Physics Letters A 380 (2016) 654–666

Contents lists available at ScienceDirect

Physics Letters A

www.elsevier.com/locate/pla

Structure and dynamics of stock market in times of crisis

Longfeng Zhao *, Wei Li, Xu Cai

Complexity Science Center & Institute of Particle Physics, Hua-Zhong (Central China) Normal University, Wuhan 430079, China

ARTICLE INFO

Article history:
Received 4 March 2015
Received in revised form 15 October 2015
Accepted 8 November 2015
Available online 26 November 2015
Communicated by F. Porcelli

Keywords:
Correlation-based network
Financial crisis
Market structure
Market dynamics
Power-law distribution
Market backbone

ABSTRACT

Daily correlations among 322 S&P 500 constituent stocks are investigated by means of correlation-based (CB) network. By using the heterogeneous time scales, we identify global expansion and local clustering market behaviors during crises, which are mainly caused by community splits and inter-sector edge number decreases. The CB networks display distinctive community and sector structures. Graph edit distance is applied to capturing the dynamics of CB networks in which drastic structure reconfigurations can be observed during crisis periods. Edge statistics reveal the power-law nature of edges' duration time distribution. Despite the networks' strong structural changes during crises, we still find some long-duration edges that serve as the backbone of the stock market. Finally the dynamical change of network structure has shown its capability in predicting the implied volatility index (VIX).

1. Introduction

Recent subprime mortgage crisis has attracted very much attention on studying the relationship between financial market structure and economic crises [1–10]. Apparently new studies are needed towards such a purpose [11]. Benefiting from the works on correlation-based (CB) network of financial asset return time series [12–16], we can apply the methodology of network theory [17] to analyzing the financial market.

As a powerful tool to understand the properties of stock return time series [12], the CB network method facilitates the stock market research by techniques of complex network analysis [18]. Some useful methods to construct CB networks are minimum spanning tree (MST) [12,19–21], threshold cutting procedure (the winners-take-all strategy) [22–25], planar maximally filtered graph (PMFG) [14], dependency network [26,27], bootstrap method [28], etc.

The researches on extreme events and economic crises have been very intriguing. Black Monday crash has been analyzed by means of asset tree with MST method. The shrink in asset tree length and degree distribution difference are observed after the crash [20]. The asset tree research on financial indices also reveals the topological structure changing from star-like to chain-like during crisis [5]. The principal components analysis indicates an increase in the strength of the relationship between several different markets during subprime mortgage crisis [2] and is also employed as a measure of systemic risk [29,30]. By analyzing the eigenvalues and eigenvectors of correlation matrices, direct links between high volatility and strong correlations can be recognized, which means the market tends to behave like the one during big crashes [3]. Regression techniques have been used to identify the strong clustering behavior distributed by geographic differences [31] for stocks from different countries and sectors. A newly introduced measure named sector dominance ratio is capable of capturing the economic sectoral activities [32]. The main information that can be inferred from the previous works mentioned above is that the clustering and correlation strengthen behaviors that can be observed during crises. However, detailed information about structure and dynamics of stock market in times of crisis is still lacking.

Here in this article, we aim at exploring the detailed structure and dynamics of stock market during crisis periods from the aspect of CB networks with planar maximally filtered graph (PMFG) method [14].

We investigate the correlations among 322 S&P 500 constituent stocks by using the stocks' daily adjusted closure price return time series between January 1994 and January 2014. In order to filter the influences of economic crises, we first use the entire historical data records (heterogeneous time scales) to calculate the correlation coefficients between pairs of stocks. The PMFG method is then adopted to construct CB networks, which yields very clear modular structures. We argue that the heterogeneous time scales can characterize the slow dynamics of the market [33] by reducing the fluctuations more efficiently as compared to the traditional truncated method. We hereby identify two of the most serious

* Corresponding author.
E-mail address: zlfccnu@mails.ccnu.edu.cn (L. Zhao).

http://dx.doi.org/10.1016/j.physleta.2015.11.015
0375-9601/

crisis periods of US stock market during recent twenty years, dot-com bubble and subprime mortgage crisis. We observe very similar trends of the market during these two periods. Namely, the clustering coefficients and the shortest-path lengths are always positively correlated under almost all time scales ranging from 1 month to 100 months, as well as under the heterogeneous time scales. The key result obtained from the PMFG topological analysis is that during the crisis periods both shortest-path length and clustering coefficient increase, a clear signal of global expansion and local clustering behaviors. This finding can be well explained by network community detection [34] and sector relationship networks. The community splits and decline of the overlap between communities have been unfolded during the crisis periods. It has been noticed that the inter-sector edge decrease and the intra-sector edge increase happen simultaneously. The statistical quantities such as modularity of PMFGs and the edges of the sector relationship networks have been used to evaluate the structural evolution. We then use graph edit distance [35] and edge statistical analysis to investigate the dynamics of PMFGs. There exist abrupt increases in edit distance at the onset time points of crises. We show that the edges' duration time follows a power-law distribution. A set of edges with very long duration time have been observed which serve as the backbones of the stock market. As an illustration of applications of our results, the successive edit distance has proved to Granger cause the implied volatility index (VIX). The successive edit distance is capable of helping predict the implied volatility. We further use the cross correlation analysis to study the correlation between successive edit distance and VIX. It turns out that VIX is similar to successive edit distance in the previous month.

