Journal of Multidisciplinary Engineering Science and Technology (JMEST)
ISSN: 2458-9403
Vol. 8 Issue 7, July - 2021

# Analysis of vital nodes in Chinese stock market network during the COVID-19 pandemic

**Hongxing Yao[1]**
School of Finance and Economics,
Jiangsu University, Zhenjiang 212013, China

**Chengxuan Guo[2]**
School of Mathematics and Science,
Jiangsu University, Zhenjiang 212013, China
Email:371401651@qq.com

***Abstract*—To analyze the impact of the COVID-19 pandemic on Chinese stock market, we use Granger causality test to construct four stock market networks in different periods. We define the influence of a stock according to the overlapping influence algorithm of Rayleigh entropy. We test the validity of our algorithm and compare the influence value with other topological features of the stock. We find that the impact of COVID-19 on Chinese stock market is much lower than that of the turbulence of 2015-2016. When the crisis occurs, the network density and the aggregation of important nodes are significantly higher. From the perspective of industry classification, "Real estate" and "Finance" are the most influential sectors. Furthermore, analyzing the influential individual stocks in the industry helps the government prevent and regulate stock market risks.**

***Keywords—COVID-19; Complex network; Granger causality test; The overlapping influence algorithm; Chinese stock market***

## I. Introduction

With the further opening of Chinese financial market, the financial crisis facing Chinese stock market has gradually increased. In the past decade, Chinese stock market has suffered from two major crises: the turbulence of the stock market from 2015 to 2016, and the social crisis caused by COVID-19. In the initial stage of the epidemic, Chinese stock market was affected to a certain extent due to the impact of social stagnation. Therefore, it is necessary to study the stock market development during the epidemic period.

Using correlation matrices to construct networks to study financial markets first appeared in Mantegna's paper[1]; Mantegna[2] . Then, many scholars used complex networks to study the spread of financial risks[3]; Smolyak et al.[4]; Wu and Duan[5]; He et al.[6]; ; Xu et al. [7]. Newman[8] believes that the stock market is complex and highly coupled, and the use of complex network method to analyze the stock market can clearly show the correlation between different stocks. Researchers treat stocks as nodes and use correlations between them to build networks of stocks. Chi et al. [9] constructed a network model of the American stock market and found that the stock price were strongly influenced by the few stocks in the stock market, and concluded that the stock market was mainly dominated by the financial industry. Jung and Chang[10] used Pearson correlation coefficient to construct a complex network and analyzed the Korean stock market from 2004 to 2014 through clustering effect. Wang et al. [11] constructed two networks based on Pearson correlation and partial correlation respectively for comparison, and analyzed the correlation and evolution of global stock markets. Hu et al.[12] used the minimum spanning tree algorithm to construct the Chinese stock market network and study the correlation between different stocks. In order to further determine the long-term stable relationship between stocks, the co-integration relationship is used in stock market research. In addition, the Granger causality test has been widely used to determine the influence relationship between stocks and construct a directed financial network[13]; [14]; Song et al.[15]; Gong et al.[16]; Memon et al. [17]. Gao et al. [18] established a comparative analysis of stock networks through co-integration relationship and Granger causality to identify the propagation path of potential risks. Wang et al.[19] used the influence of a single stock price fluctuation on the price of other stocks to establish the Granger causality network, and analyzed the influence ability and industry influence of different stock price changes. Bu et al.[20] used Granger causality test to establish Chinese stock market networks in different periods, and through the research found that there was a higher correlation between stocks in the period of financial crisis in 2008 and the period of stock market turbulence in 2015.

Identifying influential nodes is a crucial issue in complex networks. For the stock market, the identification of vital stocks is conducive to the awareness and control of risk proliferation. The influence of nodes can be analyzed based on topological characteristics, such as the degree, k-shell decomposition, centrality measures, H-index, etc[21]; Wang et al.[22]; Liu et al.[23]; L{\"u} et al.[24]; Ma and Ma [25]; Li et al.[26]. L{\"u} et al. [27] compared various node influence algorithms according to different types of networks. In order to consider the overlapping influence of nodes, the relationship between single node influence and multi-node comprehensive influence is analyzed based on Rayleigh entropy (Zhou et al. [28]).

The rest of this paper is organized as follows. In section 2, we discuss the Granger causality network construction method. In section 3, we describe the influence algorithm of vital nodes. In section 4, we test the validity of the algorithm and make an empirical analysis. Some conclusions and discussions are given in section 5.

