Physica A 395 (2014) 261–268

Contents lists available at ScienceDirect

Physica A

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

# Grouping characteristics of industry sectors in financial markets

Gabjin Oh *

*Division of Business Administration, Chosun University, Gwangju 501-759, Republic of Korea*

## HIGHLIGHTS

- US and Korean stock markets show different features in grouping coefficient of industrial sectors.
- We investigate grouping coefficient of industrial sectors.
- Grouping coefficient of industrial sectors is calculated by minimal spanning tree.
- Conglomerates play important role determining the grouping mechanism.

ARTICLE INFO

*Article history:*
Received 4 April 2013
Received in revised form 15 September 2013
Available online 7 October 2013

*Keywords:*
Grouping coefficient
Complex network
Minimum spanning tree

ABSTRACT

We investigated the grouping coefficients of industrial sectors in the stock network based on stock data for the U.S. and Korean stock markets. These complex networks were modeled using the minimal spanning tree (MST) method. We propose a novel approach based on the shortest path length (SPL) between stocks to quantify the grouping characteristics of the industrial sectors. We find that the grouping coefficients for the industrial sector in the U.S. are larger than those of the Korean stock market. In particular, for the Korean stock market the conglomerates, comprised of a diverse of industrial companies, have a significant grouping coefficient.



## 1. Introduction

Recently, the complex systems of financial markets have been increasingly studied by both physicists and economists [1,2]. The stock market has evolved due to non-trivial interactions between heterogeneous agents such as noise and fundamental traders [3] and due to events generated by the internal and external market forces. Therefore, stock price properties cannot be explained by the traditional pricing model based on the efficient market hypothesis (EMH) [4,5] and so-called "stylized facts" [6–20]; prices of individual stocks in financial markets are determined by interactions between traders who use various investment strategies and by unpredictable external shocks. Moreover, the stock market shows a complex connection structure among the formation of various stock groups according to the characteristics of specific industries [21–23].

Previous studies have found that interactions among stocks in financial markets deviate from the random interactions given by random matrix theory [24–26] and are closely related to common factors existing in the financial market [23]. In addition, Mantegna et al. [21] found that an individual stock is clustered based on the characteristics of the industrial sectors to which it belongs. Until now, most studies have focused on qualitative analyses of this clustering behavior based on industry sector, so we cannot directly compare previous results to the clustering phenomena of a minimal spanning tree (MST) structure created by diverse financial markets (foreign exchange and stock markets).

* Tel.: +82 62 230 6811.
*E-mail address:* phecogjoh@gmail.com.

0378-4371/$ – see front matter 
http://dx.doi.org/10.1016/j.physa.2013.09.031

We propose a novel method for quantifying the grouping behavior of a MST structure and measure the grouping coefficient of each industrial sector for the United States and Korean stock markets. We used the daily closing price of individual stocks traded on the U.S. and Korean stock markets from 1993.01.03 to 2006.12.31. We find that the grouping coefficient of the U.S. stock markets (S&P500) are larger than those of the Korean stock market (KOSPI). In addition, for the Korean stock market we observe higher grouping coefficients in the insurance, health care, and construction sectors, while no grouping characteristics are observed in the remaining sectors. In particular, Korean conglomerates such as Samsung, LG, Hyundai, SK, Hanwha, Kumho, and Dongbu exhibited large grouping coefficients.

In the next section, we describe the data sets and methodology used in this paper. In the 3 sector, we present our results of this study. Finally, conclusions are given in Section 4.

## 2. Data and methodology

In this section, we investigate the degree of grouping in a MST structure estimated from the correlation matrix describing individual industrial sector stocks listed on the Korean (KOSPI(294)) and the U.S. (S&P500(343)) stock markets (from Thomson Datastream). We used daily closing prices spanning approximately 14 years, from 1993.01.03 to 2006.12.31. To observe the detailed grouping characteristics, we estimate the grouping coefficient for sub-periods by repeatedly shifting 5 days, with 250 data points. We derive a return time series, $r(t)$, by using the log-difference, $r(t) \equiv \ln(P(t)) - \ln(P(t-1))$. Here, $P(t)$ is the closing price on day $t$. We construct the stock network by using the MST method proposed by Mantegna [14] from the correlation matrix describing stocks for both stock markets. To estimate the MST, we need the correlation coefficients among stocks, which are calculated as follows:

$$\rho_{ij} = \frac{\langle r_i r_j \rangle - \langle r_i \rangle \langle r_j \rangle}{\sqrt{(\langle r_i^2 \rangle - \langle r_i \rangle^2)(\langle r_j^2 \rangle - \langle r_j \rangle^2)}}, \tag{1}$$

where $\langle\rangle$ indicates the average value of the period during which the correlation coefficients lie within the range of $-1 \leq \rho_{ij} \leq 1$. If $\rho_{ij} = 1$ or $-1$, the two time series are completely correlated or anti-correlated, respectively. If $\rho_{ij} = 0$, the time series are independent. We then estimate the distance matrix for the stocks by using the correlation matrix calculated using Eq. (1), which is given by

$$d_{ij} = \sqrt{2(1 - \rho_{ij})}, \tag{2}$$

where the distance $d_{ij}$ between stocks lies within $0 \leq d_{ij} \leq 2$. The correlation between stocks increases as the distance $d_{ij}$ decreases. We calculate the MST through the distance matrix estimated using Eq. (2).

Over the last few decades, several papers have been devoted the qualitative studies of industrial sector grouping in MST structures. However, studies that quantify the degree of grouping are rare. To measure this grouping behavior, we propose a novel approach for quantifying the grouping coefficients according to the industrial sector. In previous studies of complex network theory [27,28], clustering coefficients were used to estimate the clustering effect in network [27]. Because the MST structure does not permit loops, we cannot use the clustering coefficient; thus, we consider the shortest path length (SPL) between stocks in the MST structure, which is defined as follows:

$$GC_{g_k} = \left| \frac{\langle SP_{g_k}^{external} \rangle - \langle SP_{g_k}^{internal} \rangle}{\langle SP_{g_k}^{external} \rangle + \langle SP_{g_k}^{internal} \rangle} \right|, \tag{3}$$

$$\langle SP_{g_k}^{external} \rangle = \frac{1}{n(N-n)} \sum_{i \in g_k} \sum_{j \notin g_k,\, j \neq i} SP_{ij},$$

$$\langle SP_{g_k}^{internal} \rangle = \frac{1}{n(n-1)} \sum_{i \in g_k} \sum_{j \in g_k,\, j \neq i} SP_{ij},$$

where $n$ is the number of companies belonging to the industrial sector $g_k$, $N$ is the total number of companies and $SP_{ij}$ is the SPL between the $i$-th and $j$-th stock. The average SPL between stocks in an industrial sector with higher grouping coefficients is smaller than those for stocks not belonging to that industrial sector. In the case $\langle SP_{g_k}^{internal} \rangle = \langle SP_{g_k}^{external} \rangle$, $GC = 0$. The parameter $GC$ represents the grouping coefficient of the industrial sector; if $GC = 0$, the correlation structure among individual stocks is uncorrelated, if $GC > 0$, there exists a correlation or anti-correlation among the individual stocks. In other words, individual companies that belong to industrial sectors with higher grouping coefficients should have stock prices that are driven by common factors existing in the industrial sector's market. We expect that our method can be used to elaborate upon the fractal nature of the tree structure as well as upon the classification of diverse economic systems according to the properties of industrial sector. We can observe the degree of grouping of each industrial sector in the MST structure by using Eq. (3).