The Characteristics Analysis of the Stock Network Based on Weighted Relative Values :an example of information service industry

Hu Ping
Dept. of Management Science
Xi'an Jiaotong University
Xi'an 710049, China
E-mail:helenhu@xjtu.edu.cn

Qi Jie, Wang Bingqing
Dept. of Management Science
Xi'an Jiaotong University
Xi'an 710049, China
E-mail:hello_xiaoqi@stu.xjtu.edu.cn

Abstract—**Based on the data of the information services stock network in Shanghai A-share and Shenzhen A-share, this paper established a matrix according to the correlation of the stock price fluctuation. Because China's stock market is "starting late and immature" compared with global stock markets, the market value of shares does not truly reflect the intrinsic value of the company. Pointing to the problem, it is better to use financial index system to calculate the weight instead of using stock price directly. Then the matrix is dealt with the threshold method and establishes the stock network. By analyzing the evolving characteristics of the stock network, this paper got the following conclusions: (1) the Chinese stock networks has the small-world feature, which means the relationship between the listed companies in information service is very close;(2) the stock network follows the power-law distribution, and the power-law index is 1.098 and 1.120 respectively, which indicates that they are scale-free networks and there are some stocks which have large influence on others in each network;(3) It is feasible to reflect the true value of the company by using financial index system to calculate the relatively value of company as weights.**

Keywords-complex network; stock price; small-world features; power-law features

I. INTRODUCTION

As complex networks theory mature, they have been widely applied in range of the natural world, engineering, society, such as the World Wide Web [1], the Internet [2], protein networks [3], power grid [4], the science Citation network [5], scientific collaboration network [6] and so on. It can almost describe and abstract any complex system topology, and simulate the mechanisms of evolution of the network system, in order to obtain a more profound and practical results. Since securities network has wealth real-time data and unpredictable changes, and the financial market is greatly attractive, scholars have invested a lot of effort to develop the laws behind security market. Now, complex network theory has been done to global important stocks, such as the S & P 500 [7], the Hang Seng [8], the Italian stock market [9], the Shanghai and Shenzhen stock markets [10] [11], which have been all established stock networks. Much work finds that different stock networks have many similar characteristics, such as scale-free distribution and small-world characteristics.

Scholars who have researched in stocks networks often aim at the entire stock market, such as the S & P 500, the Shanghai stock market, which you can grasp the whole topology of the stock market network in a macro view. However, each stock market consists of diverse forums, for example, the Shanghai A-shares are classified into real estate, information technology and other 11 forums by China's Securities Regulatory Commission. It is well worth exploring what kind is topology of the network built by stocks belong to the same industry, and what the differences are between its network characteristics and the entire network built by all stocks in the market. Discussion of the domestic stock network model stays in undirected and weighted stock network model, which not well correspond to the real stock market. This paper is focused on information services industry in the Shanghai and Shenzhen A-share market, build directed and weighted stock network based on calculating the correlation coefficient between stocks, setting threshold, and combining the financial index system. Then we analyze the stock network, in order to be able to know the dynamic characteristics of the networks evolution.

II. DATA SOURCE AND DATA PROCESSING

The data that we use are all second data from CSMAR solution（www.gtarsc.com）.We chose stocks in information technology and some stocks in culture communication plate in Shanghai A-share and Shenzhen A- shares, according to the classification of China's Securities Regulatory Commission. After selection, we get stocks in information service industry, including 58 stocks in Shanghai A-share and 119 stocks in Shenzhen A- shares. The study period started from 1.1.2011 to 12.31.2011.

Due to poor management and restructuring, listed company will face delisting. Hence, those stocks are removed first. In addition, the new listed companies under a year are removed, because of lacking dada.

In original closing price, if the stock is suspended due to holidays or meeting special circumstances, the closing price is zero. In order to avoid this vacancies, we use the before data to fill in.

