

Wu, Sen; Tuo, Mengjiao; Xiong, Deying

## Article

# Network structure detection and analysis of Shanghai stock market

Journal of Industrial Engineering and Management (JIEM)

## Provided in Cooperation with:

The School of Industrial, Aerospace and Audiovisual Engineering of Terrassa (ESEIAAT), Universitat Politècnica de Catalunya (UPC)

*Suggested Citation:* Wu, Sen; Tuo, Mengjiao; Xiong, Deying (2015) : Network structure detection and analysis of Shanghai stock market, Journal of Industrial Engineering and Management (JIEM), ISSN 2013-0953, OmniaScience, Barcelona, Vol. 8, Iss. 2, pp. 383-398, <https://doi.org/10.3926/jiem.1314>

This Version is available at:

<https://hdl.handle.net/10419/188688>

## Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

## Terms of use:

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<https://creativecommons.org/licenses/by-nc/3.0/>

# Network Structure Detection and Analysis of Shanghai Stock Market

Sen Wu, Mengjiao Tuo, Deying Xiong

University of Science and Technology Beijing(China)

[wusen@manage.ustb.edu.cn](mailto:wusen@manage.ustb.edu.cn), [lilyanne93@163.com](mailto:lilyanne93@163.com), [xiong.deying@hotmail.com](mailto:xiong.deying@hotmail.com)

Received: November 2014

Accepted: January 2015

## Abstract:

**Purpose:** In order to investigate community structure of the component stocks of SSE (Shanghai Stock Exchange) 180-index, a stock correlation network is built to find the intra-community and inter-community relationship.

**Design/methodology/approach:** The stock correlation network is built taking the vertices as stocks and edges as correlation coefficients of logarithm returns of stock price. It is built as undirected weighted at first. GN algorithm is selected to detect community structure after transferring the network into un-weighted with different thresholds.

**Findings:** The result of the network community structure analysis shows that the stock market has obvious industrial characteristics. Most of the stocks in the same industry or in the same supply chain are assigned to the same community. The correlation of stock prices' fluctuation in the internal community is closer than in different ones. The result of community structure detection also reflects correlations among different industries.

**Originality/value:** Based on the analysis of the community structure in Shanghai stock market, the result reflects some industrial characteristics, which has reference value to relationship among industries or sub-sectors of listed companies.

**Keywords:** complex network, stock market, community structure, GN algorithm

## 1. Introduction

Complex networks are the abstract of complex systems, such as computer network, biology network, social network and transportation network. The complexity of network systems reflects in complexity of structure, complexity of vertices and all the interactions among complex factors.

In recent years, the research of complex network gradually expanded from mathematics, physics and biology to sociology and economics. Rubinov and Sporns (2010) discussed construction of brain networks from connectivity data and described the most commonly used network measures of structural and functional connectivity. Morales, Borondo, Losada and Benito (2014) built complex network to understand the collective reaction to individual actions on Twitter. Gao and Gu (2011) applied Social network analysis to study the strength relation, small groups and coreness to detect the problems in knowledge flow within groups with the help of quantitative description, measure knowledge network. Li, He, Zhuang and Shi (2011) studied the impact of stocks on Chinese inter-bank network stability from the perspective of complex networks.

In these researches, complex networks revealed some obvious statistical characteristics, including small-world properties (Milgram, 1967), scale-free property (Pool & Kochen, 1987) and so on. A network, that has large clustering coefficient and small average distance, has small-world property. Consequently the network is called small-world network. A network with power-law degree distribution is called scale-free network. The degree distribution is the basic topological property of a network, which represents the possibility that a randomly selected node has degree  $k$  or described as the possibility of a node whose degree is  $k$  in network. The degree distribution of a lot of networks can be described by power-law distribution, which is  $P(K) \propto K^{-\gamma}$ . The power-law distribution is also known as the scale-free distribution.

Another important characteristic of complex networks is community structure (Albert, Jeong & Barabasi, 1999). Communities can be defined as groups of nodes such that there is a higher density of edges within groups than between them (Boccaletti, Latora, Moreno, Chavez & Hwang, 2006). A lot of researches' results indicated that many real networks have community structures. Namely, the whole network consists of many communities, the connection between different communities is relatively sparse, but the connection in the same community is relatively dense. Community detection utilizes topological structure of graph to analyze its modularized community structure of complex network, which is meaningful for having an insight into module, function and evolution of the whole network as well as understanding principles of organization, topologies and dynamic characteristics of complex systems.

