Physica A 523 (2019) 622–630

Contents lists available at ScienceDirect

Physica A

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

An analysis of the clustering effect of a jump risk complex network in the Chinese stock market

Sunyang Hu [a], Zongyuan Gu [b], Yifeng Wang [c], Xiaolei Zhang [d,e,*]

[a] School of Business, Nankai University, Tianjin 300071, China
[b] Faculty of Mathematics, East China Normal University, Shanghai 200241, China
[c] Shenwan Hongyuan Securities CO., LTD, Shanghai 200031, China
[d] Yunnan Normal University, Kunming 650000, China
[e] Yunnan Association for Promotion of Trans-Asian Financial Cooperation and Development, Kunming 650000, China

HIGHLIGHTS

- Jump risk is an important part of the volatility. Jump contributes more to volatility during the financial crisis.
- The jump of manufacturing industry stocks plays an important role in the entire tree map.
- The jump risk in the financial industry has the greatest significance.

ARTICLE INFO

Article history:
Received 25 June 2018
Received in revised form 20 September 2018
Available online 2 February 2019

Keywords:
Realized volatility
Realized jump
Complex network
Modular Q function
Fast unfolding algorithm

ABSTRACT

With the development of Chinese financial market, the correlation between stock price jump risks cannot be ignored. This study uses the complex network method to analyze the clustering effect of stock price jumps. Taking a sample of stocks from the CSI 300 Index, the realized jumps are extracted from the 5-minute high frequency data using the MinRV method. The authors use the Minimum Spanning Tree algorithm to construct a complex network of stock price jumps. It is found that there is a clear correlation among stocks in the entire jump network. The jump in manufacturing industry stocks plays the most important role in the network. The Modular Q function and the Fast Unfolding algorithm are used to divide the entire complex network and study the differences in jump correlation between different communities. The result shows that the correlation among the financial industry stocks is stronger, and a large fluctuation in the price of one financial stock can cause the price of another financial stock to fluctuate significantly.

1. Introduction

With the development of China's stock market, the correlation among stocks has increased, and the jump risk in stock prices plays an important role in enhancing stock correlation. Therefore, the study of the correlation of jump risk among stocks has practical significance for further research on the coordination of China's stock market. This paper uses the complex network to study the correlation among the realized jumps of Chinese stocks. The minimum spanning tree network formed by the stock price jump volatility was divided into groups to study the correlation among the inner and outer communities.

* Corresponding author at: Yunnan Normal University, Kunming 650000, China.
E-mail address: financialmath@163.com (X. Zhang).

https://doi.org/10.1016/j.physa.2019.01.114
0378-4371/

With the development of computer science and big data, the extraction of volatility from high-frequency data has become a mainstream research method. Andersen and Bollerslev [1] used five-minute high-frequency data in the foreign exchange market to apply the quadratic power variation method in the study of volatility jumps and proposed the concept of realized variation. Barndorff-Nielsen and Shephard [2] proposed a non-parametric estimation method of the realized double-power variation (RBV) and the realized multi-variable variation (MPV) of neighboring reception-sequence, which improved the estimation accuracy of the attainable realized volatility. Wang et al. [3] proposed and studied detection statistics of intra-day jumping behavior ABD statistics. Andersen et al. [4,5] proposed the minimum RV based on their study of the theory of double-power variation, and empirically verified that it has better recognition capacity. Zeng and Zuo [6] used the MCMC algorithm to establish a jump-dependent stochastic volatility and described jumps of overflow condition using the conditional overflow probability, skip overflow frequency, and jump overflow strength for jump terms. Xu et al. [7] evaluated the impact of jump components of different jump tests on the prediction of volatility by using BRV and three other jump test identification methods. Song et al. [8] researched the Shanghai composite index and estimated the size and frequency of consecutive component jumps of volatility based on C_TZ statistics.

With the development of computer-based informatization, the stock network has become more widely used in recent years, and stock network theory has become more mature. Some scholars used a minimum spanning tree algorithm (MST) to build a stock network. They then conducted a cluster analysis on it and found that the minimum spanning tree algorithm can reveal the hierarchical structure of the network [9,10]. Some scholars also proposed a method to divide communities based on Laplacian graph feature values [11]. Some scholars regarded each node as a stock and built the stock network by using the relationships between stock price returns as the edges between them in the model [12,13]. Based on previous research, scholars also developed the stock network algorithm, and the GN algorithm [14], Fast Newman algorithm [15], Fast Unfolding algorithm [16] were proposed.