III. CONSTRUCTION OF THE STOCK NETWORK.

This paper builds stock network using information services industry shares in Shanghai and Shenzhen A-share market as nodes, which is divided into two steps: ① establish the connection between the nodes. Combining

with logarithmic yields of the closing price of each stock, we calculate correlation coefficient between stocks. With introduction of a threshold indicator, we establish the connection between the nodes when the correlation coefficient is greater than the threshold. Otherwise, it is not to establish a connection. ② give node direction and the weight. Since the influence between nodes of the stock network has direction and weight, it is need to weight links between nodes. Generally, the effect of the stock node A to node B is different from node B to node A, which means that the interaction between stock nodes is not symmetrical, Introduce node influence by calculating the relative value of the company and related coefficient. ③ calculate node strength. Through the above steps, we can obtain the information service industry stock network nodes, edges, weights, threshold, and strength and other network characteristic index, which you can build a weighted complex network system.

A. link establishment of stock network

We build stock network treating information services industry shares as nodes, and the connection between nodes relate to the correlation matrix of the stock price volatility. So first calculate the stock price volatility correlation.

Definite $Y_i(t)$ as the closing price of stock i at the moment t, then the logarithm yields of the stock can be defined as:

$$S_i(t) = \ln Y_i(t+\Delta t) - \ln Y_i(t) \qquad (1)$$

Where Δt represents the time interval. As the analysis in this article is the closing price, Δt equals to 1.Due to the logarithm yield of the closing price highly depending on the time, its mean value may be interfere with external factors in market, such as bank interest rates, inflation, exchange rates, which are disturbance over time changing. In order to extract the essential characteristics of the stock relevance produced by the stock price volatility, this paper based Nitin Arora [12] and others' research, change(1) as follows transform:

$$G_i(t) = S_i(t) - \frac{1}{N}\sum_i S_i(t) \qquad (2)$$

where N represents the total number of nodes in the network. Thus, at the moment t, the price of stock i is relative value to the mean value of logarithmic yield of the overall stock market, which objectively reflect the impact leading by stock price volatility on the stock.

Using $G_i(t)$ to define the correlation coefficient of the stock nodes, we get the definition of the Pearson correlation coefficient in this paper:

$$w_{ij} = \frac{<G_iG_j> - <G_i><G_j>}{\sqrt{(<G_i^2> - <G_i>^2)(<G_j^2> - <G_j>^2)}} \qquad (3)$$

Hence, $w_{ij} \in [-1,1]$ represents the correlation coefficient based on logarithmic yield of the stock price. If w_{ij}=0, the node i is not associated with node j; if w_{ij}>0, node i is positive correlation with node j, and the greater absolute value is, the stronger relevance is.

According to the research of Kim et al. [13] on the S & P 500 stocks of the stock correlation network, they find stock networks have similar characteristics. With Table 1 and Figure 1, it can be seen that the correlation coefficient of information service industry in the Shanghai A-share and Shenzhen A-share have similar probability distribution, and the mean value is very close to 0.

Determination of the node connections needs threshold. Based on the evolution of the financial markets, Qiu et al. compare the dynamic behavior of the network of Sino-US stocks with Chinese stock. The results illustrate that the correlation coefficient matrix addressed by threshold can effectively suppress large fluctuations caused by correlation between stocks, providing a more stable topology for us.

Based on the previous studies, we set a threshold θ for the correlation coefficient of the information services industry stock. If $|w_{ij}| \geq \theta$, establish connection; if $|w_{ij}| \leq \theta$, don't connect. So after threshold processing, we redefine stock correlation coefficient matrix.

TABLE 1 CORRELATION COEFFICIENT STATISTICAL DESCRIPTION

Market	MAX	MIN	Mean value	Standard deviation
Shanghai A-shares	0.5258	-0.277596	0.00247	0.183678
Shenzhen A-shares	0.7862	-0.314	0.003	0.144589

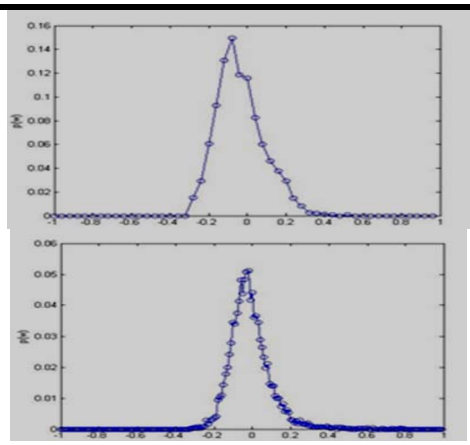


Figure 1. correlation probability distribution of information services industry stock in the Shanghai and Shenzhen A-shares stock market

B. Determination of weight of directed stock network

When determine the connection between the nodes, let us weight the links. We define that e_{ij} represents the link between node i and node j, which means the impact of stock I on stock j based on the stock price volatility.