Detecting community structure in complex networks has become attractive in many researches. Hierarchical partitioning is a common kind of algorithms for community structure

detection. According to edges are added or removed from the network, the algorithms are classified into agglomerative algorithms and divisive algorithms (Scott, 2002).

The basic idea of agglomerative algorithms is to take every node in network as a community, then merge nodes with highest similarity into a community. Repeating this step until all the nodes are in one community or satisfying other termination condition. Common agglomerative algorithms include Newman fast algorithm (Newman, 2004), CNM algorithm (Clauset, Newman & Moore, 2004), clustering algorithm combined with spectral analysis (Donetti & Munoz, 2004) and so on. On the contrary, divisive algorithms put all the nodes in a community at first, and then delete nodes with lowest similarity. Continuing the step above, the entire community will be divided into smaller communities until every node turns to an independent community or satisfying other termination condition. The famous GN algorithm (Girvan & Newman, 2001) is a divisive algorithm.

Since Mantegna (1999) analyzed hierarchical structure of the Standard and Poor's 500 index stocks correlation network, scholars around the world began the researches on the topologies of stock market networks, the conclusions of which have important reference value to reveal trends in financial market.

Kim et al. (2002) established the weighted correlation network of the Standard and Poor's 500 index stocks to study the scale-free characteristics. Boginski, Butenko and Pardalos (2005) studied American stock market and found that the correlation of the stock price is in accordance with scale-free attribute. Through construction of American stock network, Tse, Liu and Lau (2010) found that the variation of stock prices were strongly influenced by a relatively small number of stocks. Caraianni (2012) investigated the properties of the returns of the main emerging stock markets from Europe by means of complex networks. Besides the studies abroad, the application of complex network in China also develops fast these years. Zhuang, Min and Chen (2007) analyzed the topology in Shanghai stock market and found that the stock market networks had the typical statistical characteristics of complex networks. Huang, Zhuang and Yao (2009) used a threshold method to construct China's stock correlation network and then study the network's structural properties and topological stability. Han and Wang (2010) utilized improving CNM algorithm and Li and Chen (2013) utilized multi-gene method, they both found obvious community structure of stock market. The results of these researches are beneficial for economic forecast and financial supervision.

In this paper, through the community detection of correlation network, closely related communities are found in Shanghai stock market. The price fluctuation of stocks in a same community is highly correlated, which illustrates that these stocks are influenced by the same or similar economic factors. The analysis of industrial characteristics of stocks has reference value to relationship among industries or sub-sectors of listed companies.

## 2. Construction of the network

The initial network is built weighted undirected. The vertices represent stocks and the edges represent correlation coefficients of logarithm returns of stock price (Wu, Tuo & Xiong, 2014). Assuming that there are  $N$  stocks in this network, the observation time is  $[t_0, t_0 + \tau]$ . At any time point  $\tau$ , stock  $i$ 's logarithm return is

$$r_i(\tau) = \ln P_i(\tau) - \ln P_i(\tau - \Delta t) \quad (1)$$

$P_i(\tau)$  is the stock price at time point  $\tau$ .  $\Delta t$  is the time difference to calculate the logarithm returns. We choose closing price on Friday every week as the source data to calculate the logarithm returns, hence the time difference is a week. Then we calculate the correlation coefficient  $\rho_{ij}$  between any two stocks  $i$  and  $j$  by the logarithm return series, namely

$$\rho_{ij} = \frac{E[r_i r_j] - E[r_i]E[r_j]}{\sqrt{(E[r_i^2] - E[r_i]^2)(E[r_j^2] - E[r_j]^2)}} \quad (2)$$

$r_i$  is the logarithm return series of stock  $i$ .  $r_j$  is the logarithm return series of stock  $j$ .  $E[*]$  is the mathematical expectation.  $\rho_{ij} \in [-1, 1]$ . The result is an  $n$ -order coefficient matrix  $C$ .

$$C = \begin{cases} c_{ij} = \rho_{ij} & i \neq j \\ c_{ij} = 1 & i = j \end{cases} \quad (3)$$

In formula (3),  $c_{ij}$  represents the element in  $i^{\text{th}}$  row and  $j^{\text{th}}$  column in matrix  $C$ .