The research has three advantages over the existing research. (1) The object of the study is more detailed. The existing research is mainly to study the correlation between volatility. In this paper, the jump is extracted from the volatility by using the realized jump method. The correlation between jump risks is studied. (2) The research method is more reasonable. In this paper, we extract the volatility and jump by using the realized method. The traditional GARCH model needs to set a specific form of the model, which has the model risk. Realized volatility and realized jump method is a non-parametric method. It does not need to set the form of the model and can completely avoid the model risk. When we study the correlation, we use the complex network method. Comparing with the correlation coefficient, Granger causality test and regression model, the complex network has the advantages of intuition and high efficiency. (3) The results are more in-depth. This paper constructs a complex network of jump correlation and then uses the community method to study the clustering effect of jump correlation. Compared with the existing research, the results of this paper are more perfect.

This study has the following conclusions. (1) Jump volatility is an important part of the realized volatility of stock prices. Jump volatility contributes more to overall volatility during the financial crisis. (2) From the minimum spanning tree, it can be seen that when a stock price jumps, it will cause a jump in the price of a stock that is directly or indirectly connected to it; this shows that there is a certain correlation among stock price jumps. (3) It can be found that the manufacturing industry plays an important role in the minimum spanning tree network by calculating the closeness of the nodes. (4) After classifying the communities using the Modular Q function and Fast Unfolding algorithm, it is found that the average correlation coefficient of the stock price jump of the financial stocks community is the highest, and the average correlation coefficient of the communities formed by the retail and media industry stocks is the lowest.

2. Realized jump and complex network

2.1. Realized volatility

According to Andersen et al. [4,5], suppose the ith price on day t is $p_{t,i}$. The intraday return can be expressed as

$$r_{t,j} = p_{t,j} - p_{t,j-1}$$

By using the intraday high frequency return, the realized variance on that day can be obtained by the follow method:

$$RD_t = \sum_{j=1}^{n} r_{t,j}^2$$

RD_t is the realized variance. n represents the number of returns on that day. We can use the following method to get the realized volatility.

$$RV_t = \sqrt{W \cdot RD_t}$$

RV_t is the realized volatility. W is the number of the stock trading days in a year.

2.2. MinRV-Based jump separation method

The realization of double-power variation (BRV) and realized multiple power variation (MPV) are commonly used for separating realized jumps from the realized volatility. The effect of the jump achieved by multi-power variation is better than that of the jump achieved by double-power variation. Andersen et al. [4,5] proposed the MinRV method, using the idea of neighborhood truncation in view of MPV research:

$$MinRV_t = \frac{\pi}{\pi - 2}\frac{n}{n-1}\sum_{j=1}^{n-1} min\left(\left|r_{t,j+1}\right|, \left|r_{t,j}\right|\right)^2, t \in [0, T] \tag{1}$$

Andersen et al. [4,5] found that MinRV is an unbiased estimate of integral volatility (IV). Therefore, when n approaches infinity, $\mathrm{MinRV_t} \xrightarrow{P} \int_t^{t+T} \sigma_s^2 ds$. Assuming that the stock price jumps significantly at t, by continuously comparing two consecutive stock price returns and picking the one with the smallest absolute value, their sum of squares is MinRV.

Define MinRQ as:

$$MinRQ_t = \frac{\pi n}{\pi - 2}\frac{n}{n-1}\sum_{j=1}^{n-1} min\left(\left|r_{t,j+1}\right|, \left|r_{t,j}\right|\right)^4, t \in [0, T] \tag{2}$$

According to the results obtained by Andersen et al. [4,5], the corresponding intra-day jump detection $\mathrm{MinZ_t}$ in expressed as:

$$\mathrm{MinZ_t} = \frac{(\mathrm{RV_t} - \mathrm{MinRV_t})\,\mathrm{RV_t^{-1}}}{\sqrt{(3.81 - 2)\,\frac{1}{n}\max\left(1, \frac{\mathrm{MinRQ_t}}{\mathrm{MinRV_t}}\right)}} \to \mathrm{N}(0, 1),\ \ \mathrm{t} \in [0, \mathrm{T}] \tag{3}$$

2.3. Complex network

Complex networks can more systematically illustrate the correlations among realized jumps. They illustrate the correlation among jumps with network graphs. The minimum spanning tree algorithm is a common and simple network construction method in the complex network theory. It is highly efficient and stable, and the built-in network has good invulnerability. At present, the Kruskal algorithm, the Prim algorithm, the Circle-breaking method, the Sollin algorithm and the Dijkstra algorithm are commonly used to generate the minimum spanning tree. Prim's algorithm has the advantages of being concise and easy to understand, it is quick to implement, and it is difficult to make mistakes.

After constructing a complex network using the Prim algorithm, it is vital to analyze the importance of each node in the network. This paper uses closeness to the center to analyze the importance of each node in a complex network. The closeness to the center describes the importance of this node in the network by measuring the distance from one node to all the other nodes. It is one of the indexes that measure the importance of a node. Its formula is:

$$\mathrm{C_c}(\mathrm{v_i}) = (\mathrm{N} - 1)/[\sum_{\mathrm{i=1,j\neq 1}}^{\mathrm{N}} \mathrm{d_{ij}}] \tag{4}$$

where $C_c(v_i)$ is the degree of the closeness to the center, N is the number of nodes, and d_{ij} is the minimum number of edges between two nodes. i and j are belong to [1, N]. The greater the closeness of the node, the more the node is in the center of the network, and the more importance it has in the network.