When it comes to the weight of edges, scholars have done much research. Kim et al [13] directly think that e_{ij}=w_{ij}, which treats the correlation coefficient of stock as the weight of the edge, and use the data of S & P 500 listed

companies during the period from 1994 to1997 to analyze. The results show that all the nodes having greater strength are from certain plate (like semiconductor plate). However, the large companies, such as GE, P & G, have smaller strength. Nitin et al [12] propose improvement that he introduces the company's market value of shares (stock price multiplies by the number of shares) to make a more accurate calculation of the strength influence of the stock node. Assume A and B, two listed companies, if Company A has influence on bigger companies while Company B impacts small companies, the company A and B should not have the same impact. After adjustments the calculation of the strength is closer to the stock in the reality, such as Exxon Mobil (XOM), Procter & Gamble (PG), and Johnson & Johnson (JNJ) ranking ahead. The market value changes with the stock price volatility. In the strong and mature capital market, the value of listed companies can be reflected by its market value of shares such as S&P500. However, China's Shanghai and Shenzhen stock markets are basically not yet mature and just a weakly efficient capital market, where the stock market does not fully reflect the fundamentals, and the company's value often deviates its intrinsic value deviates from. According to Liu Jianrong [14], Chen Jie, Zhou Xin, calculate the company's intrinsic value measured by constructing the system of financial indicators of listed companies.

According to the characteristics of the Chinese stock market, drawing on previous research on building a market value of financial indicators system, the paper selects listed company solvency, operating capacity, profitability and development capacity as the first level indicators, covering 13 second level financial indicators (Table 2), and use TOPSIS method, objective weighting evaluation, whose process can be found in the literature [16] [17] to evaluate relative intrinsic value of the listed companies.

We can calculate the relative value of each stock (between 0 and 1) based on the index system. For analyzing and calculating in the back, time relative value 100 as the relative intrinsic value M of various listed companies. Many companies rankings behind are always warned, ST or companies on the verge of delisting, which illustrates poor management or other reasons, relative value of the company decreases, due to less competition of the stock complex network node, lower fitness can't get more connections, finally those companies will face the risk of elimination, and remove from the network.

On the foundation of Nitin research, combining with financial index system to calculate the value of M, this paper introduces the node strength and propose a more suitable method of edge weight for the special case, the Chinese stock market network. Normally, the interaction between nodes has direction instead symmetrical. Since the effect of the stock node i is not the same with that of node j, strength of node is defined as follows:

$$Infl(i \to j) = w_{i,j} \times \frac{M_i}{M_i + M_j} \quad (5)$$

where Mi representatives relative value of listed companies i. Similarly, we also define the influence of stock node j to node i. From equation (5), it can be seen that the influence of the stock node i to node j actually uses the relative value of the listed companies to adjust w_{ij}. Obviously, the greater the relative value M is, the greater its influence is.

C. Calculation of strength of stock network

In weighted network, an important concept is strength. Since side is given a different weight, degrees can not completely reflect the connectivity of a node. The definition of the strength distribution is similar with the degree distribution. Define strength of node i in the entire network as:

$$s_i = \sum_j Infl(i \to j) \times M_j = \sum_j w_{i,j} \times \frac{M_i}{M_i + M_j} \times M_j \quad (6)$$

indicates that the impact of the node in the entire network is related to the relative value of its neighbors stock node, namely, the greater the relative value of neighbor nodes, the greater strength nodes play a role in.

Hence, we define node, edge weights, threshold, as well as strength and other indicators of network characteristics in network of information service industry to build a complex directed and weighted network.

