Iranian Journal of Finance

Print ISSN 2676-6337
Online ISSN 2676-6345

The Network Analysis of Tehran Stock Exchange using Minimum Spanning Tree and Hierarchical Clustering

Salman Abbasian-Naghneh
Associate Prof., Department of Financial Management, Masjed-Soleiman Branch, Islamic Azad University, Masjed-Soleiman, Iran. (Email: salman_abasian@yahoo.com)
Reza Tehrani *
*Corresponding author, Prof., Faculty of Management, University of Tehran, Tehran, Iran. (Email: rtehrani@ut.ac.ir)
Mohammad Tamimi
Associate Prof.,Department ofAccounting, Dezful Branch, Islamic Azad University, Dezful, Iran. (Email: tamimi.mohammad@gmail.com)

Document Type: Original Article 2020, Vol. 4, No. 2. 1-18
Received: 2019/08/09 Accepted: 2020/09/22 Published: 2020/10/13

Abstract

Nowadays, financial markets in Iran have attracted the attention of many managers, investors and financial policymakers. Therefore, in order to make the optimal decision and reduce the risks in such a market, it is important to identify and analyze the network behavior of the financial markets at different times to obtain the optimal decision. The current study aims to answer the following research question; how is it possible to use the minimum spanning tree and hierarchical clustering in the network analysis of the Tehran Stock Exchange? The period examined was 2013 to 2018. The population consisted of all the companies accepted in Tehran Stock Exchange. The sampling was selected purposefully and contained the companies which had at least one trading day in the time span from the beginning of 2013 to the end of 2018. The stock of the investigated companies was considered as the vertexes of one graph and the coherent information criterion was considered as the weight of the edge. First, the minimum spanning tree of the graph was calculated. The results revealed that the stocks of DarooAbuReihan, DarooPakhsh and Alborzdaroo had a high influence on directing the prices of the other stocks. Furthermore, the results of hierarchical clustering classified the stocks of the companies into 8 clusters. This study presents a viewpoint about the modern method designed for the analysis of complex financial networks. Moreover, the study offers an analysis of Iran's stock market structure which can be the center of finance researchers and analysts' attention.

Keywords: Network behavior of stock; Minimum spanning tree; Hierarchical clustering; Behavioral similarity.

DOI: 10.22034/ijf.2020.193709.1042 Publisher: Iran Finance Association

Introduction

As for human intervention, it is believed that financial markets make complicated comparative systems. Since the economy has not a theory to explain the behavior of financial markets, this discipline is left with the assumption that prices are progressing randomly which is recognized as an efficient market (Tubin, 1969). In this pattern, the evolution of stock price is only explained with the random hypothesis. Then, it is natural to ask whether these random processes are discreet or continuous and based on some basic reasons. The researches which discuss the correlation between financial instruments respond to question (Samuelson, 2016).

Regardless of interdependences in financial markets which were concentrated on financial affairs, currently, the use of complicated networks in these investigations in literature is new (Foris & Rigion, 2002). For the first time, Chai et al (2010) presented these complicated networks to the research community of financial markets that analyzed stock of America in 2010. Tumino and Mangana (2010) studied hierarchical networks in financial markets. The literature on econometrics neglected networks that explain complications of markets based on available data. Instead, it considers more traditional economic subjects like playing on networks (Charnes et al., 2014).

System complexity is a high volume of information in a system which can challenge the extract of system's suitable and basic information. Accordingly, the approaches are required that while preserving basic relations, make the system simple and minimize complexity. One of these approaches is to analyze network by Graph theory. Benano et al. (2003) stated that network analysis shows those specifications of the market which simulation and agent methods can not reveal. Thus, in this research, by network analysis, the behavior of company' stock accepted in Tehran stock exchange was discussed and in this way, the relationships and behaviors in this system are explained.

The problem which is raised in this stud is that in Tehran stock exchange, the stock of many companies is sold and purchased. Price of the stock is announced as daily for different companies, the question is that if stock of companies is considered as a network which is located in Iran stock system, and economic and political system influence on stocks, how is behavior of stocks in different periods? Price movements of which stock have a high effect on other price movements? From the behavioral similarity of prices, how to classify them? Drawing upon these questions, this study attempts to answer the following research question:

- How is the analysis of behavior network of companies accepted in Tehran stock exchange by using a minimum spanning tree and hierarchical analysis?