## II. Network model construction

In this section, we first briefly describe the data processing of the stock closing price before establishing the network model. Then, we use Granger causality to test the relationship between different stocks and build a directed network. All of the statistical tests used in our research are assessed using the significance level of 0.05. The process of how to build the model is as follows.

To make the data more stable, we first perform the logarithmic differential treatment to the time series of the closing price of each stock. We define $R_t^i$ as the closing price of stock $i$ at time $t$, the equation of which is shown as follows:

$$\varDelta \ln R_i^t = \ln R_i^t - \ln R_i^{t-1} = \ln \frac{R_i^t}{R_i^{t-1}} \quad (1)$$

We use the stationary of price series to set up the Granger causality test to avoid spurious regression (Granger [29]). Therefore, the logarithmic differential time series is tested for stationarity using ADF test to obtain a stationary time series (Phillips and Perron [30] ).

We use the ADF test to filter out the stable time series and perform the Granger causality test. Our research focuses on the effect of one stock on other stocks. Therefore, we ignore its autocorrelation in the stock network. The equation of which is shown as follows (Granger[31] ):

$$R_X^t = \alpha + \sum_{i=1}^{m} \lambda_i R_X^{t-i} + \sum_{j=1}^{n} \theta_j R_Y^{t-j} + \varepsilon_t \quad (2)$$

$$R_Y^t = \delta + \sum_{i=1}^{r} c_i R_x^{t-i} + \sum_{j=1}^{s} d_j R_Y^{t-j} + \eta_t \quad (3)$$

Through the above tests, a stock is regarded as a network node, and the causal relationship between different stocks is regarded as an edge. We define the set of nodes as $V$ and the set of edges as $E$. Then, we can construct a directed stock network, which is presented by an adjacency matrix of $A = (a_{ij})$. If stock $i$ Granger causes stock $j$, we record it as $a_{ij} = 1$, otherwise it is recorded as $a_{ij} = 0$. The stock network is denoted $G = (V, E)$.

## III. Algorithms of vital nodes

Identifying a minority of vital nodes in the network is the key to analyzing the stock network, which helps control the further spread of the crisis when it occurs. To determine the comprehensive influence of important nodes, the overlapping influence algorithm based on Rayleigh entropy is used (Zhou et al.[32] ).

Definition 1 Let the influence of an edge $e_{ij}$ is $S_e$,$\mathrm{d}_i$ is the degree of node $\mathrm{i}$ and $N_i$ is the set of neighbor nodes of node $\mathrm{i}$. The equation of which is shown as follows:

$$s_{e_{ij}} = d_i \sum_{k \in N_j} d_k + d_j \sum_{k \in N_i} d_k \quad (4)$$

Definition 2 The influence of a single node $m$ is actually the sum of the influence of all the edges connecting the node. The equation of which is shown as follows:

$$S_m = d_m \sum_{i,j} a_{m_j} a_{jk} d_k + \sum_{j,k} a_{jm} a_{m_k} d_j d_k \quad (5)$$

Considering that the Granger Causality Network is a directed network, and we focus on studying the influence of one stock on another stock in the market. Thus, we should consider the importance of nodes with out-degrees. Therefore, we let $d_i$ is the out-degree of node $i$.

## IV. Data and empirical analysis

### A. Data

In this study, we select stocks in the CSI 300 Stock Index and collected stock trading data from January 1, 2014 to January 30, 2016 and November 1, 2018 to October 1, 2020.Among them, some companies in the CSI 300 Stock Index are not listed or suspended trading during that period and filter to get 208 stocks. Therefore, we select 92 stocks from the Shanghai and Shenzhen A-shares, and use these 300 stocks as research data. In order to study the impact of COVID-19 on Chinese stock market and compare the Chinese stock market's turbulence during the period from 2015 to 2016, we choose four periods as the data for this study.Period 1:before the turbulence of 2015-2016 (from January 1, 2014 to January 1, 2015);Period 2: during the turbulence of 2015-2016 (from January 30, 2015 to January 30, 2016);Period 3: before the COVID-19 pandemic (from November 1, 2018 to November 1, 2019);Period 4: during the COVID-19 pandemic (from December 1, 2019 to December 1, 2020).