As formula (3) shows,  $\rho_{ij} = \rho_{ji}$ . Therefore, matrix  $C$  is symmetric. So we just take the lower triangular matrix to study.

## 3. Community Structure Detection

### 3.1 GN Algorithm

GN algorithm is a kind of community detection algorithm based on divisive idea. It keeps deleting the edge with biggest betweenness. The steps are:

- Calculate all the edges' betweennesses.
- Find the edge with biggest betweenness and delete it from network.
- Repeat step 2, until every node in network becomes an independent community.

Since it is difficult to judge when to stop GN algorithm if the amount of communities is unknown, Newman and Girvan (2004) came up with a concept modularity  $Q$  to solve this problem.

$$Q = \sum_i (e_{ij} - a_i^2) = \text{Tre} - \|e^2\| \quad (4)$$

In formula (4),  $a_i$  represents sum of each row (or column).  $\text{Tre}$  represents sum of every element on the diagonal.  $\|x\|$  represents sum of all the elements in matrix  $x$ .

Typically,  $Q$  is calculated for each split of a network into communities. The values in dendrogram's peak indicate particularly satisfactory splits. The height of a peak is a measure of the strength of the community division.

### 3.2 Community Structure Detection

GN algorithm is applied to un-weighted network, so we determine a threshold to turn the network into un-weighted. Assuming the threshold is  $\delta$ , the conversion process is as follows:

$$\begin{cases} C_{ij}=1 & |C_{ij}| \geq \delta \\ C_{ij}=0 & |C_{ij}| < \delta \end{cases} \quad (5)$$

The key to the conversion is the choice of threshold. Due to the strong correlation is the only interested factor in this study, we set threshold according to coefficient distribution (Newman & Girvan, 2004; Newman, Strogats & Watts, 2001). In the built network, when  $|C_{ij}| \geq \delta$ , vertices  $i$  and  $j$  are linked. Otherwise there's no edge. We regard the influence between two stocks as the same, without considering the direction.

When using the GN algorithm, the threshold  $\delta$  decides the number of edges so it leads to the structure of stock network. And it also determines the result of community structure detection to a large extent. Different stock networks have different structures. The multi-threshold method tries to find the most reasonable community structure with probing the distribution of the network.

## 4. The Empirical Analysis

### 4.1 Data Preparation

The SSE 180-index was listed on July 1st, 2002, which selects the most marketing representative 180 sample stocks from Shanghai A-share Composite. In the choosing process, stocks are deleted in following conditions:

- stocks listed less than a season;
- stocks are suspended;
- stocks that have operational abnormality or have serious loss in recent financial report;

- stocks whose price is in a sharp volatile situation or whose market are obviously manipulated;
- other stocks through identification of experts committee.

The basic data for the empirical analysis is selected from SSE 180-index component stocks from Jan. 9th, 2009 to Apr. 8th, 2011. To ensure integrity, the stocks that listed after Jan. 9th, 2009 are deleted. After deletion, 164 vertices are left in the stock network. For each stock, the weekly closing price means the daily closing price on Friday. If there are any stocks whose closing prices are missing, we replace it with the closing price on last day that week. If all this weeks' data is missing, we replace it with last week's data to ensure the consistency of transaction data. After completion, there are 118 weeks' data, a total of 19532 daily closing prices.

According to Listing Corporation Industry Classification guide, listed companies are classified into 13 categories. 164 stocks to be analyzed in this paper are included in these industries, as Table 1 shows. According to data from Table 1, in SSE 180-index component stocks, the industries having more stocks are manufacturing (C), real estate (J), financial insurance (I) and mining industry (B). These four categories take 72% of all the 164 stocks, others take 28%.

Industry code	Name of industry	Amount of stocks
A	Agriculture, forestry, animal, husbandry and fishing industry	2
B	Mining industry	20
C	Manufacturing industry	52
D	Electricity, gas and water production and supply industry	4
E	Construction industry	4
F	Transportation and warehousing industry	11
G	Information technology industry	8
H	Wholesale and retail trade industry	6
I	Financial insurance industry	22
J	Real estate industry	24
K	Social service industry	2
L	Communication and culture industry	1
M	Comprehensive industry	8
Total		164

Table 1. Listing Corporation Industry Classification Guide

#### 4.2 The Frequency Distribution of Correlation Coefficients

There are 13,530 correlation coefficients between all the 164 stocks. The biggest one is 1, the smallest is -0.2041. The mean value is 0.3834 and the standard deviation is 0.1771. According to correlation coefficients, the frequency distribution histogram and the probability density distribution curve are as Figure 1 shows.