Background

A graph is an ordered pair G network (V, E). Members of V are named as headers G and members E and edges' G. Graphs are used in mathematics and computer sciences. There are many structures which are displayed by graph. Minimum spanning tree (to weight graph) is a tree that among all spanning trees, among all the spanning trees of that graph, the total weight of its edges has the lowest possible value (Bolubas, 1979). Minimum spanning tree is the lowest cost tree in problems to create a network that should be paid to establish a link between each of its members and we want to make a relation between two members. For example, suppose that, in a country, we want to construct a road between cities and the cost of road construction is specified. In order to find the lowest cost method, we shall calculate the minimum spanning tree. Since minimum spanning tree leaves the network from the complexity, while maintaining the basic relations of indexes, by calculating the minimum possible distance between indices, it is used in stock market analysis, currency and oil market.

Kulti 2016 compared the minimum spanning tree of Italian markets. In this research, by using four different methods, a network of 100 Italian companies during 2001-2011 was created. The obtained minimum spanning trees were compared for methods and sections in industry. Young et al., (2014) calculated correlation coefficients of stock based on maximum spanning tree. In this research, data pertaining to the daily finalized price of 268 stocks for S*P500, 221 stocks from London stock exchange and 148 stocks from Shanghai were used. Ji and Fan (2016) discussed the integration of price oil of oil exportation countries by graph theory and concluded that the market moved into integration before the global crisis in 2008. Irjit and Irjit (2009) conducted research on 143 indicators from 59 countries and showed that global integration of indicators is moved into progress and it is pertinent to Europe and North America. Mc Donald et al. (2005) assembled minimum spanning tree on the equal rate of exchange (110-time series) as 1000 hours and concluded that the converged clusters based on the behavioral similarity of currency exchange rates are largely justified by the geographic location of countries (American-European-China).

In recent studies, network structure was being considered more and by studying a system and network structure, suitable information was obtained. Structure of financial markets is one of the issues which has been discussed recently and was suitable in financial markets analysis. From researches which were done are research studies on New York stock exchange (Gan, Dejari, 2015, Himo et al., 2007, Anala et al., 2003, Tuminlo et al., 2005). Moreover, network analysis was performed on the stock of other countries (Cronlo 2005, Galaska, 2011, Hoang et al., 2009, Kantara et al., 2011, Machaba Gosel 2016, Nobi et al., 2014, Tabek et al., 2010, Wilsinki et al., 2013). In these researches, the stock exchange was discussed as a network and suitable information was extracted. Abbasian-Naghneh, et al., (2019) investigated the effect of JCPOA on the network behavior analysis of Tehran Stock Exchange indexes. Yoo et al. (2015) assembled network analysis on Shanghai stock exchange and, by non-linear scales, showed that stock has similarity for price behavior. Am et al. (2009) discussed specifications of topological methods for stock exchange in terms of minimum spanning tree and considered random matrix theory in financial time series. Raei et al. (2019) analyzed the collective behavior of Iran banking sector by random matrix theory. Namaki et al. (2011) had a network analysis of a financial market based on genuine correlation and threshold method. Namaki et al. (2019) analyzed of Iran banking sector by multi-layer approach.

In data mining and statistics, hierarchical clustering (also, hierarchical cluster analysis) is a clustering method to construct clusters. Each level of the hierarchy represents a category of data which is seen as a tree. Each tree leaf represents an initial observation and the root of the tree is the set of all observations. The results of a hierarchical clustering generally appear in the form of a dendrogram. Mantega (1999) considered a hierarchical analysis of the stock in the financial market. Tomnlo et al. (2010) suggested a method to obtain hierarchal clustering which used network structure of stock and correlation matrix in the stock exchange. Batacharji et al. (2017) discussed the dynamism of financial markets and connectedness for Asian capital markets. Kantara et al. (2011) compared hierarchical structure analysis for Turkish foreign commerce by using real exportation and importation price. Ghanbari et al. (2010, 2011) and Jahanshahloo et al. (2011a, 2011b) used hierarchical clustering in their research.