The paper is organized as follows. In Sec. 2, we describe the data, the methodology and the selection of proper time scales by using different quantities in both short and long time scales. In Sec. 3, we analyze the evolution of topology parameters of CB networks under the heterogeneous time scales. We also show the topological structural evolution filtered by PMFG method. In Sec. 4, we discuss the community and sector structural evolution. In Sec. 5 we analyze the dynamics of PMFGs by means of edit distance, as well as the statistical properties of edges. In Sec. 6 an application of our results is demonstrated. The last section is discussion and summary.

2. Data, correlation-based networks, and time scales

2.1. Data

Our data sets include 322 stocks (see Appendix A). They are the constituent stocks of S&P 500 between 1st January 1994 and 1st January 2014. We adopt the logarithm return defined as

$$r_i(t) = \ln p_i(t+1) - \ln p_i(t), \tag{1}$$

where $p_i(t)$ is the adjusted closure price of stock i at time t. We then compute the mutual correlation coefficients between any pair of return time series at time t by using the past return records sampled with different length Δ ranging from 1 month to 100 months (amounting to nearly 2500 trading days). Following Ref. [20] one can evaluate the similarity between stocks i and j at time t with the Pearson correlation coefficient by,

$$\rho_{ij}^{t,\Delta} = \frac{\langle R_i^t R_j^t\rangle - \langle R_i^t\rangle\langle R_j^t\rangle}{\sqrt{\left[\langle R_i^{t^2}\rangle - \langle R_i^t\rangle^2\right]\left[\langle R_j^{t^2}\rangle - \langle R_j^t\rangle^2\right]}}, \tag{2}$$

where Δ is the estimation interval, and $\langle \ldots \rangle$ is the sample mean over co-trading days of stocks i and j in the logarithm return series vector $R_i^t = \{r_i(t)\}$ and $R_j^t = \{r_j(t)\}$. Then we obtain the $N \times N$ matrix $C^{t,\Delta}$ at time t with estimation interval Δ, and N is the number of stocks. The matrix entries of $C^{t,\Delta}$ are the correlation coefficients $\rho_{ij}^{t,\Delta}$ between all pairs of stocks.

2.2. Correlation-based networks and time scales

The planar maximally filtered graph (PMFG) method [14] is employed to construct networks based on the correlation matrices $C^{t,\Delta}$. The algorithm is implemented as follows,

(i) Sort all of the $\rho_{ij}^{t,\Delta}$ at time t with time scale Δ in descending order to obtain an ordered list l_{sort}.

(ii) Add an edge between nodes i and j based on the order in l_{sort} only if the graph is still planar after edge addition.

(iii) A graph, $G_P^{t,\Delta}(V,E)$, is formed with $NUM_e^{single} = 3(N-2)$ edges under the constrain of planarity.

As described in Ref. [14] PMFGs not only keep the hierarchical organization of minimum spanning tree (MST) but also generate some cliques. We calculate the basic topological parameters such as clustering coefficient C, shortest-path length L, as well as average correlation coefficient ρ. In addition, we adopt a quantity γ [36] to measure the heterogeneity of PMFGs, whose definition is given by,

$$\gamma = \frac{N - 2\sum_{ij\in e}(k_i k_j)^{-1/2}}{N - 2\sqrt{N-1}}, \tag{3}$$

where k_i is the degree of vertex i.

The magnitude of Δ is crucial for analyzing the dynamics of financial market [33,37,38]. Short estimation interval is suitable for analysis of fast dynamics but with relatively large statistical uncertainty [33]. This means the influence of financial crises on structure of CB networks could be indistinguishable under improper estimation intervals. In Figs. 1 and 2 we present the influences of different estimation intervals varying from 1 month to 100 months via the evolution of four quantities ρ, C, L and γ. Here we only show four quantities during the time window from January 1999 to January 2014. If the time window is enlarged, the estimation interval cannot be extended to 100 months. We notice that when the estimation interval is shorter than 10 months, the fast dynamics can be observed. With the increase of Δ, fluctuations of quantities decrease and the influences of major economic crises can be recognized. The influences of small shocks will be smeared out under large estimation intervals. We use different slide window sizes in Figs. 1 and 2, which is 1 week and 1 month respectively. This setting is a tradeoff between accuracy and computation load due to the algorithm complexity of planarity test [39]. When the estimation interval is large, slow dynamics will dominate the evolution of those quantities. So larger slide window size is more proper for capturing the evolution of the market under larger time scales. In Figs. 1 and 2, the average correlation ρ behaves very differently during dot-com bubble and subprime mortgage crisis. No abrupt increase of average correlation can be observed during dot-com bubble even under very short estimation intervals. This might be caused by the bankruptcy of many internet companies during dot-com bubble. However, ρ increases very fast at Lehman's failure in September 2008. The clustering coefficient C and shortest-path length L show clear signals about major crises under estimation interval larger than 60 months in Fig. 2.

Figs. 3(a)–(d) are the topological structures of PMFGs under different estimation intervals at 01\01\2014. The Fruchterman–Reingold layout [40] has been used to demonstrate the structures of networks. Stocks from different sectors do not show clear structures under estimation interval from 10 months to 100 months.