2.4. Community clustering algorithm

A large number of studies have shown that there are many different communities in a complex network, and there are many similar nodes in one community. Nodes in the same community are more densely connected, and node connections in different communities are sparse. By splitting the complex network, it is easy to analyze the strength of the correlation among different nodes and the size of the clustering effect in the network. Therefore, this paper will divide the complex networks into groups and study the jump correlations among the inner and outer communities. There are two issues that need to be resolved when dividing the community: (1) the number of communities to be divided. This paper uses the Modular Q Function to solve this problem. (2) How to group jumps into a community. This paper uses the Fast Unfolding algorithm to group the communities.

2.4.1. The modular Q function

To quantitatively describe the structure of the community in a complex network and to discuss the pros and cons of its division, Girvan and Newman introduced the concept of a modular Q function in the process of studying complex community

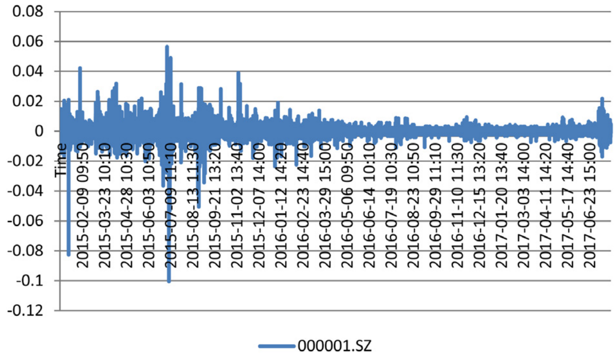

Fig. 1. 000001.SZ's 5-minute High-Frequency Return Series.

structures. Assuming that a complex network is divided into several communities by the community division algorithm, the proportion of edges connected to each other within communities can be calculated using the following formula:

$$\frac{\sum_{i,j} a_{ij}\delta(\sigma_i, \sigma_j)}{\sum_{i,j} a_{ij}} = \frac{\sum_{i,j} a_{ij}\delta(\sigma_i, \sigma_j)}{2M} \tag{5}$$

Among them, σ_i, andσ_j represent the community numbers of the nodes v_i and v_j, respectively, and they take any element a_{ij} in the adjacency matrix. If the nodes v_i and v_j are connected, $a_{ij} = 0$. If the nodes v_i and v_j are not connected, $a_{ij} = 0$. If the nodes v_i and v_j belong to the same community, $\delta\left(\sigma_i, \sigma_j\right) = 1$. If the nodes v_i and v_j do not belong to the same community, $\delta\left(\sigma_i, \sigma_j\right) = 0$. That is, $\sigma_i \neq \sigma_j$. M represents the total number of edges in a complex network, i.e., $M = \sum a_{ij}$. If the community structure is fixed, $k_i k_j/2M$ represents the possibility of connecting node v_i and v_j in a network with a randomly connected edge, where k_i is the degree of node v_i, and k_j is the degree of node v_j.

The expression of the Modular Q function is:

$$Q = \frac{1}{2M}\sum_{i,j}\left[\left(a_{ij} - \frac{k_i k_j}{2M}\delta(\sigma_i, \sigma_j)\right)\right] \tag{6}$$

The closer the Q value is to 1, the better the community division.

2.4.2. Fast unfolding algorithm

This study selected the Fast Unfolding algorithm to divide the network into communities. This algorithm is simple, efficient and easy to use. The Fast Unfolding algorithm is divided into two stages. The first stage is to pre-set the communities until each community belongs to a complex network and no change occurs. The second stage is to build a new image and repeat the first stage of the operation until the Q value no longer increases. A Modular Q function reduction formula can be expressed as:

$$Q = \frac{\sum in}{2M} - \left(\frac{\sum tot}{2M}\right)^2 \tag{7}$$

$\sum$ *in* represents the total number of edges connected to all nodes within the same community, and $\sum$ *tot* represents the total number of edges connected to nodes of the same community.

3. Empirical analyses

3.1. Data

This study uses the stocks' price of the CSI 300 Index from January 5, 2015 to July 28, 2017 to do the empirical research. The CSI 300 Index includes 300 stocks from industries such as manufacturing, finance, real estate construction, agriculture and animal husbandry, and service retailing. The data is processed as follows : (1) Remove the data of "fuse" period. (2) Remove stocks with incomplete trading days. There are 274 samples left. (3) Select the 5-minute frequency data for the analysis. (4) Assume that the jump is 0 during the suspension.