### B. Stock market network

Firstly, we construct the stock network diagrams of four periods through Granger causality, and use Fruchterman-Reingold algorithm to make the network diagram layout. Secondly, we use the influence value of the stock to sort the network node sizes in order. In addition, in order to better highlight the aggregation of important nodes, four colors of red, purple, blue and yellow will be used to arrange according to the influence value of nodes from high to low. Fig.1 shows that the stock market has undergone significant changes. The aggregation of important nodes in the network diagram during period 2 is higher than period 1. Similarly, the network diagrams of the period 3 and 4 also showed the same situation. This indicates that the aggregation of important nodes increases when the crisis occurs. During the critical period, due to the high aggregation of nodes, important nodes are impacted, and risks will spread rapidly among these nodes, which will cause the entire stock network to be greatly affected in a short time. Compared with period 2, the aggregation of important nodes in period 4 was lower, indicating that the risk spread intensity and speed of stock network in period 4 were much lower than in period 2.

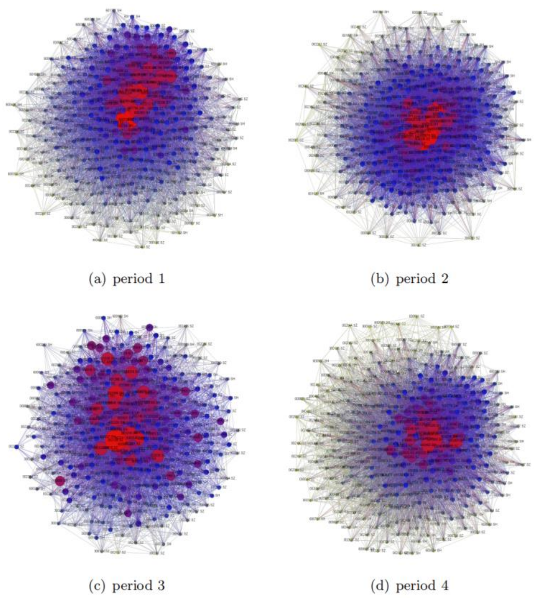



Fig.1: Causal network diagrams of four different periods obtained by the FR algorithm

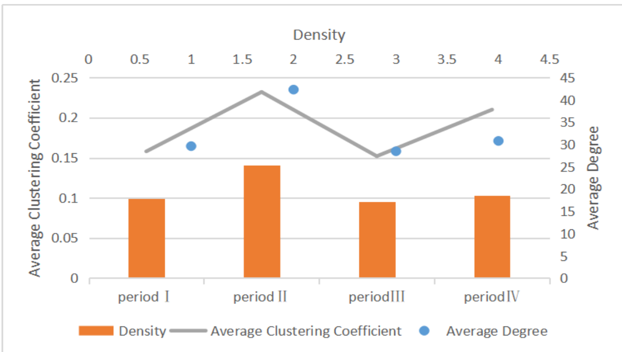


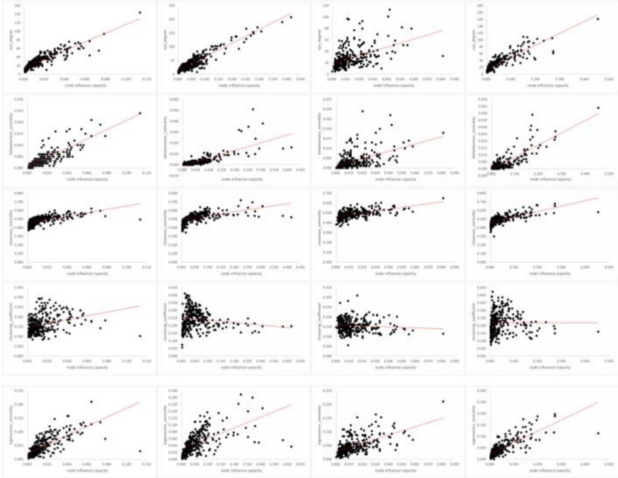

Fig.2: Topological characteristic statistical graphs of network graphs in four periods

We calculated the topological characteristics of the network, such as the network density, average clustering coefficient, and average degree. It can be seen from Fig.2 that all the topological characteristic values in period 2 are the highest, while all the topological characteristic values in the periods 1, and 3 are lower, and the average clustering coefficient in the periods 2, and 4 is close to each other. From the perspective of network density, there was no significant difference between the network tightness in the periods 1, 3, and 4, while the network tightness in period 2 increased significantly, which indicated that the network was more densely connected and the risk was more easily spread. From the perspective of average clustering, both period 2 and period 4 are relatively high, which can also be seen from the aggregation of important nodes in the network. With the linkages of stocks increase during the turbulence, the financial system becomes more closely connected. In terms of average degree, period 2 is the highest, while the other three periods are almost the same, which indicates that the correlation between nodes increased during period 2.