Research Methodology

Research statistical population consists of all companies accepted in Tehran stock exchange. Sampling method was purposeful. The companies were selected for this research which, from the beginning of 2013 to the end of 2018, had at least one trading day. The number of companies which had this characteristic was 135. The network structure of 135 stocks is considered as a complete graph. Vertices of this graph are 135 stocks of companies accepted in Tehran stock exchange and there is one edge between every two vertices. The edge weight between index i and index j was the same mutual information criterion between the monthly return rate of the two stocks i and j. The complete graph was obtained with 135 vertices which include 9045 edges. Such a network with complete communication and this volume of complexity does not provide useful information.

Therefore, using the minimum spanning tree, firstly, simplifies the system and, secondly, allows others to understand the most important relationships (joint behavior of companies) in the system. Minimum spanning tree of a graph is a tree that contains all vertices so that the total weight on the edges is minimized. There are two main algorithms to calculate the spanning tree; Kruskal and Prime. In this article, the Kruskal algorithm, whose details were explained by Dabroski and Polka (1998) was used.

By using minimum spanning tree and degree, betweenness centrality and farness of stock are used to calculate main stocks. The important point is that, compared to a degree, stock betweenness centrality and farness of stock are so important because they considered the weight of edges. How to calculate degree, betweenness centrality and farness of stock are as follows:

***i.* Index degree**

The degree of an index in the MST is the number of edges in that stock. Therefore, K(i), the index degree of i^{th}, is calculated by the formula of $K(i) = \sum_{j=1}^{n} a_{ij}$. If the stocks i and j are interconnected by an edge in MST, $a_{ij} = 1$ and otherwise $a_{ij} = 0$. n is the number of stocks in the MST. When the vertex degree for a stock becomes high, it indicates the importance and strength of the stock in directing other stocks.

***ii.* Betweenness centrality of stock**

Sieczka and Hołyst (2009) identified betweenness centrality as an important and useful measurement for determining the centrality of the stock. This index is calculated as follows:

$$B(i) = \frac{2}{N(N-1)} \sum_{(j,l)} \frac{\sigma_{jl}(i)}{\sigma_{jl}} \quad i \neq j \neq l \qquad (1)$$

Where $B(i)$ is the betweenness centrality of the i^{th} index, $\sigma_{jl}(i)$ is the number of shortest paths from j to l passing through i and σ_{jl} is the number of shortest paths from j to l. The betweenness centrality for constructed MST will be computed. Since the shortest path in the MST for each pair of vertices is unique, the value of $\frac{\sigma_{jl}(i)}{\sigma_{jl}}$ is zero if the transient path from j to l does not pass through i and it is one if it passes through i. $B(i)$ reflects the amount of degree that other stocks rely on i stock. The higher value of $B(i)$ represents the higher betweenness centrality.

iii. Stock farness

Sabidussi (1966) defines the notion of stock farness i^{th} as the sum of the distances of that stock to all other stocks. Therefore, $Farness = \sum_{(i,j)} R_{ij} \;\; i \neq j$.

Where R_{ij} is the shortest distance from i to j in the MST. The lower amount for i^{th} stock farness indicates that the degree of stock centrality i is higher in all stocks and the stock i is closer to other stocks.

According to Hughes (2004), the minimum spanning tree and ultrametric space are equal. It means that there is a tree that corresponds to each ultrametric and exactly there is an ultrametric space for each tree. With the help of obtained ultrametric by the MST, a hierarchical cluster can be constructed to reveal the internal cluster relationship of stocks and the near-dispersed connections between them. Hierarchical cluster helps to achieve the overall structure and main risks of the market.