3.2. Realized volatility and jump extraction results

Before the empirical analysis, statistical properties of the sample need to be analyzed. Taking 000001.SZ as an example, Fig. 1 shows the time series of the return of Ping An Bank (000001.SZ).

Table 1
Statistical Properties of the 5-Minute High-Frequency Return of Some Stocks.[a]

Stock Symbol	Mean	Maximum	Minimum	Standard Deviation	Skewness	Kurtosis	ADF (*P*-value)
000001.SZ	1.86E−8	0.0565	−0.1004	0.0031	1.1622	87.6144	0.0001
000002.SZ	1.86E−5	0.0764	−0.1054	0.0039	−0.9045	83.7781	0.0001
…	…	…	…	…	…	…	…
601998.SH	−4.22E−6	0.0685	−0.0833	0.0040	0.1850	35.7287	0.0001
603993.SH	2.96E−5	0.0961	−0.1050	0.0059	1.1772	42.1293	0.0001

[a]Note: due to space limitations, this paper only reports on the statistical nature of the first two and last two stocks.

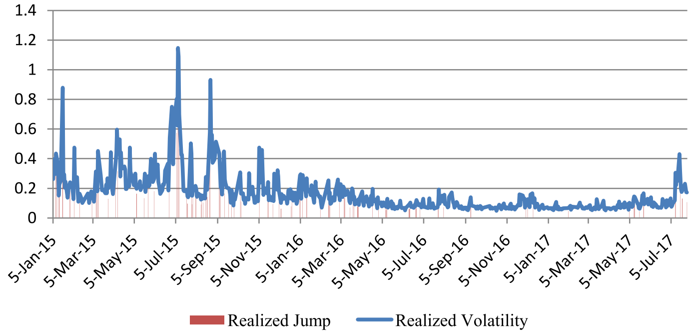

Fig. 2. Realized Volatility and Realized Jumps for 000001.SZ.

Table 2
Statistical Properties of Realized Volatility.

Stock Symbol	Mean	Maximum	Minimum	Standard Deviation	Skewness	Kurtosis	ADF (*P*-value)
000001.SZ	0.2632	1.1461	0.0480	0.2131	2.9292	12.5369	0.0000
000002.SZ	0.3242	1.8156	0.0000	0.2825	1.2934	2.9122	0.0000
…	…	…	…	…	…	…	…
601998.SH	0.5377	2.3124	0.0000	0.3476	1.5788	3.3188	0.0000
603993.SH	0.3215	2.0790	0.1586	0.1923	3.7644	20.5277	0.0000

Table 3
Statistical Properties of Realized Jumps.

Stock Symbol	Mean	Maximum	Minimum	Standard Deviation	Skewness	Kurtosis	ADF (*P*-value)
000001.SZ	0.0398	1.0248	0.0000	0.1315	5.9438	50.4805	0.0000
000002.SZ	0.0756	1.6746	0.0000	0.1839	3.3948	15.5357	0.0000
…	…	…	…	…	…	…	…
601998.SH	0.1229	1.6432	0.0000	0.2720	2.7368	8.0758	0.0000
603993.SH	0.0258	1.0452	0.0000	0.1122	5.1908	29.8758	0.0000

During the stock-market-crash period in 2015, the 5-minute high-frequency yield of 000001.SZ fluctuates significantly, so we can deduce that the stock price jumps significantly during this period. Since April 2016, stock market tends to be stable. We can deduce that during this period, there is hardly a big jump in stock prices. This study conducts a statistical analysis of the series of high-frequency returns of the first three and the last three stocks of the CSI 300 Index, as shown in Table 1:

Take 000001.SZ as an example. The *p*-value of the ADF test is 0.001. This result indicates that the return of 000001.SZ is stationary. The series can be used to calculate the realized volatility and jumps, as shown in Fig. 2:

The results in Fig. 2 confirm the previous assumptions: during the stock-market-crash period, the volatility of 000001.SZ reaches 114%, and the jump in its stock price contributed significantly to the volatility. After April 2016, the stock market was more stable, and the volatility of Ping An Bank decreased to 20%–30%. At this time, there was no significant jump. Tables 2 and 3 show the statistical properties of the realized volatility and jumps.

3.3. *Results of the jump network*

3.3.1. *The coefficient matrix and distance matrix of the jump correlation*

The jump correlation coefficient matrix $\rho_{\mathrm{ij}}\left(\Delta \mathrm{t}\right)$ can be directly constructed. The results are shown in Table 4:

The jump correlation coefficient matrix $\rho_{\mathrm{ij}}\left(\Delta \mathrm{t}\right)$ is transformed into the corresponding distance d (i, j):

$$d\left(i,j\right)=\sqrt{2\left[1-\rho_{ij}\left(\Delta t\right)\right]}$$

Table 4
Correlation Coefficient Matrix among Jumps.

