Correlation analysis of topology of stock volume of Chinese Shanghai and Shenzhen 300 index

Yiqi Wang [a], Zhihui Yang[b*]

School of Statistics and Applied Mathematics, Anhui University of Finance and Economics, Bengbu 233030, China

[a] email: 916082741@qq.com, [b]email: zhi-hui-yang@163.com

* the corresponding author

Key Words: stock volume; associated network; threshold; small world; centrality

Abstract: In this paper, firstly use the data of 224 component stocks of Shanghai and Shenzhen 300 index from January 2, 2014 to December 31, 2014 a total of 245 days to establish the network. Then use the threshold method of general election to build a number of different associated networks for different thresholds, and verify the nature of small world. Finally, we calculate various indicators of central and together to draw some conclusions.

Introduction

On the macroscopic view, with the end of 2014 to the first half of 2015. The price of the stock was rose rapidly at the beginning, and then after a series of adjustments and several large crash, the stock market has become increasingly received attention. Several small new industries' grow also need to inject funds, in particular, some high-tech industries which capital return period is long and needn't a large amount of funds. But these small businesses are difficult to loans from banks. So it is imminent to solve their survival problem to achieve the goal of entrepreneurship to create jobs and maintain social stability and development. The voice of fully open the National Equities Exchange and Quotations (NEEQ) is increase quickly. At the same time some large funds, brokerages and other financial institutions are also trying to expand new business and to find suitable investment opportunities and internet. In this context, the study of the stock market is very meaningful.

Social network generally exists in our social and economic life, it plays a central role in the transmission of information of job opportunities. And it is essential for many trades of goods and services [1]. With the advent of the internet age, the development of the internet makes data collection and analysis of huge amount of numbers easier than before. Thus, social network analysis is rapidly flourished, and became one of the hottest methods. Now it has been a large number of applications in the life sciences, humanities and other fields.

The establishment of associated network about stock volume of Shanghai and Shenzhen 300 index

Combined with previous literature, we use the threshold method to building the associated network. But there is a problem of threshold method. How much to take the most appropriate threshold in the end? For this question, it did not reach a general consensus and selection criteria. Mostly they choose a similar value, but it tends to make different conclusions because of the different threshold selections. So, in this paper we choose threshold method of general election,

which is selecting several thresholds to build different associated networks (for details, see Figure 1-5), to analysis together and draw conclusions.

Volume between any two stocks can draw the correlation coefficient by calculating. So, all the shares are associated. The discrimination is some strong and some weak. The method used in this paper is to remove those correlation coefficients which are less than a certain value (the threshold), and to retain those which are great than or equal the value [2]. Then, make all maintenances equal to one to build a network without direction and weight. But the network may not connect. It is meaningless to study branches with fewer points. So, we just analysis the network with the biggest branch.

The network is created through, (1) collected some data of daily volume of several stocks over time, and does some certain screening; (2) correlation coefficient was calculated for stock volume between each two; (3) select the appropriate threshold to filtering, then for different thresholds, to obtain different initial network; (4) finally, processing of initial network and reach the largest branch network. The largest branch network is seen to be the associated network of stock volume.
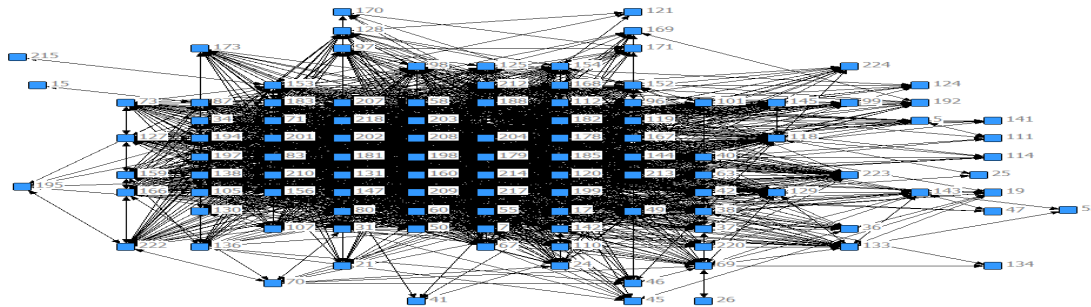