In this research, in order to discuss the linear and non-linear correlation between two stocks, mutual information criterion was used. Mutual information criterion is a derived concept from Shanon entropy which means uncertainty measurement is a random variable. Entropy can be extended for randomly variable. Accordingly, if x, y are mutual information by torque as p(x, y), therefore,

$$H(X,Y) = -\sum_{x,y} p(x,y) Ln\, p(x,y) \tag{2}$$

Where H(x,y) is entropy between two random variables of X, Y. For two random variables, x and y are mutual information criterion

$$I(X,Y) = \sum_{x,y} p(x,y) Ln\, \left(\frac{p(x,y)}{p(x)p(\,y)}\right) \tag{3}$$

Is defined, in which $p(x)$ is density function for x and $p(y)$ is marginal density function (Fertzel & Pomp 2007). If variables are continuous, this index is defined as follows:

$$I(X,Y) = \iint p(x,y) log \left(\frac{p(x,y)}{p(x)p(y)}\right) dxdy, \tag{4}$$

Two independent random variables x and y don't have shared information. In fact, knowing one does not provide information about another. One can easily see that for two independent variables, mutual information criterion is equal zero, because $p(x,y) = p(x)p(y)$ is normal correlation coefficient to describe the behavior of two random variables which have a linear correlation. If mutual information criterion can estimate the non-linear correlation of two random variables, it can be shown that:

$$I(X,Y) = H(X) + H(Y) - H(X,Y) \tag{5}$$

$$I(X,Y) = H(X,Y) - H(Y|X) - H(X|Y) \tag{6}$$

Where $H(X|Y)$ is conditional entropy of x on y. To convert mutual information criterion to metric index, relation $d(X,Y) = H(X,Y) - I(X,Y)$ is used.

Thus, $d(X,Y) = H(X|Y) + H(Y|X)$ is proved, and therefore, the less the distance between two random variables, the more the predictability of one of the variables with knowing the position of another variable (Beghli, 2006).

Research Findings

To calculate the minimum spanning tree, the monthly return was calculated for the period 2013 to 2018 for each stock. 72 monthly efficiencies for each stock were calculated for the specified period. Then, 135 stocks were considered as 135 vertices of a complete graph and the mutual information criterion for both vertices in the network was considered as their connecting edge weights. Minimum spanning tree was extracted using Kruskal algorithm. Minimum spanning tree is shown in Fig. 1 and its vertices and edges are revealed in Table 1.

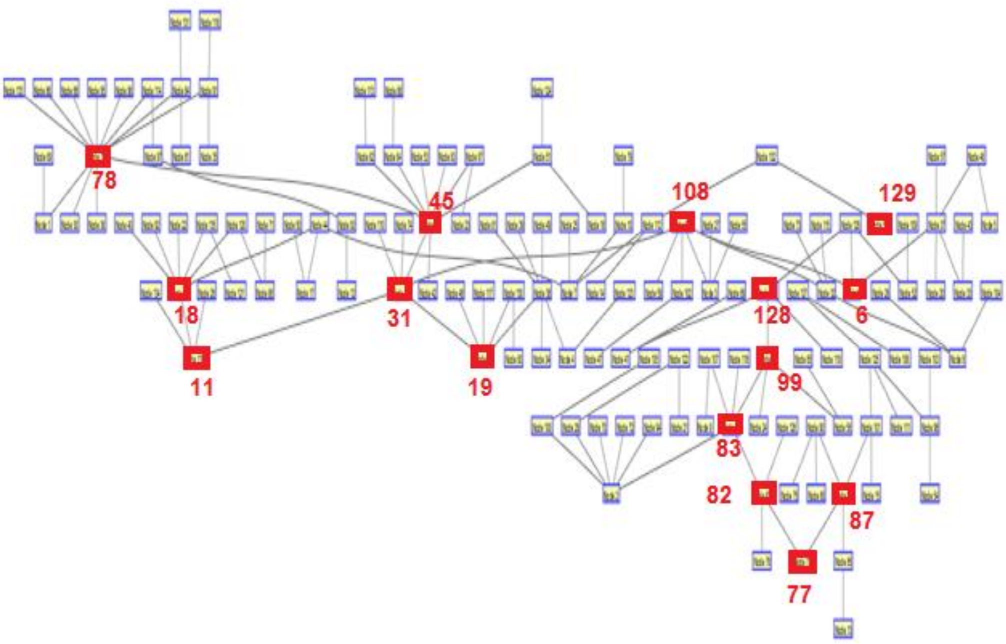

Figure 1. The minimum spanning tree of the examined network

The created minimum spanning tree has vertices, edges and d_{ij} weights, as shown in Table 1. The weight d_{ij} shows the mutual information criterion between the two vertices. This criterion shows the linear and nonlinear correlation between two different stocks.