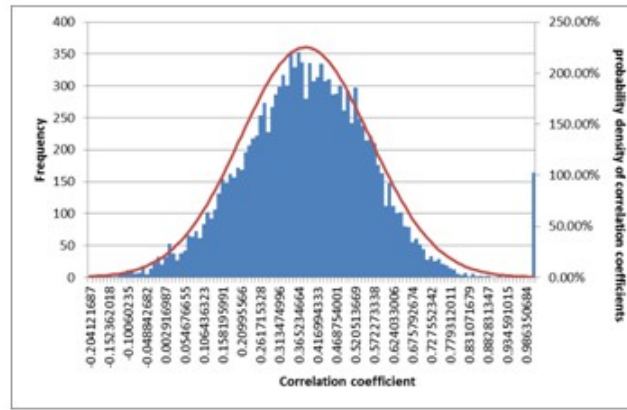


Figure 1. The Frequency Distribution Histogram and The Probability Density Distribution Curve

In Figure 1, X axis represents the value of correlation coefficients. The left side in vertical axis represents the frequency of correlation coefficients. The right side represents the probability density of correlation coefficients. As we can see, the probability density distribution curve approximately obeys the normal distribution.

Using SPSS to verify, the distribution composed of each data point approximates to the linear distribution as Figure 2 shows. Thus, it proves the frequency distribution of correlation coefficients among 164 stocks obeys the normal distribution.

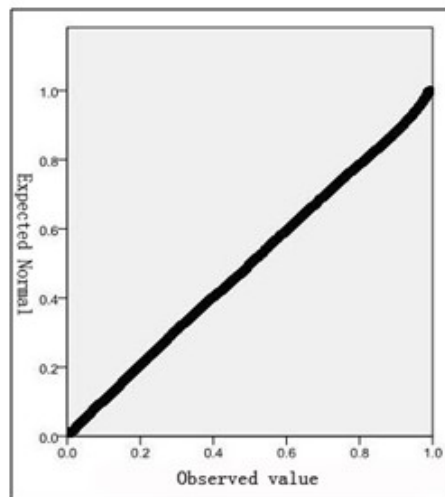


Figure 2. P-P Plot Test of Correlation Coefficients

#### 4.3 The Community Structure Analysis

After removing edges whose correlation coefficient is 1, there are 13,366 edges left. By using different thresholds to transfer network into un-weighted, we try to find the most reasonable community structure. Table 2 shows the amount and percentage retention of edges with



different thresholds. As Table 2 shows, when threshold is bigger, the amount of remaining edges in network is smaller. This means, it reflects closing correlation between vertices better. However, too few remaining edges are not enough to reflect correlation among stocks. So we choose some reasonable thresholds to analyze their community structures.

Thresholds	Amount of retaining edges	Percentage of retaining edges
0.4	6,070	45.41%
0.5	3,167	23.69%
0.6	1,088	8.14%
0.7	253	1.89%

Table 2. Retaining Amount and Percentage of Edges with Different Threshold

When selecting threshold as 0.4, there are 45.41% edges left, as Figure 3 shows. The nodes in the upper left corner in Figure 3 represent outliers in network. Apparently, the community structure is not clear because of too many edges.

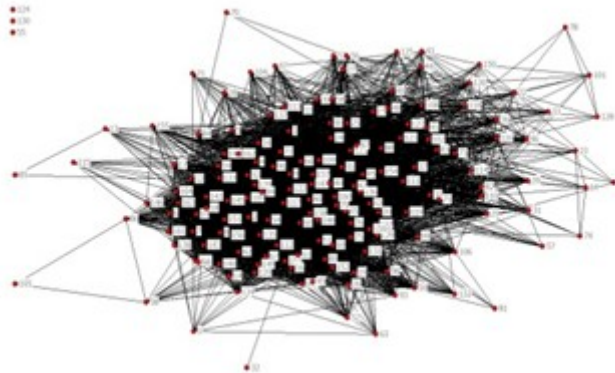


Figure 3. The SSE 180-Index Sample Stock Network Topology with Threshold 0.4



Figure 4. The SSE 180-Index Sample Stock Network Topology with Threshold 0.5

There are three outliers in the network, which means the nodes have no edges connected with other nodes. When threshold gets larger, the amount of outliers gets larger too. After investigation, we found that these three outliers are all controlled by State-owned enterprises.