In general, when a crisis occurs, the topology characteristics of the network will increase to a certain extent, and the linkage between nodes will also increase. However, the network topology values during the COVID-19 pandemic are all lower than the turbulence of 2015-2016, because the crisis period of the COVID-19 pandemic is much shorter than the turbulence of 2015-2016. The main reason is that when the epidemic broke out, Chinese government exercises management by sealing off entities, which greatly reduced the spread of the virus in a short period of time and quickly restored confidence in the stock market.

### C. Influence of stocks

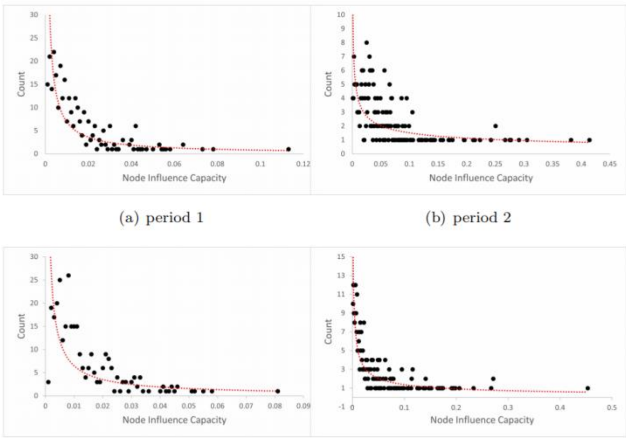
To verify the effectiveness of the node influence algorithm, we compare the value of stock influence with other topological features, such as the out-degree, betweenness centrality, closeness centrality, cluster coefficient, and eigenvector centrality. These topological features show the direct influence, indirect influence, intermediary location, average distance, and aggregation advantage of nodes respectively.



(a) period 1 (b) period 2 (c) period 3 (d) period 4

Fig.3: We compare influence values with other topological features in the scatter plot

Fig. 3 shows that the influence value of the stock is highly correlated with the out-degree, betweenness centrality, and eigenvector centrality. The out-degree directly expresses the influence of a node on other nodes. The betweenness centrality reflects the number of the shortest path through a node. The eigenvector centrality highlights the importance of the neighbor nodes of a node. According to these three topological features, it is shown from the side that the algorithm for calculating the influence of stocks is of validity, and can better highlight the influence level of a single stock in the network.



(a) period 1 (b) period 2

(c) period 3 (d) period 4

Fig.4: Scatter plots of influence values in four periods. The red line is the power law regression

To analyze the distribution of stock influence, we draw a scatter chart of influence in each period. Fig.4 shows that the distribution of nodes in the periods 1 and 3 is close, and the situation in the periods 3 and 4 is similar. Additionally, it can be found that the value of the maximum node influence in the periods 1 and 3 is 0.113 and 0.081 respectively, while the value of the maximum node influence in the periods 2 and 4 is 0.414 and 0.453 respectively. This indicates that when the stock market crisis occurs, the node influence will generally increase. The greater influence of nodes in four periods, the fewer amount of nodes, and the nodes with lower influence account for the majority. The comparison between the periods 2 and 4 shows that the amount of nodes with influence greater than 0.2 in period 4 is significantly lower than that in period 2, indicating that the influence level of key nodes in period 4 is much lower than that in period 2. Therefore, it proves that the stock market crisis in period 4 is better than in period 2.