Table 1. Edges and their weights in the minimum spanning tree

Edge	weight	Edge	weight	Edge	weight	Edge	weight
(69, 1)	2.4426	(93, 12)	2.2932	(115, 32)	2.3742	(130, 60)	2.3902
(78, 1)	2.2204	(85, 13)	2.2948	(37, 33)	2.3390	(113, 62)	2.4968
(29, 2)	2.3228	(112, 14)	2.4778	(43, 33)	2.5011	(66, 64)	2.4224
(70, 2)	2.3638	(101, 15)	2.3819	(38, 34)	2.2739	(71, 68)	2.2566
(72, 2)	2.3809	(51, 16)	2.3982	(91, 35)	2.1568	(120, 68)	2.3303
(83, 2)	2.3122	(44, 17)	2.2602	(78, 36)	2.1374	(90, 75)	2.1831
(94, 2)	2.3055	(50, 17)	2.3382	(46, 37)	2.2893	(82, 76)	2.2952
(100, 2)	2.2705	(22, 18)	2.2888	(57, 37)	2.4010	(82, 77)	2.2806
(27, 3)	2.3911	(40, 18)	2.2801	(39, 38)	2.4223	(87, 77)	2.1153
(55, 3)	2.3811	(44, 18)	2.3069	(49, 38)	2.2736	(84, 78)	2.3277

(108, 3)	2.2894	(92, 18)	2.3727	(61, 38)	2.2482	(86, 78)	2.1939
(38, 4)	2.1251	(120, 18)	2.3411	(89, 41)	2.3801	(88, 78)	2.3285
(123, 4)	2.0177	(135, 18)	2.2469	(126, 41)	2.3505	(91, 78)	2.2518
(46, 5)	2.3886	(31, 19)	2.1953	(45, 42)	2.4272	(95, 78)	2.2654
(37, 6)	2.3520	(38, 19)	2.1275	(51, 45)	2.2997	(96, 78)	2.3280
(108, 6)	1.9793	(48, 19)	2.3180	(53, 45)	2.4504	(114, 78)	2.3015
(129, 6)	2.0936	(117, 19)	2.3289	(62, 45)	2.3240	(133, 78)	2.3745
(10, 7)	2.2872	(130, 19)	2.2871	(63, 45)	2.3544	(90, 80)	2.3325
(25, 7)	2.4381	(37, 20)	2.4231	(64, 45)	2.3181	(84, 81)	2.2124
(44, 7)	2.2961	(122, 21)	2.3102	(67, 45)	2.2617	(83, 82)	2.1215
(112, 7)	2.3119	(67, 23)	2.3506	(78, 45)	2.2244	(128, 82)	2.3924
(26, 8)	2.4344	(99, 24)	2.3233	(102, 47)	2.3261	(99, 83)	2.2884
(32, 8)	2.4570	(122, 29)	2.2945	(124, 51)	2.4111	(107, 83)	2.2801
(104, 8)	2.5642	(78, 30)	2.2315	(109, 52)	2.2866	(118, 83)	2.4315
(107, 9)	2.3497	(45, 31)	2.2179	(129, 52)	2.3412	(131, 84)	2.3292
(79, 10)	2.4308	(74, 31)	2.2735	(98, 54)	2.3521	(87, 85)	2.3675
(18, 11)	2.2859	(108, 31)	2.1885	(132, 56)	2.2762	(90, 87)	2.2578
(28, 11)	2.2977	(110, 31)	2.2518	(65, 58)	2.2551	(101, 87)	2.3916
(31, 11)	2.2816H	(73, 32)	2.3240	(99, 58)	2.3063	(116, 91)	2.2836
(134, 11)	2.3456	(108, 32)	2.1415	(108, 59)	2.3735	(97, 93)	2.3328
(114, 97)	2.2592	(105, 100)	2.4747	(125, 111)	2.2645	(127, 125)	2.3649
(103, 98)	2.5137	(125, 101)	2.3937	(132, 112)	2.3525	(129, 126)	2.2188
(125, 98)	2.3651	(108, 102)	2.2916	(126, 119)	2.4303		
(126, 99)	2.2333	(127, 106)	2.3500	(135, 121)	2.4509		