IV. ANALYZING CHARACTERISTICS OF STOCK NETWORK

A. Small –world characteristic

The small-world characteristic generally describes by two network characteristics, the average path length and clustering coefficient. In this paper, specific formula of two network characteristics refers to [4] [20].

Based on constructed stock network, in the case of setting different threshold value θ, calculate the stock average shortest path in the network and clustering coefficient respectively. The result of index characteristic of the Shanghai and Shenzhen A-share are shown in Table 3.

Comparing with the characteristics of the ordinary model network in literature [6] [11] [19]. In Table 3, average path length is smaller and clustering coefficient is greater which indicates that the network of information service industry stock market of Shanghai A-share and Shenzhen A-shares have typical small-world characteristics. Smaller average shortest path length illustrates any two stocks can be associated through a complex network. Since the stock network constructed in this paper is based on the stock price volatility, it states that the individual stock price volatility in the information services industry stock network easily impact others through the network; Meanwhile, larger clustering coefficient describes better gathering of this stock network which means connection completeness of the local With the increasing of the threshold θ, the average shortest path is increasing, while the clustering coefficient is continuously decreasing which proves that

threshold θ of the stock limits small world network and characteristics of the network will decrease as the threshold gradually increasing.

Compare to the complex network characteristics of the two markets, with the same threshold θ, it can see the Shanghai stock market has a smaller average path length and greater clustering coefficient, which suggests that the volatility of the stock price of information service industry have greater impact on the Shanghai Stock. Moreover, "clusters" the Shanghai Stock internally connected more closely and small world characteristic is even more significant.

The emergence of this result corresponds to the reality. The reason is the information services industry in this emerging industry contact closer than traditional industry companies which especially the leading enterprises in the industry played a role indeed affect the whole industry and small world characteristics is more obvious. Information services industry in Shanghai A shares are listed earlier, while information service companies in the Shenzhen A shares come from GEM, with shorter listed time, so connection between enterprises in Shenzhen compared to Shanghai A-shares is weaker.

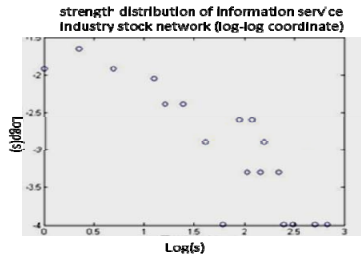
B. Power-law characteristic

Figure 1 describe stock correlation coefficient w_{ij} distributes in (-0.2, 0.2). According to method in reference [11] approach and conclusions of this section set the threshold θ = 0.2 to analyze power-law characteristics of information services industry shares. When the correlation coefficient matrix w addressed by threshold, combining with equation (5), and the formula about the relative intrinsic value of all listed companies get the matrix of node mutual influence the stock in the weighted stock network, and finally substituted into equation (6) to obtain strength s of nodes in the network of the information service industry stock. Strength of Shanghai and Shenzhen A-share information services industry listed companies is shown in Table 4.

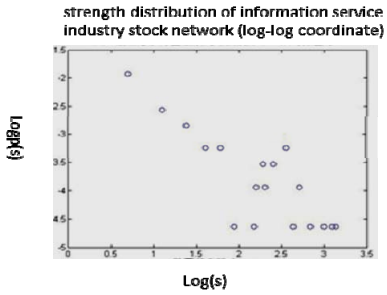
From Table 4, companies with the greater strength in information services industry in Shanghai stock market are Neusoft Group, the wave of software, China Unicom, which listing a long time and "leading" enterprises, while companies with lower strength are warned, and treated as ST even face delisting. Therefore, ranking top companies have much greater impact

Similarly, Shenzhen stock market has similar performance, which gap between the maximum and minimum strength value is big. In Shenzhen A-share many stocks node set up late far behind mature large companies, which also shows the greater volatility of the information services market in Shenzhen.

According to strength, make the picture of strength distribution of information services industry shares in Shanghai and Shenzhen A-shares in Figure 3. In the double logarithmic coordinates, the relation of strength and probability is approximate linear distribution which shows good power-law characteristics. By using the method of least squares, we can get power-law index γ = 1.098 and γ = 1.120 respectively.