Selecting threshold as 0.5, there are 23.69% edges left. Compared to network with threshold 0.4, most edges with small correlation coefficients have already been deleted. However, its community structure is still not obvious. Using GN algorithm to analyze this network, the result is similar to network with threshold 0.4. As Figure 4 shows, the community structure is not obvious because of too many retaining edges. The result of detection is not ideal.

Adjusting threshold to 0.6, as Figure 5 shows. The number of edges in network is 8.14% of the original. Only a small number of edges corresponding to larger coefficient are left. Compared to those with thresholds 0.4 and 0.5, this result is better to reflect the structure that is made of stocks with strong correlation. Using GN algorithm to analyze, after sixth partition, the community structure of SSE 180-index network is basically stable. The whole network is divided into 13 small communities and 1 big community. The rest of vertices are outliers. The biggest community includes 100 vertices after 6th partition, as Table 3 shows.

Compared to detection results with threshold 0.4 and 0.5, a situation appeared that most nodes come into the same biggest community. However, because of the lack of vertices in other communities, it is difficult to find their common characteristics.

After adjusting threshold to 0.7, as Figure 6 shows, the vertices with larger coefficients and their edges remain in the network. The community structure begins obvious, compared with previous structures.

Applied GN algorithm to this network, the community structure is more obvious. What's more, there are no outliers around large communities any more. In the dividing process, the community structure becomes stable after fifth partition, as Figure 7 shows.



Figure 5. The SSE 180-Index Sample Stock Network Topology with Threshold 0.6

Community No.	Vertices No. inside the community
1	1, 138, 2, 6, 7, 8, 9, 14, 61, 67, 117, 122, 134, 135, 140, 143, 146, 149, 151, 161, 163, 164, 5, 12, 47, 66, 109, 147, 148, 153, 160, 4, 10, 40, 54, 68, 100, 136, 142, 145, 156, 158, 17, 19, 20, 24, 34, 42, 59, 62, 65, 69, 72, 75, 76, 79, 80, 89, 91, 92, 94, 96, 112, 119, 121, 129, 133, 139, 141, 144, 152, 154, 157, 162, 16, 18, 45, 51, 60, 98, 103, 107, 120, 131, 85, 87, 159, 111, 90, 3, 35, 95, 113, 137, 115, 155, 108, 114, 27, 81
2	11, 30, 49, 15, 41, 25
3	21, 88
4	22, 118
5	23, 28, 64
6	29, 82
7	37, 71, 99
8	38, 105
9	39, 70
10	43, 97, 132
11	48, 83, 102, 127
12	56, 84
13	57, 116
14	123, 125
Total number of vertices	135

Table 3. The Community Structure Distribution by GN Algorithm with Threshold 0.6



Figure 6. The SSE 180-Index Sample Stock Network Topology with Threshold 0.7

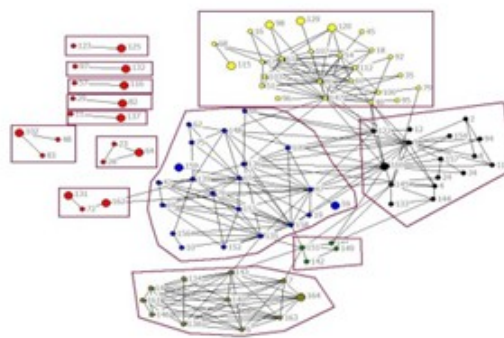


Figure 7. The Community Structure of SSE 180-Index Sample Stock Network with Threshold 0.7

According to the result of community structure detection, there are 13 communities. Nine of them include 2 to 4 vertices, which is too few for community detection. So, we just analyze the first four communities.

As Table 4 shows, the 68% stocks in community 1 belong to real estate development and management. Meanwhile, stock No. 79 belongs to civil engineering construction of construction industry. Stock No. 98 belongs to IT industry, which has close business contact with stock No. 79. Apparently, stock No. 79 and stock No. 98 are both closely related to the real estate industry. Community 1 shows that real estate industry, especially the development and management part and other sectors belong to real estate have strong industrial characteristics. The fluctuations of the stock price belonging to the same industry are highly correlated, but less affected by other industries.