	000001.SZ	000002.SZ	000008.SZ	…	601992.SH	601998.SH	603993.SH
000001.SZ	1.0000	0.6941	−0.0113	…	0.4668	0.3131	0.3760
000002.SZ	0.6941	1.0000	−0.0012	…	0.3832	0.1159	0.3286
000008.SZ	−0.0113	−0.0012	1.0000	…	0.2541	0.1213	0.1838
…	…	…	…	…	…	…	…
601992.SH	0.4668	0.3832	0.2541	…	1.0000	0.3246	0.3914
601998.SH	0.3131	0.1159	0.1213	…	0.3246	1.0000	0.1584
603993.SH	0.3760	0.3286	0.1838	…	0.3914	0.1584	1.0000

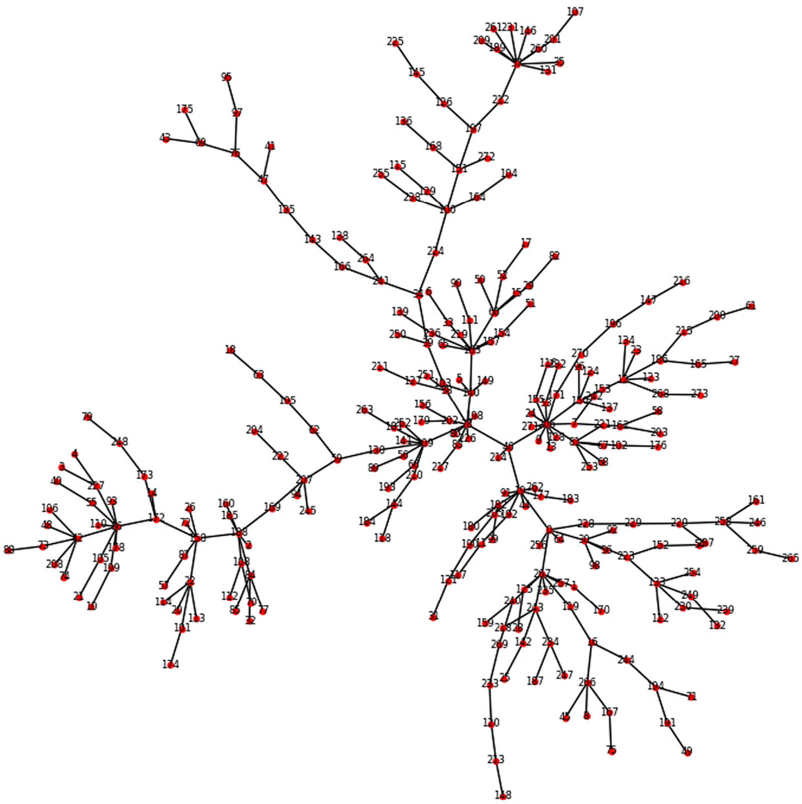

Fig. 3. Minimum Spanning Tree Network for Stock Volatility Jump.

Table 5
Distance Matrix between Jumps.

	000001.SZ	000002.SZ	000008.SZ	…	601992.SH	601998.SH	603993.SH
000001.SZ	0.0000	0.7822	1.4222	…	1.0327	1.1721	1.1171
000002.SZ	0.7822	0.0000	1.4150	…	1.1107	1.3297	1.1588
000008.SZ	1.4222	1.4150	0.0000	…	1.2214	1.3257	1.2777
…	…	…	…	…	…	…	…
601992.SH	1.0327	1.1107	1.2214	…	0.0000	1.1622	1.1032
601998.SH	1.1721	1.3297	1.3257	…	1.1622	0.0000	1.2974
603993.SH	1.1171	1.1588	1.2777	…	1.1032	1.2974	0.0000

The greater the distance d (i, j) is, the weaker the correlation between stocks, and vice versa. Table 5 shows the results:

3.3.2. Constructing the stock volatility jump network

As shown in Fig. 3, this study uses the minimum spanning tree algorithm to build the jump network. The distance between each pair of nodes represents the correlation between nodes. From Fig. 3 we can see that: (1) The nodes that live in the center of the complex network have stronger influence, such as nodes 78 and 79. When the price of these nodes jumps, they cause a lot of jumps in the stock price associated with them. (2) The nodes at the edge of the network are less influential in the network, such as node 107 and 265. When the stock price of these nodes jumps, they will only refer to a small number of