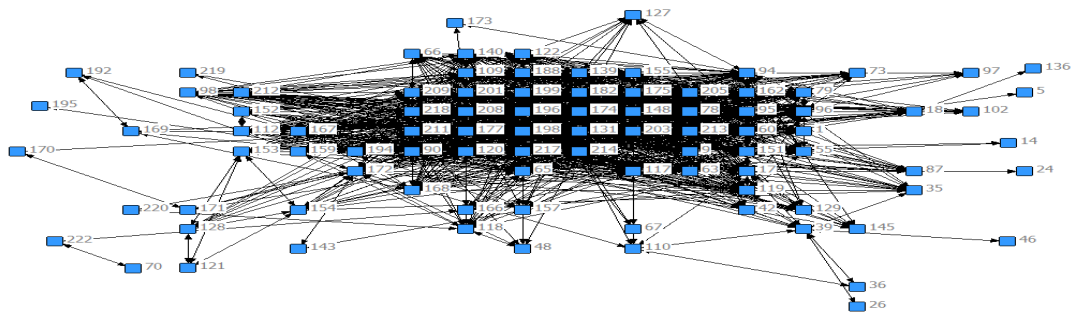Figure 1. a network topological structure of Threshold 0.5

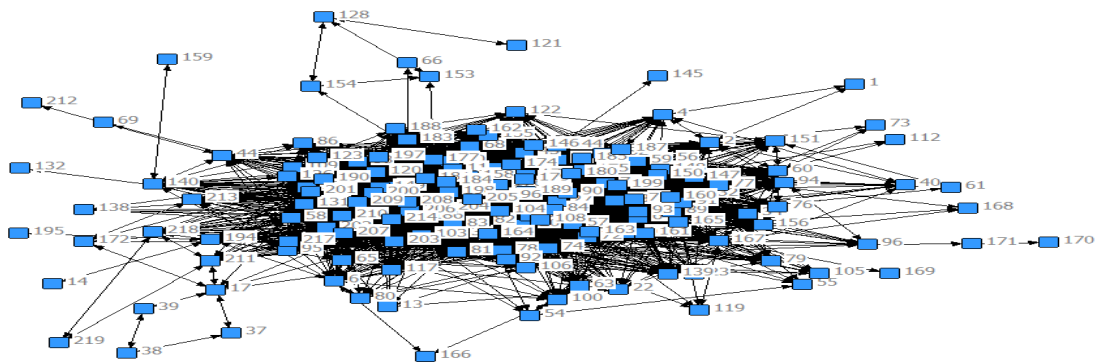Figure 2. a network topological structure of Threshold 0.6

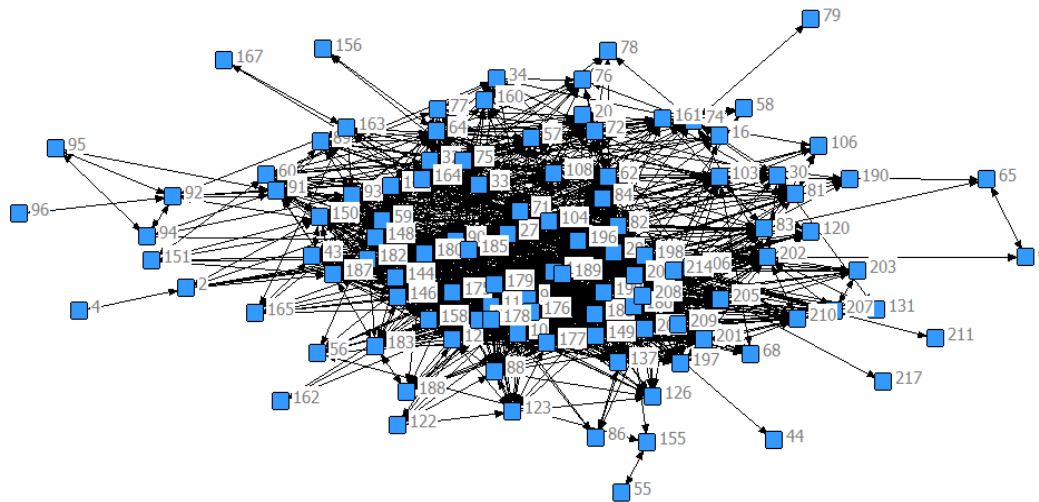Figure 3. a network topological structure of Threshold 0.7