Stock No.	Name of Industry	Industry code
40	Manufacturing industry	C
115	Electricity, gas and water production and supply industry	D
79	Construction industry	E
98	Information technology industry	G
95	Financial insurance industry	I
16	Real estate industry	J
18	Real estate industry	J
51	Real estate industry	J
54	Real estate industry	J
60	Real estate industry	J
61	Real estate industry	J
67	Real estate industry	J
68	Real estate industry	J
80	Real estate industry	J
92	Real estate industry	J
100	Real estate industry	J
103	Real estate industry	J
107	Real estate industry	J
112	Real estate industry	J
120	Real estate industry	J
129	Real estate industry	J
147	Real estate industry	J
35	Social service industry	K
45	Comprehensive industry	M
96	Comprehensive industry	M
<b>Total number of stocks in community 1</b>		<b>25</b>

Table 4. Stocks in Community 1

As Table 5 shows, the 81% stocks in community 2 belong to mining industry and metal rolling processing of manufacturing industry. Mining industry includes metal mining, coal mining, oil and gas exploration. The mining industry owns the most stock nodes in the community. They are easily affected by the price fluctuations inside the industry. The rolling metal processes of manufacturing and mining industry companies are in the same supply chains. Therefore, the correlation of the price fluctuation between these two industries is large, which illustrates these sectors (sub-sector) are vulnerable to the same economic factors. Besides the vertices above, some commercial brokerage, agency industry and Securities & Futures industry are also included. These stocks are closely related to mining and manufacturing industry.

Stock No.	Name of Industry	Industry code
159	Mining industry	B
42	Mining industry	B
69	Mining industry	B
76	Mining industry	B
135	Mining industry	B
152	Mining industry	B
153	Mining industry	B
158	Mining industry	B
10	Mining industry	B
156	Mining industry	B
75	Mining industry	B
139	Mining industry	B
47	Manufacturing industry	C
65	Manufacturing industry	C
148	Manufacturing industry	C
62	Manufacturing industry	C
59	Manufacturing industry	C
19	Wholesale and retail trade industry	H
109	Wholesale and retail trade industry	H
66	Financial insurance industry	I
136	Financial insurance industry	I
<b>Total number of stocks in community 2</b>		<b>21</b>

Table 5. Stocks in Community 2

As Table 6 shows, the 75% stocks in community 3 are from transportation industry and manufacturing industry. Through analyzing these industries deeply, manufacturing industry actually includes ship manufacturing industry, electrical machinery and equipment manufacturing industry, steel rolling industry. Transportation industry includes airport and air transport auxiliary industry, water transport industry, railway transport industry and ocean shipping industry. Manufacture of all kinds of transportation tools in transportation industry is closely related to steel rolling industry. It is quite clear that ship manufacturing industry, water transport industry and ocean shipping industry have a close relationship. The fact that all these stocks are in the same community confirms the internal correlation among their price fluctuations. Moreover, like community 2, community 3 also includes several stocks of securities industry and futures industry, which illustrates the connection between securities and futures industry and transportation and manufacturing industry.

Stock No.	Name of Industry	Industry code
34	Manufacturing industry	C
154	Manufacturing industry	C
2	Manufacturing industry	C
8	Manufacturing industry	C
117	Manufacturing industry	C
24	Manufacturing industry	C
94	Electricity, gas and water production and supply industry	D
145	Construction industry	E
4	Transportation and warehousing industry	F
9	Transportation and warehousing industry	F
133	Transportation and warehousing industry	F
144	Transportation and warehousing industry	F
157	Transportation and warehousing industry	F
160	Transportation and warehousing industry	F
12	Financial insurance industry	I
122	Financial insurance industry	I
<b>Total number of stocks in community 3</b>		<b>16</b>

Table 6. Stocks in Community 3

The industrial characteristics are most obvious in community 4, as Table 7 shows. All the nodes belong to banking. Although bonds and insurance belong to financial insurance industry, they are excluded in this network. Thus, the price fluctuation of banking stocks is not only immune to other types of industries, but also to Securities & Futures industry and insurance industry. Compounding analysis above, stocks belong to bonds, futures and trust sub-sectors all appear in other communities, which means most stocks of finance and insurance industry (except banking) are highly correlated to their specific industry (like real estate industry) instead of industry they belong to (financial insurance industry). In general, except macroeconomic factors, stocks in finance and insurance industry (except banking sub-sector) are easily affected by other industries.