The vertices of the minimum spanning tree are the same 135 stock exchanges in Tehran which are connected by the edges shown in Table 1. Moreover, the weight of each edge which represents the mutual information criterion between the two stocks is given. For example, the edge of (1, 69) shows that there is an irrefutable relationship between vertex one (A.S.P. stock) and the edge of 69 (Tehran investment housing stock). The intensity of this relation (the weight of the edge) is equal to 2,446.2.

Degree centrality, betweenness centrality, and eccentricity are used to examine the strength of different stock in the orientation of the stocks of other companies. The results are presented in Table 2.

Table 2. Indicators' power criteria

Stock NO.	Stock Name	Betweenness centrality	Degree	Farness
31	DarooAbuReihan	1.337313	6	1445.32
108	Daroopakhsh companies	1.134771	6	1460.639
6	Alborzdaroo	0.995246	3	1525.955
129	DarooPakhsh	0.872084	3	1628.542
126	Magsal	0.84356	4	1750.579
45	Salmin	0.78806	9	1598.352
99	Sanati Minoo	0.787839	4	1891.277
83	Siman Shargh	0.762852	5	2053.751
11	Iran Daroo	0.553897	4	1648.379
78	Siman Khash	0.518076	12	1809.67
18	Pars Khodro	0.507905	7	1865.536
82	Siman Sepahan	0.468988	4	2259.538
77	Siman Tehran	0.394693	2	2494.445
87	Siman Fars va Khuzestan	0.382753	4	2716.554
19	Pars Daroo	0.306468	5	1689

Table 2 lists 15 companies that have the highest degree of centrality and betweenness centrality as well as the least eccentricity criterion. The shares of DarooAbuReihan, Daroopakhsh and Alborzdaroo companies had the most betweenness centrality, respectively, in terms of the criterion of betweenness centrality. In terms of eccentricity, the shares of DarooAbuReihan, Daroopakhsh and Alborzdaroo companies had the least eccentricity. As seen in the period of research, the shares of active companies in the field of medicine had a great influence on the market according to two criteria of eccentricity and betweenness centrality. In other words, the market has relied on the stock of pharmaceutical companies, and the changes in their prices directed the market and identified the prospects for future market developments. Furthermore, several companies in the field of cement, including Salmin Company, Industrial Mino in the field of food and Magsal companies in the field of the industry have played a major role in the market and were reliance for other stocks. In terms of degree centrality, the shares of Siman Khash, Siman Salmin, Pars Khodro, DarooAbuReihan and Daroopakhsh companies had the closest relationship with other companies. The location of the fifteen companies listed in Table 2 is shown in the minimum spanning tree chart indicated in red squares.

After calculating the meta-metrics distances between the vertices of the minimum spanning tree, the clustering of stocks based on behavioral similarity and their hierarchical cluster representation is presented in Fig. 2.

