Physica A 413 (2014) 523–533

Contents lists available at ScienceDirect

Physica A

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

Dynamics of cluster structures in a financial market network

Anton Kocheturov [a,*], Mikhail Batsyn [a], Panos M. Pardalos [a,b]

[a] Laboratory of Algorithms and Technologies for Network Analysis, National Research University Higher School of Economics, 136 Rodionova, Nizhny Novgorod, Russian Federation

[b] Center of Applied Optimization, University of Florida, 401 Weil Hall, P.O. Box 116595, Gainesville, FL 32611-6595, USA

HIGHLIGHTS

- We introduced a new approach for the financial market analysis.
- The approach reveals financial crises in the American and Swedish markets.
- It provides a more contrasting picture of a crisis than other approaches.
- It detected the difference in behavior of Swedish and US companies during crisis periods.

ARTICLE INFO

Article history:
Received 6 February 2014
Received in revised form 7 May 2014
Available online 5 July 2014

Keywords:
Dynamics
Financial markets
P-Median Problem
Clustering
Crisis

ABSTRACT

In the course of recent fifteen years the network analysis has become a powerful tool for studying financial markets. In this work we analyze stock markets of the USA and Sweden. We study cluster structures of a market network constructed from a correlation matrix of returns of the stocks traded in each of these markets. Such cluster structures are obtained by means of the P-Median Problem (PMP) whose objective is to maximize the total correlation between a set of stocks called medians of size p and other stocks. Every cluster structure is an undirected disconnected weighted graph in which every connected component (cluster) is a star, or a tree with one central node (called a median) and several leaf nodes connected with the median by weighted edges. Our main observation is that in non-crisis periods of time cluster structures change more chaotically, while during crises they show more stable behavior and fewer changes. Thus an increasing stability of a market graph cluster structure obtained via the PMP could be used as an indicator of a coming crisis.

1. Introduction

The main goal of the paper is to analyze the dynamics of a cluster structure in financial markets and find its connection to crisis and non-crisis periods. Network-based approaches have become a significant tool for the analysis of complex dynamic systems arising in finance such as financial markets and interbank networks. These approaches reveal a great deal of useful information: Minimum Spanning Trees (MSTs) can, for example, provide us with the hierarchical structure existing in a market [1] or detect critical banks in banking networks and estimate their roles [1–4]. Cliques [5], community structures [2], MSTs [3] and cluster structures [6] of a market network can be used in size reduction, core nodes detection, and supervising of the whole network. Independent sets can serve as a tool for solving the diversified portfolio selection problem [5].

There are a number of papers committed to analyzing stock markets of certain countries and their dynamics [7–11]. The majority of them use the network-based approach introduced by Mantegna [1] who suggested building the Minimum

* Corresponding author. Tel.: +7 9527740309.
E-mail addresses: antrubler@gmail.com (A. Kocheturov), mbatsyn@hse.ru (M. Batsyn), pardalos@ufl.edu (P.M. Pardalos).

http://dx.doi.org/10.1016/j.physa.2014.06.077
0378-4371/

Spanning Tree on the similarity matrix of all stocks in the DJIA index. The similarity between two stocks is measured as the estimation of the Pearson's correlation coefficient between returns of these stocks. Some authors apply other network structures or characteristics to study markets [5,12–14]. Some works (e.g., Ref. [14]) analyze the behavior of several network characteristics which are not based on special network structures over different time periods. This work is based on and is a continuation of an approach of Goldengorin et al. [6] to use the P-Median Problem for studying financial markets. We pay special attention to the paper of Buccheri et al. [15] in which the results confirm our observations for crisis periods: the authors traced the behavior of the second eigenvalue of the correlation matrix between 49 industry indices of US stocks and found the peaks of it during the Dot-com bubble in 2000–2001 and the first phase of the world financial crisis in 2008. During these crises we also observe a specific behavior of the cluster structures we study—the increasing stability of these structures.

We suggest a new approach which reveals a cluster structure in a market graph and evaluates its dynamics over time. In this work we study two financial markets: the stock markets of the USA and Sweden. We use the end of day (EOD), or the closure, prices of stocks traded in these markets in order to estimate the similarity matrix of stocks. We apply this measure to build the network structure which has a predefined number of connected components where every component is a star: it has a central node, or median, which is indeed some stock and a number of other nodes connected to this median. From a general point of view such a structure is a forest of stars with weighted edges where every node represents a stock and weights are equal to the measured similarity between stocks.