Stock No.	Name of Industry	Industry code
6	Financial insurance industry	I
7	Financial insurance industry	I
14	Financial insurance industry	I
134	Financial insurance industry	I
138	Financial insurance industry	I
140	Financial insurance industry	I
143	Financial insurance industry	I
146	Financial insurance industry	I
161	Financial insurance industry	I
163	Financial insurance industry	I
164	Financial insurance industry	I
<b>Total number of stocks in community 4</b>		<b>11</b>

Table 7. Stocks in Community 4

Above all, in the process of GN algorithm, the result of community detection is affected by thresholds. When threshold gets larger, the amount of retaining vertices gets smaller, the result of community detection is much clearer. Stocks in community 4 are all belong to banking industry, which indicates that banking has the most obvious industrial characteristics. The next is real estate industry, then mining industry, manufacturing industry and transport industry.

## 5. Conclusions

In this paper, GN algorithm is applied to find community structure of Shanghai stock market in order to analyze industrial characteristics inside community. Through changing different thresholds, the result shows that along with the threshold getting larger, community structure of network becomes much more obvious. When threshold turns to 0.7, the network is divided into 13 communities, where the first four get the most nodes.

The results of community structure detection indicate that, first of all, the stocks belonging to the same industry are inclined to be in the same community, such as real estate, manufacturing and mining industry. The most obvious is the banking sector of financial insurance industry. Second, the stocks of the same sub-sectors prefer to be in the same

community. For example, the community including banking sector does not include financial trust or securities. Finally, the industries in the same supply chain tend to be in the same community. For example, stocks of all kinds of metal rolling processing industry belonging to mining and manufacturing industry are divided into the same group.

In short, the price fluctuation of stocks in a same community is highly correlated, which suggests that these stocks are influenced by identical economic factors. Therefore, the stock price is vulnerable to the same industry or affected by the volatility of stock prices on the same production chain; at the same time, besides stock prices are affected by macroeconomic factors, they are more susceptible to effects of the same economic factors.

## Acknowledgments

This work is supported by the National Natural Science Foundation of China under Grant No.71271027, by the Fundamental Research Funds for the Central Universities of China under Grant No. FRF-TP-10-006B and by the Research Fund for the Doctoral Program of Higher Education under Grant No.20120006110037.

## References

- Albert, R., Jeong, H., & Barabasi, A.L. (1999). Diameter of the world-wide web. *Nature*, 401, 130-131. <http://dx.doi.org/10.1038/43601>
- Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., & Hwang, D.U. (2006). Complex networks: Structure and dynamics. *Physics Reports*, 424, 175-308. <http://dx.doi.org/10.1016/j.physrep.2005.10.009>
- Boginski, V., Butenko, S., & Pardalos, P.M. (2005). Statistical analysis of financial networks. *Computational Statistics & Data Analysis*, 48(2), 431-443. <http://dx.doi.org/10.1016/j.csda.2004.02.004>
- Caraiani, P. (2012). Characterizing emerging European stock markets through complex networks: From local properties to self-similar characteristics. *Physica A: Statistical Mechanics and its Applications*, 391(13), 3629–3637.
- Clauset, A., Newman, M.E.J., & Moore, C. (2004). Finding community structure in very large networks. *Physics Review E*, 70, 066111. <http://dx.doi.org/10.1103/PhysRevE.70.066111>
- Donetti, L., & Munoz, M.A. (2004). Detecting Network Communities: a new systematic and efficient algorithm. *Journal of Statistical Mechanics: Theory and Experiment*, 10, P10012.