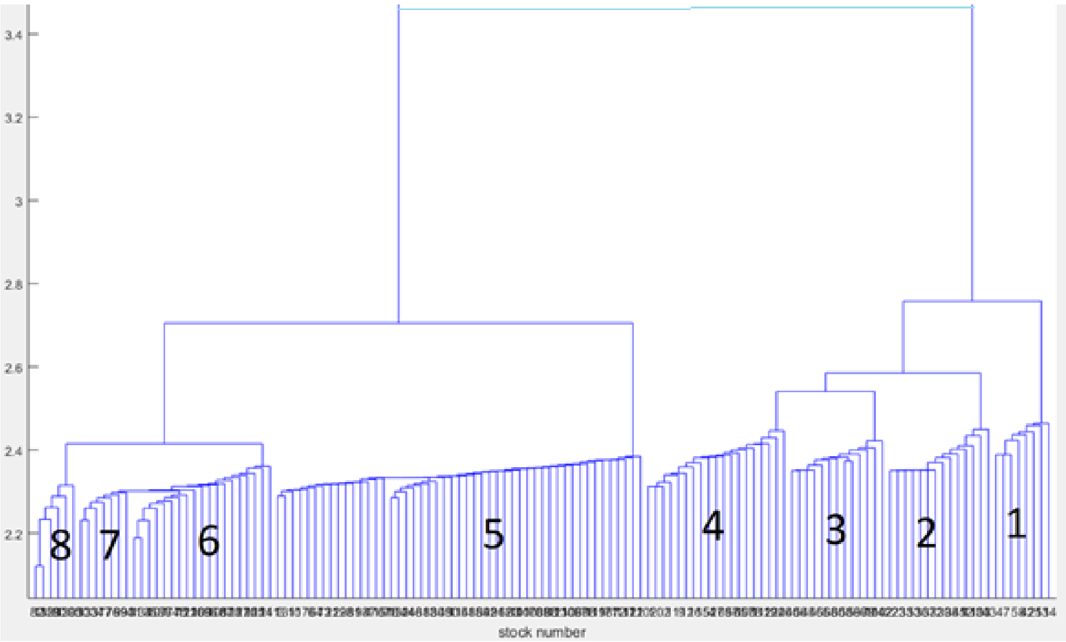

Figure 2. Hierarchical clustering based on behavioral similarity

Distinctive clusters in hierarchical clustering include:

Group 1: Bank Tejarat, Bama, Pegah Esfahan, Rayan Saipa, Ama, Afranet, Ertebatat Sayar

Group 2: Rahshad Sepahan, Foulad Khouzestan, Goltash, Sar Bahman, Derakhshan Tehran, Daroo Farabi, Daroo Loghman, Daroo Osveh, Daroo Sobhan, Daroo Zahravi, Sar BouAli, Daroo Amin, Ghand Hegmatan, Pars Minoo

Group 3: Ferrosilis Iran, Ghand Sabet Khorasan, Sar Ghadir, Sar Toseae Shomal, Sar Toseae Azerbaijan, Sar Toseae Sanati Iran, Sar Maskan, Sar Petroshimi, Sobhan Daroo, Salmin, Sar Pardis, Kowsar Pharmaceuticals, Sanati Mino

Group 4: Takin, Pashm Shisheh Iran, Pars Daroo, Bahman Group, Siman Khash, Nishapour Sugar, Sar Sanat & Madan, Sar Tos Gostar, Sar Tehran Maskan, Jam Daroo, Sar Pars Toshe, Karafarin Bank, Behsaram, Airka Part Sanat, Iran Khodro, Iran Daroo, Abadgaran, Pars Ceram, A.S.P.

Group 5: Leasing Khodro, Leasing Sanat & Madan, Pipes and Machine-building, Sar Naft, Sar Meli, Leasing Iran, Leasing Iranian, Sanati RavAn Fan Avar, Butane Industrial, Maroon Petrochemical, Sina Kashi, Mokhaberat Iran, Siman Kurdistan, Siman Kerman, Sar

Melat, Omran & Toseae Fars, Fanavari Mavad-e-Madani, Sar Shahed, Sar bank Refah, Sina Daroo, Siman Sufiyan, Siman West, Sar Iran Khodro, Glucozan, Sar Omid, Sar Alborz, Siman Darab, Siman East, Sar Saman Gostar Isfahan, Sar Etebar Iran, Zamyad, Naft Behran, Pars Khodro, Behshahr SarGorooh, Lamiran, Sepanta, Razi Pharmaceutical Glass, Siman Dorood, Shomal Hafari, Khak Chini Iran, Pars Switch, Sar Noor Kowsar Iranian, Sar Sandogh Pazneshashtegi, Alborz Daroo, Parsian Insurance Company, Iran Khodro, Motogen, Azarab

Group 6: Calciuminum, Bahonar Mess, Rooy maaden, Magsal, Meli Sorb & Rooy, Meli Sanaye Mes Iran, Fars and Siman Khuzestan, Navard-e-Ghataat-e-Foladi, Daroopakhsh factories, Industrial Barez, Alvand Kashi, Leabiran, Niromohareke Mashinsazi, Siman Urmia, Sar Niroo, Siman Fars Nou, NirooMohareke, Droo Abu-Reihan