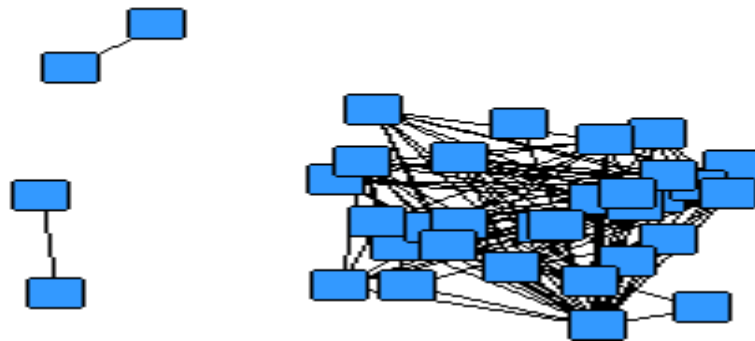Figure 4. a network topological structure of Threshold 0.8

Figure 5. the MDS diagram of Threshold 0.8

Test and verify Small World nature of the associated network

In generally speaking, experts and professors explain the Small World nature when building associated networks. The huge social networks usually reflect this nature, and it is one of the earliest, best-known and most widely used in social networking researching.

The small world network is interested by so many people; the main reason is it has the following features [3]:

(1) Social networks have large-scale. The number of the real world contains a billion magnitudes.

(2) The network is sparse. In other words, for the common life, quantities of everyone can recognize or be able to meet people with respect to the population around the world is very small. And this is consistent with the actual.

(3) Core of the network does not exist, that is to say there will not be any one who at the center of the world.

(4) The whole network is highly clustering. Most of friend networks have overlapping parts.

The step to using UCINET software to measuring small world nature is: Network → Cohesion → Clustering coefficient. We can obtain the associated network. Then following Network → Cohesion → Distance to meet the "Distance Matrix", and reaching the mean length through Tools → Univariate Stats → Matrix [4].

Four clustering coefficient of associated networks were 0.875, 0.868, 0.793, 0.721. They are relatively high, so they are in line with small world properties of associated networks.

Computing of centrality

Power can be seen as an ability of an individual influences others. In other words, it is dependent on the extent of an individual for another individual. The impact of other individuals has different sizes, so we should to quantify [5]. In this way, we can distinguish the manager of a small department and the chief executive of a large area. For different things we should make different tactics, so that we can capture the core element and do things easier.

Several common centralities include degree centrality, closeness centrality and betweenness centrality. Degree centrality is mainly to test the ability of direct connected of a node with other nodes. When changes in its volume, there is an impact with nodes of direct linked. Also these nodes will in turn have a direct impact on it. Closeness centrality is a measure of the speed of impact to other nodes when a node's volume changed. Betweenness centrality measures the frequency of a node on the shortest path of other nodes. And investigate whether the node's "location" is "very good" [6].

The results of three centralities of top ten stocks for each threshold are describing in table 1, table 2 and table 3 respectively.

Table 1. Degree centrality of top ten stocks' number for each threshold

	0.5	0.6	0.7	0.8
1	600166	601607	600000	601088
2	600252	000725	600150	000623
3	601607	600166	000623	600029
4	600118	601118	600010	601668
5	002353	000400	600663	600019
6	600886	000623	600188	600030
7	000651	600663	000725	601117
8	600518	600118	601088	601288
9	000725	600150	600029	600642
10	000623	600111	600111	600999

Table 2. Closeness centrality of top ten stocks' number for each threshold

	0.5	0.6	0.7	0.8
1	600166	601607	600000	601088
2	600252	000725	600150	000623
3	601607	600166	000623	600029
4	600118	601118	600010	601668
5	002353	000400	600663	600019
6	600886	000623	600188	600030
7	000651	600663	000725	601117
8	600518	600118	601088	601288
9	000725	600150	600029	600642
10	000623	600111	600111	600999