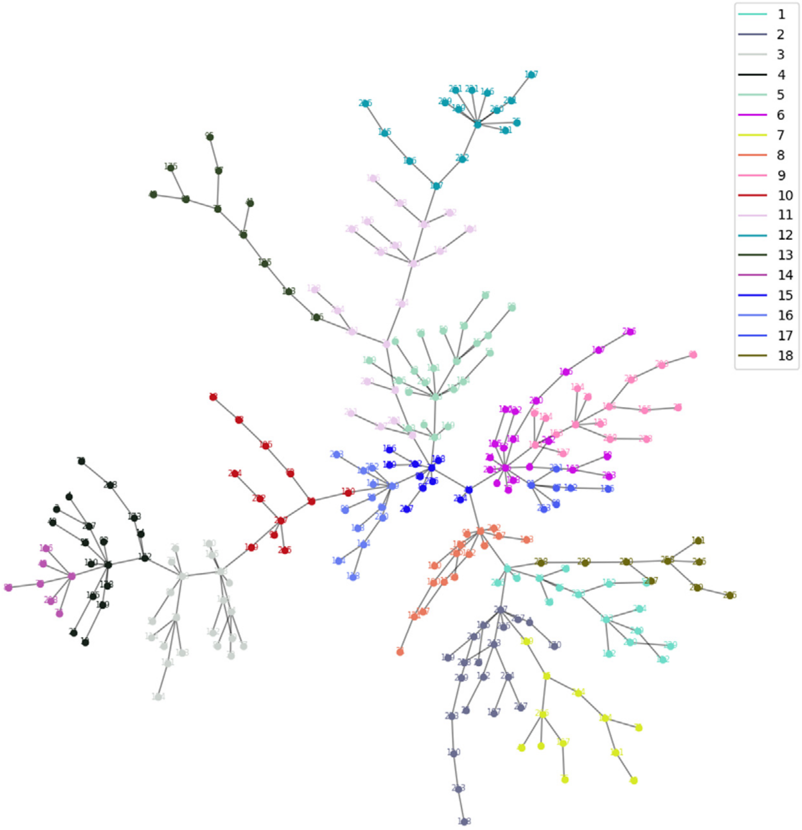

Fig. 4. Community Division Diagram of Stock Volatility Associated Network.

Table 6
Stock Volatility Jump Associated Network Closeness Centrality Top 10 Ranking Table.

Rank	Stock Symbol	Industry	Closeness Centrality	Rank	Stock Name	Industry	Closeness Centrality
1	002304.SZ	Manufacturing	0.1701	6	600795.SH	Energy	0.1498
2	000895.SZ	Manufacturing	0.1668	7	600089.SH	Manufacturing	0.1489
3	300124.SZ	Manufacturing	0.1573	8	600718.SH	Communication	0.1458
4	000538.SZ	Manufacturing	0.1550	9	002415.SZ	Manufacturing	0.1456
5	000776.SZ	Finance	0.1533	10	002352.SZ	Transportation	0.1454

the stock prices to jump. (3) The shorter the edge between two nodes, the stronger the correlation between the jumps of the two nodes. (4) If a node is not connected to others, the jumps of this node will not affect any others, so the node will not appear in the complex network. From the results of this paper, there are 274 nodes in this complex network. We can see that there is correlation among all the nodes.

3.3.3. Analysis of the importance of network nodes based on closeness

Based on the minimum spanning tree, this study proceeds to use the closeness centrality to analyze the effect of each node on the stability and invulnerability of the network. The greater the closeness of the node, the more important the node is in the network and the closer it is to other nodes. In Table 6 we can see that the top four stocks with the highest closeness centrality values are 002304.SZ, 000895.SZ, 300124.SZ, and 000538.SZ. These four stocks belong to the manufacturing industry. Manufacturing stocks account for more than half of the top ten stocks. This shows that manufacturing is prominent in the minimum spanning tree network.

3.4. Community structure division

In this study, the Fast Unfolding Algorithm is used to classify the jump network. The closer the Modular Q value is to 1, the better the community division within this network. The results are shown in Table 7.

As shown in the table above, when the stock volatility jump network is divided into 18 communities, the Modular Q value reaches 0.8877, which is close to 1, indicating the best community division at this time (see Fig. 4).

Table 7
Modular Q Statistics.

Number of Organization	Modular Q Value	Number of Organization	Modular Q Value	Number of Organization	Modular Q Value
89	0.6981	29	0.8748	18	0.8877

Table 8
Statistics of Community Average Correlation Coefficients.

