably long sample period, we collect the daily data of CSI300 index constituent stocks adjusted on 22 December 2009 from 28 September 2009 to 30 March 2012 (excluding weekends and market closed dates). After collecting historical end of