Table 3. Betweenness centrality of top ten stocks' number for each threshold

	0.5	0.6	0.7	0.8
1	600166	601607	600000	601088
2	600252	000725	600150	000623
3	601607	600166	000623	600029
4	600118	601118	600010	601668
5	002353	000400	600663	600019
6	600886	000623	600188	600030
7	000651	600663	000725	601117
8	600518	600118	601088	601288
9	000725	600150	600029	600642
10	000623	600111	600111	600999

Summary

Comprehensive analyze with centrality result which obtained by calculating. we can draw conclusions as follows:

(1) JiLinAoDong which stock number is 000623, ZhongGuoShenHua which stock number is 601088 and NanFangKangKong which stock number is 600092 are all in the top ten list for each case of threshold value. BaoGangGuFen which stock number is 600010, ShanXiZhengQuan which stock number is 002500 and ZhongGuoLianTong which stock number is 600050 are all in the top twenty list for each case of threshold value. So, JiLinAoDong can directly affect the largest number of component shares of Shanghai and Shenzhen 300 index, and it also affected greatest impact by other equities. There are also a great number of shares which ZhongGuoShenHua and NanFangKangKong affect directly. They are ranked the second and the third place, and they also suffer a great impact on other equities. BaoGangGuFen, ShanXiZhengQuan and ZhongGuoLianTong are ranked in four to six. They have a large direct influence with other stocks.

(2) JiLinAoDong which stock number is 000623, ZhongGuoShenHua which stock number is 601088 and NanFangHangKong which stock number is 600092 are all in the top ten list for each case of threshold value. BaoGangGuFen which stock number is 600010 and ShanXiZhengQuan which stock number is 002500 are all in the top twenty list for each case of threshold value. So, JiLinAoDong can impact other component stocks of Shanghai and Shenzhen 300 index at the fastest speed, and affected other stocks' volatility is also the largest. The speed of ZhongGuoShenHua which ranked at the second place affect other shares is quickly. And fluctuations by other stocks are also huge. BaoGangGuFen, ShanXiZhengQuan and NanFangHangKong are ranked in three to five. They are affected by fluctuations of other stocks as well.

(3) JiLinAoDong which stock number is 000623 is all in the top ten list for each case of threshold value. JiangXiTongYe which stock number is 600362 is all in the top twenty list for each case of threshold value. ZhongGuoShenHua which stock number is 601088 and NanFangHangKong which stock number is 600092 are also more forward for each case of threshold value. So JiLinAoDong is on the best location of component stocks of Shanghai and Shenzhen 300 index. And the dependence is on the highest level. The location of JiangXiTongYe is also very good, and the dependence is ranked second. ZhongGuoShenHua and NanFangHangKong are in the third and fourth place.

On the whole, we must follow the actual situation to choose the most appropriate centrality. Then we can make choice with the result of one kind of centrality or together with all three centralities.

References

[1] Matthew O. Jackson, Social and Economic Networks, Princeton University Press, Princeton, 2008.

[2] Laijun Zhang, Zhihui Yang, Feifei Lu, An empirical analysis of stock index association based on complex network theory, J. Chinese Journal of Management Science. 12 (2014) 85-92.

[3] Watts, D.J, Small world: The dynamics of network between order and randomness, Princeton University Press, Princeton, 1999.

[4] Jun Liu, Overall network analysis handout, Shanghai People's Press, Shanghai, 2009.

[5] Yafang Li, Caiyan Jia, Jian Yu, A parameter-free community detection method based on centrality and dispersion of nodes in complex networks, Physica A: Statistical Mechanics and its Applications. 15(2015) 321-334.

[6] Alireza Abbasi, Jorn Altmann, Liaquat Hossain, Identifying the effects of co-authorship networks on the performance of scholars: A correlation and regression analysis of performance measures an social network analysis measures, Journal of Informetrics. 5(2011) 594-607.