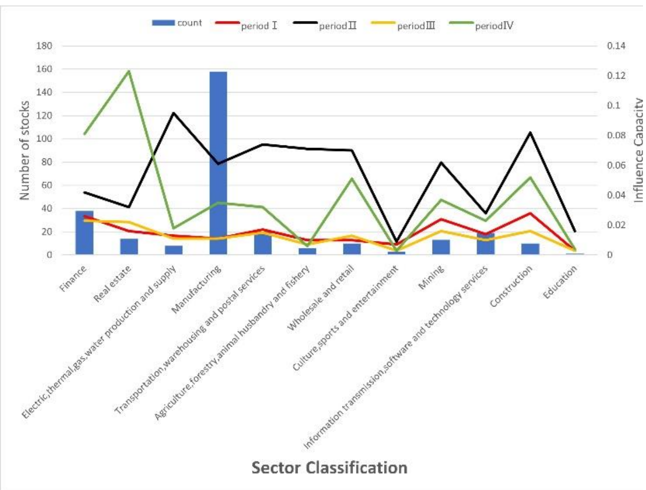



Fig.5: The influence distribution of each industry in the four periods

From Fig.5, it can be found that the top two industries with influence in each period are as follows, period 1: “Construction” and “Finance”, period 2: “Electric, thermal, gas, water production and supply” and “Construction”, period 3: “Finance” and “Real estate”, period 4: “Real estate” and “Finance”. From the perspective of period 1 and 2, the construction industry has mainly benefited from the country's heavy investment in infrastructure over the past years, finance mainly includes major state-owned banks, and the energy industry is an important part of blue chip stocks. Therefore, these three industries have shown significant influence in the stock market from 2014 to 2016. From the perspective of the periods 3 and 4, “Real estate” and “Finance” dominate, which is in line with the actual development situation in China. Because of the rapid growth of Chinese real estate industry in recent years, the stock market is confident in the real estate industry. In addition to covering the stocks of many state-owned banks, the financial industry also benefits from the support of insurance and securities companies. In addition, Chinese financial market will begin to open up in an all-round way in 2020. Although manufacturing occupies a huge advantage in terms of quantity, its influence in each period is at an average level. Because there are more traditional manufacturing industries and fewer state-of-the-art manufacturing industries.

## V. Conclusion

This paper mainly constructs four stock market networks in different periods through Granger causality. First, we found the commonalities and differences between the turbulence of 2015-2016 and COVID-19 epidemic period by comparing the network diagrams. Second, we use the node influence algorithm to calculate the influence of stocks in different periods. Then, we compare the node influence value with different topological features and get the correlation between them to demonstrate the effectiveness of the algorithm. Finally, we classify different stocks by industry, and analyze the influence of the industry in different periods and the important influential stocks in the industry. The following conclusions are of practical significance for policy decisions and prevention of the spread of stock market risks.

(1) During the COVID-19 pandemic, China's stock market was hit, but compared with the stock market crash from 2015 to 2016, it is still in a relatively stable state. The main reason is that Chinese government exercises management by sealing off entities, which greatly reduced the spread of the virus in a short period of time and quickly restored confidence in the stock market.

(2) When the nodes of higher influential stocks are clustered, the entire stock market network is in a moment of crisis. At this time, once it is impacted by risks, it will cause the stock market to crash.

(3) From the analysis of industry classification, the real estate and financial industries have been the most influential industries in the past two years. This is consistent with the facts of Chinese development.

(4) Because of the impact of the COVID-19 pandemic and the opening of Chinese financial market, the influence of pharmaceutical and securities stocks has increased.

**Acknowledgment**

This work was supported by the National Natural Science Foundation of China no. 71701082. Authors are grateful to the editor and the anonymous referees for their insightful remarks which greatly improved this paper. We would like to thank the National Natural Science Foundation of China. This work would not have been possible without their support.

**References**

1.Mantegna, R.N. and H.E. Stanley, *Introduction to econophysics: correlations and complexity in finance*. 1999: Cambridge university press.

2.Mantegna, R.N., *Hierarchical structure in financial markets.* The European Physical Journal B-Condensed Matter and Complex Systems, 1999. **11**(1): p. 193-197.

3.Tabak, B.M., T.R. Serra, and D.O. Cajueiro, *Topological properties of stock market networks: The case of Brazil.* Physica A: Statistical Mechanics and its Applications, 2010. **389**(16): p. 3240-3249.

4.Smolyak, A., et al., *Interdependent networks in Economics and Finance—A Physics approach.* Physica A: Statistical Mechanics and its Applications, 2018. **512**: p. 612-619.

5.Wu, B. and T. Duan, *Nonlinear dynamics characteristic of risk contagion in financial market based on agent modeling and complex network.* Complexity, 2019. **2019**.

6.He, Y., S. Wu, and M. Tong, *Systemic risk and liquidity rescue in complex financial networks: Pit hole and black hole of liquidity.* Physica A: Statistical Mechanics and its Applications, 2019. **536**: p. 121005.

7.Xu, R., et al., *Complex network construction of Internet finance risk.* Physica A: Statistical Mechanics and its Applications, 2020. **540**: p. 122930.