- Gao, L., & Gu, Q.X. (2011). Social network analysis on knowledge sharing of scientific groups. *Journal of System and Management Sciences*, 1(3), 79-89, available online in: [http://www.aasmr.org/jsms/Vol1/No%203/JSMS\\_Vol1\\_No3\\_9.pdf](http://www.aasmr.org/jsms/Vol1/No%203/JSMS_Vol1_No3_9.pdf)
- Girvan, M., & Newman, M.E.J. (2001). Community Structure in social and biological networks. *Proceedings of The National Academy of Sciences of The United States*, 99, 7821-7826. <http://dx.doi.org/10.1073/pnas.122653799>
- Han, H., Wang, J., & Wang, H. (2010). Improving CNM algorithm to detect community structures of weighted network. *Computer Engineering and Applications*, 46(35), 86-89.
- Huang, W.Q., Zhuang, X.T., & Yao, S. (2009). A network analysis of the Chinese stock market. *Physica A: Statistical Mechanics and its Applications*, 388(14), 2956-2964. <http://dx.doi.org/10.1016/j.physa.2009.03.028>
- Kim, H.J. et al. (2002). Weighted scale-free network in financial correlations. *Journal of the Physical Society of Japan*, 71(9), 2133-2136. <http://dx.doi.org/10.1143/JPSJ.71.2133>
- Li, K.S., & Chen, G.H. (2013). Stock networks community structure division algorithm based on multi-gene families encoding. *Computer Engineering & Science*, 35(7), 87-94.
- Li, S.W., He, J.M., Zhuang, Y.M., & Shi, Y.M. (2011). Stability of Inter-bank Markets Based on Complex Networks. *Journal of Industrial Engineering and Management*, 25(2), 195-199.
- Mantegna, R.N. (1999). Hierarchical Structure in Financial Markets. *The European Physical Journal B*, 11(1), 193-197. <http://dx.doi.org/10.1007/s100510050929>
- Milgram, S. (1967). The small world problem. *Psychology Today*, 2, 60-67, available online in: [http://measure.igpp.ucla.edu/GK12-SEE-LA/Lesson\\_Files\\_09/Tina\\_Wey/TW\\_social\\_networks\\_Milgram\\_1967\\_small\\_world\\_problem.pdf](http://measure.igpp.ucla.edu/GK12-SEE-LA/Lesson_Files_09/Tina_Wey/TW_social_networks_Milgram_1967_small_world_problem.pdf)
- Morales, A.J., Borondo, J., Losada, J.C., & Benito, R.M. (2014). Efficiency of human activity on information spreading on Twitter. *Social Networks*, 39, 1-11. <http://dx.doi.org/10.1016/j.socnet.2014.03.007>
- Newman, M.E.J. (2004). Fast algorithm for detecting community structure in networks. *Physics Review E*, 69, 066133. <http://dx.doi.org/10.1103/PhysRevE.69.066133>
- Newman, M.E.J., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physics Review E*, 69, 026113. <http://dx.doi.org/10.1103/PhysRevE.69.026113>
- Newman, M.E.J., Strogatz, S.H., & Watts, D.J. (2001). Random graphs with arbitrary degree distributions and their applications. *Phys Review E*, 64, 026118. <http://dx.doi.org/10.1103/PhysRevE.64.026118>
- Pool, L., & Kochen, M. (1987). Contacts and influence. *Social Networks*, 1, 1-48.



- Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage*, 52, 1059-1069. <http://dx.doi.org/10.1016/j.neuroimage.2009.10.003>
- Scott, J. (2002). *Social Network Analysis: A Handbook* (2nd Ed.). London: Sage Publications.
- Tse, C.K., Liu, J., & Lau, F. (2010). A network perspective of the stock market. *Journal of Empirical Finance*, 17(4), 659-667. <http://dx.doi.org/10.1016/j.jempfin.2010.04.008>
- Wu, S., Tuo, M.J., & Xiong, D.Y. (2014). Community Structure Detection of Shanghai Stock Market Based on Complex Networks. *International Conference on Logistics, Informatics and Service Science, Carlifonia (USA)*, 23-26(July, 2014),1398-1402.
- Zhuang, X.T., Min, Z.F., Chen, S.Y. (2007). Characteristic Analysis of Complex Network for Shanghai Stock Market. *Journal of Northeastern University: Natural Science*, 28(7), 1053-1056.

Journal of Industrial Engineering and Management, 2015 ([www.jiem.org](http://www.jiem.org))



Article's contents are provided on a Attribution-Non Commercial 3.0 Creative commons license. Readers are allowed to copy, distribute and communicate article's contents, provided the author's and Journal of Industrial Engineering and Management's names are included.

It must not be used for commercial purposes. To see the complete license contents, please visit

<http://creativecommons.org/licenses/by-nc/3.0/>.