(a) Strength distribution of network of information service industry in Shanghai A-share stock



(b) Strength distribution of network of information service industry in Shenzhen A-share stock

Figure 2. strength distribution of Shanghai and Shenzhen A-share stock information services industry network

TABLE 2 RELATIVE INTERNAL VALUE OF LISTED COMPANIES' FINANCIAL INDEX SYSTEM

solvency	current ratio	quick assets ratio	currency ratio	
operating capacity	total assets turn over	current assets turnover	account receivable turnover rate	
profitability	operating profit ratio	profit rate of asset	rate of return on net assets	rate of return on assets
development capacity	operating income growth rate	the growth rate of net profit	basic earnings per share growth rate	

TABLE 3 SHANGHAI AND SHENZHEN A-SHARES OF STOCK NETWORK INFORMATION SERVICE INDUSTRY.

threshold θ	**The average shortest path length**	
	Shanghai A shares	***Shenzhen A shares***
0.05	1.3033	1.416
0.06	1.3613	1.4906
0.08	1.4675	1.6126
0.1	1.5828	1.731
0.12	1.6925	1.8943
0.14	1.8344	2.1453

0.15	1.9182	2.2448
0.16	2.0475	2.2606
0.18	2.3068	2.1888
0.2	2.457	2.1957

Continue to Table 3

threshold θ	Clustering coefficient	
	Shanghai A shares	*Shenzhen A shares*
0.05	0.735	0.6213
0.06	0.6902	0.5655
0.08	0.6205	0.4869
0.1	0.5394	0.4329
0.12	0.4933	0.3842
0.14	0.4675	0.3352
0.15	0.421	0.3036
0.16	0.3982	0.2758
0.18	0.3846	0.2412
0.2	0.3406	0.248

TABLE 4 A STRENGTH OF INFORMATION SERVICE COMPANY IN SHANGHAI AND SHENZHEN TWO STOCKS

(a)Shanghai		
Rink	***Stock***	***Strength***
1	Dong Ruan Group	180.5291
2	Wave Software	178.0581
3	China Unicom	170.6871
4	Fang Zheng Technology	167.734
5	Long River Communication	157.7359
……	……	……
49	ST Hua Guan	42.30953
50	ST Hu Ke	41.8296
51	Xin Zheng Stock	40.04089
52	Guo Dian Nan Rui	32.70911
53	ST Vanguard	32.52746
(b)Shenzhen		
Rink	***Stock***	***Strength***
1	Wu Han Fan Gu	410.6424
2	Mei Tai Ruo	409.0651
3	ZTE Communication	406.5838
4	Purple light Share	401.4437
5	Digital Video	391.5502
……	……	……
98	ST Tai Guang	57.76148
99	Super Graph Software	53.54683
100	New Beiyang	51.46408
101	Zhuo Yi Technology	50.58385
102	*ST Golden Horse	48.35763

Note: translation of stock name accords to the name of company.

V. CONCLUSION

This paper builds a stock based on the stock price volatility and threshold adopting financial index system including solvency, operating capacity, profitability and development capacity to adjust the weak and immature Chinese stock market, and adjust the weight deviation by direct calculating stock market value. What's more, based on the real situation of the stock market, we find strength of each stock node i in the network is related to the strength of its neighbors stock node. Combining with financial indicators system designs a weight calculating method and build a directed and weighted stock network information services. The study conclusions were as follows:

(1) Analysis of small-world characteristics: the network has a smaller average shortest path length and large clustering coefficient, which network of information service industry in Shanghai and Shenzhen A-share market has typical characteristics of small-world networks; the threshold θ limits small world characteristics and the connectivity of the network will decrease as the threshold gradually increasing; under the same threshold value , small world properties of the Shanghai A shares are more obvious.