Community ID	Stock Symbol	Average Correlation Coefficient
Community 1	000001.SZ,002142.SZ,002673.SZ,601800.SH,000750.SZ,600048.SH,600050.SH, 601186.SH,601390.SH,001979.SZ,601766.SH,600188.SH,601668.SH,600019.SH, 002500.SZ,000686.SZ,601088.SH	0.4432
Community 2	000002.SZ,601818.SH,601288.SH,600383.SH,601318.SH,601328.SH,601939.SH, 600104.SH,601988.SH,601398.SH,601601.SH,600153.SH,600886.SH,601628.SH, 600016.SH,000625.SZ,600999.SH,600585.SH,000651.SZ,600028.SH,600276.SH	0.5494
Community 3	000008.SZ,000540.SZ,000559.SZ,000627.SZ,600271.SH,000718.SZ,600297.SH, 600446.SH,600570.SH,002027.SZ,600588.SH,002195.SZ,002230.SZ,002299.SZ, 002424.SZ,002426.SZ,002456.SZ,300017.SZ,300104.SZ,300168.SZ,300315.SZ	0.3331
Community 4	000009.SZ,000060.SZ,601118.SZ,601800.SH,600703.SH,000792.SZ,300059.SZ, 000166.SZ,600415.SH,600436.SH,000413.SZ,002310.SZ,300133.SZ,600498.SH, 000555.SZ,002007.SZ,601633.SH,002508.SZ	0.3149
Community 5	000062.SZ,000069.SZ,600085.SH,600089.SH,000415.SZ,000425.SZ,600157.SH, 600208.SH,600256.SH,000671.SZ,000725.SZ,600340.SH,000961.SZ,000963.SZ, 000977.SZ,002065.SZ,002146.SZ,002411.SZ,002470.SZ,601006.SH,300027.SZ, 300144.SZ,601333.SH	0.3600
Community 6	600031.SH,600332.SH,600685.SH,000100.SZ,601225.SH,000157.SZ,600406.SH, 600737.SH,000402.SZ,600795.SH,601989.SH,601992.SH,601600.SH,600150.SH, 600008.SH,000983.SZ,000623.SZ,002044.SZ,600221.SH,600895.SH	0.5215
Community 7	600674.SH,002202.SZ,000156.SZ,600372.SH,601933.SH,002292.SZ,000876.SZ, 000423.SZ,000959.SZ,601607.SH,600015.SH,600649.SH	0.4024
Community 8	600660.SH,002714.SZ,601888.SH,000333.SZ,000858.SZ,600741.SH,600489.SH, 000538.SZ,600519.SH,600009.SH,600535.SH,600547.SH,600018.SH,002476.SZ, 600637.SH,000709.SZ	0.4435
Community 9	002594.SZ,600893.SH,600362.SH,600060.SH,000630.SZ,600068.SH,600704.SH, 000338.SZ,601958.SH,603993.SH,600170.SH,600196.SH,000568.SZ,600583.SH, 600021.SH,600023.SH,002074.SZ	0.4143
Community 10	600037.SH,600682.SH,600376.SH,600739.SH,600804.SH,000503.SZ,601608.SH, 002081.SZ,002049.SZ,601021.SH,002555.SZ,002131.SZ	0.3399
Community 11	600036.SH,600066.SH,601901.SH,600074.SH,601998.SH,600177.SH,000728.SZ, 600352.SH,000776.SZ,000783.SZ,600373.SH,600837.SH,601099.SH,002736.SZ, 300033.SZ,601377.SH,601555.SH,600000.SH,601669.SH,601688.SH,601788.SH, 600030.SH	0.5158
Community 12	601111.SH,000738.SZ,600038.SH,000768.SZ,600028.SH,601216.SH,601872.SH, 600705.SZ,601877.SH,300072.SZ,600115.SH,600118.SH,600820.SH,600871.SH, 600606.SH,600029.SH	0.4681
Community 13	002602.SZ,002558.SZ,002183.SZ,000793.SZ,000839.SZ,300251.SZ,600109.SH, 600369.SH,600482.SH,600061.SH	0.1281
Community 14	000826.SZ,000938.SZ,002236.SZ,002241.SZ,002265.SZ,300070.SZ,600816.SH	0.3508
Community 15	601117.SH,600690.SH,600718.SH,002304.SZ,002352.SZ,002415.SZ,000895.SZ, 600233.SH,600518.SH,600887.SH,600900.SH	0.4106
Community 16	600663.SH,002152.SZ,601899.SH,300124.SZ,600100.SH,600010.SH,600111.SH, 600827.SH,600522.SH,002008.SZ,002466.SZ,601718.SH,600549.SH	0.4137
Community 17	002153.SZ,002174.SZ,002385.SZ,300024.SZ,600485.SH,601018.SH,601727.SH	0.4237
Community 18	600297.SH,601166.SH,601169.SH,601336.SH,601009.SH,601618.SH,601857.SH, 601866.SH,601919.SH	0.4206

It can be seen from the figure that different colors represent different communities. By calculating the average correlation coefficients of jumps within each community, the jump correlation between communities can be compared, as shown in Table 8.

Table 8 shows that the average correlation coefficient of community 2 is the highest, and most stocks in community 2 are from the financial industry. This indicates that when volatility jumps, the correlations among financial stocks are relatively strong. Fluctuations in the prices of such stocks will cause the price of other financial stocks to fluctuate. Second, the average correlation coefficients of communities 6 and 11 are greater than 0.5 (0.5215 and 0.5158, respectively). Community 6 is mainly composed of manufacturing and energy industry stocks, indicating the close relationship between the manufacturing and energy industries. When the prices of manufacturing stocks fluctuate significantly, this may cause energy stocks to fluctuate as well. Community 11 is mainly composed of financial and manufacturing stocks, indicating a strong correlation between volatility jumps of financial and manufacturing stocks. In addition, the average correlation coefficient of community 13 is 0.1281, which is mainly composed of retail and media industry stocks. The correlation of stock price jump is weaker than that of other communities.