We build such structures by means of the P-Median Problem (a detailed survey of the problem and methods to solve it can be found in Refs. [16,17]). This problem allows us to split all stocks of a market into a predefined number of stars in such a way that the total similarity, or the sum of all weights of edges in this structure, is maximized. In other words, we have a set of connectivity components and inside every component the similarity between the stocks and the median node is large.

We divide observations into subsets for different time intervals whose length equals one calendar year. On average, there are 251 observations per year for the USA and 261 observations for Sweden. Then we build the cluster structure for every time interval and for every p where $p = 1, \dots, n-1$, where n is the number of stocks in a market. For instance, in the Swedish market we have 145 stocks, and 13 time intervals (13 years of observations); hereby we have to calculate $(145 - 1) * 13 = 1872$ cluster structures. We have decided not to use a time window because there will be too many structures to be clustered and we will not be able to calculate them in reasonable time to study the dynamics.

The next step is to compare cluster structures for a different t to evaluate the dynamics. Assume, $G_{p,t}$ is a p-cluster structure (a cluster structure of p stars) for the time period t. In order to reveal the dynamics we compare the cluster structures in the following pairs: $\{(G_{p,1}, G_{p,2}), \dots, (G_{p,T-1}, G_{p,T})\}\ \forall p = 1, \dots, n$.

We introduce two similarity measures between two cluster structures in order to compare these pairs. These measures demonstrate a specific behavior of cluster structures in crisis periods: the cluster structures tend to be more stable during crises than at the usual time.

2. Data and similarity measures

In our paper we consider American companies which were in the S&P100 index at the end of the year 2012. We took 13 years of observations from January 3, 2000 till December 31, 2012 which include 3269 EOD prices for 90 stocks. We split this time period into smaller ones covering a year of observations. Such a splitting gives us intervals big enough to accumulate statistics and small enough to study dynamics. We have not taken into consideration the remaining 10 stocks of the S&P100 index because they were traded less than 80% of all trading days during at least one of the time subintervals. The second market we study is the Swedish financial market with 266 companies. Only 145 of them were traded not less than 80% of all trading days from January 3, 2000 till January 1, 2013. This time period consists of 3392 trading days or 13 calendar years.

If we have missing data for a certain stock we fill in omissions in the following way: assume that we know the prices $P_{i,t}$ and $P_{i,t+k}$ of the stock i in the trading days t and $t+k$ and there are missing values for days $t+1, t+2, \dots, t+(k-1)$. Due to the fact that a company is not traded these days we can assume that the price of its share stays constant. Hereby we set prices $P_{i,t+1}, P_{i,t+2}, \dots, P_{i,t+(k-1)}$ equal to $P_{i,t}$. If a time interval of the stock i starts with missing values and the first known price is $P_{i,t}$ then we set prices $P_{i,1}, P_{i,2}, \dots, P_{i,t-1}$ equal to $P_{i,t}$.

We use the estimation of Pearson's correlation coefficient between stock returns as a similarity measure between stocks. In order to calculate the correlation matrices for the markets we apply the following formula:

$$\rho_{ij} = \frac{E\left\{(R_i - E\{R_i\})\left(R_j - E\left\{R_j\right\}\right)\right\}}{\sqrt{var\,(R_i)\,var\,\left(R_j\right)}}, \tag{1}$$

which gives the correlation coefficient between prices of two stocks i and j. R_i is a stock return obtained from the original set of prices according to the following formula: $R_i(t) = \ln \frac{P_i(t)}{P_i(t-1)}$, $P_i(t)$ is a closure price of the stock i at the day t.

3. Cluster structures and the P-Median Problem

A method of clustering consists in dividing all stocks presented in the market into several groups in which stock returns are strongly correlated (for details see Ref. [6]). For this purpose we calculate the correlation matrix $\mathrm{P} = \left[\rho_{ij}\right]_{n\times n}$ with formula (1). The main idea of the clustering is to find a set S of stocks or medians (which will be centers of stars) of a predefined