(2) Analysis of node strength: the strength distribution of the two stock networks follow power law distribution, and power-law indexes are $\gamma = 1.098$ and $\gamma = 1.120$. comparing with the Results of Nitin et al [12] on the U.S. S & P 500 stock which power law index is 2.2, and the model Kim et al [13] calculates that the power-law index is 1.8 our power-law index is smaller. This indicates that our network is smaller scale-free networks with smaller power law index. Fixing threshold specified there are more hub nodes in information service industry in Shanghai and Shenzhen A-shares. At the same time it indicates that in the Information Services industry, some of these influential stocks impact others, even the entire stock network. in the case of setting the threshold value as 0.2, heading companies are industry ranking, ranking mostly been warned , ST handle facing delisting the company. In addition, the impact strength "strength" in the information services industry in the stock market, network distinction is very clear, deep gap is even greater.

Directed and weighted stock network extract from the actual situation and apply to the actual which are available for subsequent analysis of formation, evolution mechanism of the information services industry stock market network as well as suggestions and recommendations for improving existed problems.

R.B.G. THANKS

This work was supported in part by the National Natural Social Science Foundation of China under grants

10BGL032, and the electronic and information department of the Chinese academy of Engineering in 2013(2013-02-XY-003）

REFERENCE

[1] Barabasi AL, Albert R, et al. Power-law distribution of the world wide web[J]. Science, 2000, 287: 2115a.

[2] Faloutsos M, Faloutsos P, et al. On power-law relationships of the internet topology[J]. ACM SIGCOMM computer communication review, 1999,29(4): 251-262.

[3] Montoya JM, Sole RV. Small world patterns in food webs[J]. Journal of Theoretical Biology, 2002,214: 405-412.

[4] Albert R, Barabasi AL. Statistical mechanics of complex networks[J].Reviews of Modern Physics, 2002, 74(1):47-97.

[5] Redner S. How popular is your paper? An empirical study of the citation distribution[J]. European Physical Journal B, 1998,4:131-134.

[6] Newman MEJ. The structure and fuction of complex networks[J]. SIAM Rev.,2003, 45(2): 167-256.

[7] Boginski, V., S. Butenko, et al. Statistical analysis of financial networks[J]. Computational Statistics & Data Analysis,2005, 48(2): 431-443.

[8] Wang, B. H. and P. M. Hui . The distribution and scaling of fluctuations for Hang Seng index in Hong Kongstock market[J]. The European Physical Journal B-Condensed Matter and Complex Systems, 2001, 20(4): 573-579.

[9] Caldarelli, G., S. Battiston, et al.Emergence of complexity in financial networks[J].Complex Networks: (2004),399-423.

[10] Xintan Zhuang, Zhifeng Min, the analysis of complex network of Shanghai stock market[J]. Journal of northeastern university (natural science edition. 2007,33(07): 1053-1056.

[11] Weiwei Lu, Zhengchun Lin, the analysis of Shanghai and Shenzhen A-share based on complex network theory [J]. Science, technology and engineering.2009,(5): 2859-2863.

[12] NitinArora, Narayanan B, Pau l S. Financial influences and scale-freenetworks.Lecture Notes in Computer Science[J], 2006: 16-23.

[13] Kim H J, Lee Y, Kahng B,et al. Weighted scale- free network in financial correlations [J].JphysSocJpn, 2002, 71(9): 2133- 2136.

[14] Jianrong Liu, Heping Pan,the analysis of investment value based on the financial index system of China's listed company [J]. Journal of management scientists (academic version) .2010, (05): 16-25.

[15] Xin Zhou, The relevance of accounting information and stock price of listed companies to empirical analysis based on the real estate industry research [J]. Economic Forum. 2009(12): 105-107.

[16] Jingbo Shao,Bozhou Li, Xiaoli Zhou, comparison the model of the potential of science and technology of China and Russia based on weighted principal component TPOSIS value function [J]. Journal of China soft science.2008（9）：132-139..

[17] Xin Xia,Huiping Zhu, Using principal component analysis combined with weighted TOPSIS method for hospital medical work quality comprehensive evaluation [J]. China hospital statistics. 2008，15（4）:343-346.

[18] Watts D, Strogatz S. Collective dynamics of small-world networks[J]. Nature, 1998, 393: 440-442.

[19] Erdos and Renyi, .On the evolution of random graPhs[J] Publ.Math. Inst. Hung. Acad. Sci. 5(1960) ,17.

[20] Newman M. Mixing patterns in networks[J]. Physical Review E, 2003, 67(2):255-261