4. Conclusions

In this study, the minimum spanning tree algorithm was used to build a complex network of the CSI 300 Index stock jumps, enabling the authors to reach the following conclusions. (1) We use the realized volatility and the realized jump method to extract the volatility and jump of the sample. In this research, we do not set any model form. We calculate the realized volatility and jump by using the information from the high-frequency data. It has the characteristics of high information utilization and low model risk. Jump volatility is an important part of the realized volatility of the stock price. Jump volatility contributes more to overall volatility during a crisis. (2) We construct a complex network by using the minimum spanning tree method, which can visually see the correlation between stock price jumps. From tree, it can be seen that a stock price jump will cause price jumps in the stocks that are directly or indirectly connected to it, revealing a certain correlation between stock price jumps. (3) By calculating the value of the closeness centrality of the entire node network, it was found that the manufacturing industry has an important position in the minimum spanning tree network. The result shows that manufacturing industry is at the core of the whole industry. In order to reduce the jumping risk of the stock market, it is very important to further develop the manufacturing industry. (4) After classifying the communities using the Modular Q function and the Fast Unfolding algorithm, it was found that the average correlation coefficient of the stock price jump of financial stocks is the highest, while the average correlation coefficient of the retail media stocks in community 13 is the lowest. The result shows that the jump risk contagion in the financial industry is higher than that in other industries, and further financial risk prevention is needed to prevent the chain reaction of jumping risk in the financial industry.

References

[1] T.G. Andersen, T. Bollerslev, Deutsche mark–dollar volatility: intraday activity patterns, macroeconomic announcements, and longer run dependencies, J. Finance 53 (1) (1998) 219–265.
[2] O.E. Barndorff-Nielsen, N. Shephard, Econometric analysis of realized co-variation: high frequency based covariance, regression, and correlation in financial economics, Econometrica 72 (3) (2004) 885–925.
[3] C.F. Wang, Y. Ning, Z.M. Fang, et al., An empirical research on jump behavior of realized volatility in chinese stock markets, Syst. Eng. 26 (2) (2008) 1–6.
[4] T.G. Andersen, T. Bollerslev, X. Huang, A reduced form framework for modeling volatility of speculative prices based on realized variation measures, J. Econometrics 160 (1) (2011) 176–189.
[5] T.G. Andersen, T. Bollerslev, X. Huang, A reduced form framework for modeling volatility of speculative prices based on realized variation measures, J. Econometrics 160 (1) (2011) 176–189.
[6] Z.F. Zeng, J. Zuo, Research on relationship of jump and volatility spillover between shanghai and hong kong stock market returns ——based on svcj model by mcmc algorithm, Chinese J. Manag. Sci. 21 (S1) (2013) 334–340.
[7] W.J. Xu, F. Ma, Y. Wei, Investigating the performance of the high-frequency volatility models using the nonparametric jump tests, Syst. Eng. 34 (12) (2016) 10–16.
[8] Y.Q. Song, S.O. Wang X.J. Economics, et al., Impact of jump intensity on the modeling and forecasting of stock market volatility, J. Finance Econ. 31 (01) (2016) 85–95.
[9] R.N. Mantegna, Hierarchical structure in financial markets, Eur. Phys. J. B 11 (1) (1999) 193–197.
[10] X.W. Zhuang, X. Jin, The research of correlation between network topological index and volatility of shanghai stock market, J. Dongbei Univ. 36 (3) (2015) 453–456.
[11] A. Pothen, H.D. Simon, K.P. Liou, Partitioning sparse matrices with eigenvectors of graphs, SIAM J. Matrix Anal. Appl. 11 (3) (1990) 430–452.
[12] K.T. Chi, J. Liu, F.C.M. Lau, A network perspective of the stock market, J. Empirical Finance 17 (4) (2010) 659–667.
[13] T.K. Peron, L.F. Costa, F.A. Rodrigues, The structure and resilience of financial market networks, Chaos 22 (1) (2012) 193.
[14] M.E.J. Newman, Fast algorithm for detecting community structure in networks, Phys. Rev. E 69 (6 Pt 2) (2003) 066133.
[15] M.E.J. Newman, M. Girvan, Finding and evaluating community structure in networks, Phys. Rev. E 69 (2) (2004) 026113.
[16] W.W. Zhan, J.L. Xi, Z.X. Wang, Hierarchical agglomerative community detection algorithm based on similarity modularity, J. Syst. Simul. 29 (5) (2017) 1028–1040.