8.Guimerà, R., *Networks: an introduction.* Oxford University Press, 2012.

9.A, C.K.T., J.L.A. B, and F.C.M.L. A, *A network perspective of the stock market.* Journal of Empirical Finance, 2010. **17**(4): p. 659-667.

10.Jung, S.S. and W. Chang, *Clustering stocks using partial correlation coefficients.* Physica A: Statistical Mechanics and its Applications, 2016. **462**: p. 410-420.

11.Wang, G.J., X. Chi, and H.E. Stanley, *Correlation Structure and Evolution of World Stock Markets: Evidence from Pearson and Partial Correlation-Based Networks.* Computational Economics, 2018. **51**(3): p. 1-29.

12.Sh, A., et al., *An analysis of the clustering effect of a jump risk complex network in the Chinese stock market.* Physica A: Statistical Mechanics and its Applications, 2019. **523**: p. 622-630.

13.Billio, M., et al., *Econometric Measures of Connectedness and Systemic Risk in the Finance and Insurance Sectors.* Social Science Electronic Publishing, 2012. **104**(3): p. 535-559.

14.Vyrost, T., Lyócsa, and E. Baum?Hl, *Granger Causality Stock Market Networks: Temporal Proximity and Preferential Attachment.* Physica A Statistical Mechanics & Its Applications, 2015. **427**: p. 262-276.

15.Song, et al., *Time-varying causal network of the Korean financial system based on firm-specific risk premiums.* Physica, A. Statistical mechanics and its applications, 2016.

16.Gong, X.L., et al., *Financial systemic risk measurement based on causal network connectedness analysis.* International Review of Economics & Finance, 2019. **64**.

17.Memon, B.A., H. Yao, and R. Tahir, *General election effect on the network topology of Pakistan's stock market: network-based study of a political event.* Financial Innovation, 2020. **6**.

18.Gao, X., et al., *Modelling cointegration and Granger causality network to detect long-term equilibrium and diffusion paths in the financial system.* Royal Society Open Science, 2018. **5**(3): p. 172092.

19.Wang, Z., et al., *Identifying influential nodes based on fluctuation conduction network model.* Physica A: Statistical Mechanics and its Applications, 2019. **514**: p. 355-369.

20.Bu, H., W. Tang, and J. Wu, *Time-varying comovement and changes of comovement structure in the Chinese stock market: A causal network method.* Economic Modelling, 2019.

21.Liu, J.G., Z.M. Ren, and Q. Guo, *Ranking the spreading influence in complex networks.* Physica A Statistical Mechanics & Its Applications, 2013. **392**(18): p. 4154-4159.

22.Changjiang, et al., *Fast ranking influential nodes in complex networks using a k-shell iteration factor.* Physica, A. Statistical mechanics and its applications, 2016.

23.Wei, et al., *Identifying influential spreaders by weight degree centrality in complex networks.* Chaos Solitons & Fractals Applications in Science & Engineering An Interdisciplinary Journal of Nonlinear Science, 2016.

24.Lü, L., et al., *The H-index of a network node and its relation to degree and coreness.* Nature Communications, 2016. **7**: p. 10168.

25.Qian and Jun, *Identifying and ranking influential spreaders in complex networks with consideration of spreading probability.* Physica, A. Statistical mechanics and its applications, 2017. **465**: p. 312-330.

26.Li, C., et al., *Identification of influential spreaders based on classified neighbors in real-world complex networks.* Applied Mathematics & Computation, 2017. **320**: p. 512-523.

27.Lü, L., et al., *Vital nodes identification in complex networks.* Physics Reports, 2016. **650**: p. 1-63.

28.Zhou, et al., *Overlapping influence inspires the selection of multiple spreaders in complex networks.* Physica, A. Statistical mechanics and its applications, 2018.

29.Granger, C., et al., *Investigationg causal relations by econometric models: cross spectral methods.* 1969.

30.Phillips, P.C.B. and P. Pierre, *Testing for a Unit Root in Time Series Regression.* Biometrika, 1988(2): p. 335-346.

31.Granger, C., *Testing for causality : A personal viewpoint.* Journal of Economic Dynamics and Control, 2001. **2**(1): p. 329-352.

32.ZHOU, M., et al., *A novel method to identify multiple influential nodes in complex networks.* SCIENTIA SINICA Informationis, 2019. **49**(10): p. 1333-1342.