Group 7: Hamedan Shishe, Siman Behbahan, Siman Tehran, Daroo Jaber-ibn-Hayan, Informatics Khadamat, Siman Caspian, Daroo Razak

Group 8: Siman Sepahan, Niroo Trans, Mavad-e-Drugpakhsh, Siman Hegmatan, Motorsazan Tractor, Kermanshah Petrochemical Industries

Error calculation

In this paper, there are error display numbers and computational error. In order to minimize the error, the "format long" command in MATLAB was used, which in the scientific display of numbers, considers Mantis to mean up to 16 digits. For example, the weight of the edges in Table 1 was calculated 15 decimal places, but in this article up to 4 decimal digits was given.

And if the cut-off method is used to display numbers, the error rate will be at most 10^{-n+1}. So the maximum error for calculating the weight of the edges will be 10^{-15}, which indicates the double-precision of the numbers' presentations.

It is also necessary to check the computational error. In the calculations, the formula $log\left(\frac{p(x,y)}{p(x)p(y)}\right)$ has been calculated and it is necessary to check its error. We have a logarithmic relationship as follows.

$$log\left(\frac{p(x,y)}{p(x)p(y)}\right) = \log p(x,y) - \log(p(x)p(y)) = \log p(x,y) - (\log p(x) + \log p(y)) \tag{7}$$

So it is enough to check the error for logarithmic addition and subtraction.

If $\widetilde{a}$ is cut a to n the decimal digit and $\widetilde{b}$ is decomposed b to n a decimal number, we have:

$X = \tilde{A} - \tilde{B}$ in which $\tilde{A} = \log \tilde{a}$ and $\tilde{B} = \log \tilde{b}$. Therefore,

$e(\tilde{A} \pm \tilde{B}) \leq e(\tilde{A}) + e(\tilde{B})$

in which,

$$e(\tilde{A}) = A - \tilde{A} = \log a - \log \tilde{a} = \log \frac{a}{\tilde{a}}$$

$$e(\tilde{B}) = B - \tilde{B} = \log b - \log \tilde{b} = \log \frac{b}{\tilde{b}}$$

Regarding that for $x > 0$, $\ln x \leq x - 1$, and if $x \geq 1$, $|\ln x| \leq |x - 1|$, and if $\tilde{a}$ is the cut a, we have $\frac{a}{\tilde{a}} \geq 1$

And if $\tilde{A}$ is cut A to n the decimal digit, we have: $\left|\frac{A-\tilde{A}}{A}\right| \leq 10^{-n+1}$

Then,

$$|e(\tilde{A})| = \left|\log \frac{a}{\tilde{a}}\right| \leq \left|\frac{a}{\tilde{a}} - 1\right| \leq 10^{-n+1}$$

$$|e(\tilde{B})| = \left|\log \frac{b}{\tilde{b}}\right| \leq \left|\frac{b}{\tilde{b}} - 1\right| \leq 10^{-n+1}$$

$$|e(\tilde{A} \pm \tilde{B})| \leq |e(\tilde{A})| + |e(\tilde{B})| \leq 10^{-n+1} + 10^{-n+1} = 2 \times 10^{-n+1}$$

Therefore, the maximum logarithmic addition and subtraction error is 2×10^{-15}.

Conclusion

This study aims to analyze the network of Tehran Stock Exchange using a minimum spanning tree and hierarchical clustering. The vertices of the network indicated the shares of companies and the weight of each edge represented the mutual information criterion of two vertices in two edges. This article sought to examine the combined behavior of stock companies in different time periods. For this purpose, a metric was defined on the network edges using the mutual information criterion and, then, the minimum spanning tree and hierarchical clustering were calculated to reduce network complexity. In the following, the degree centrality, betweenness centrality and eccentricity on the minimum spanning tree were calculated to identify the powerful and influential market shares in the market orientation. The results showed that, according to the two criteria of eccentricity and betweenness centrality, the shares of the companies of Daroo AbuReihan, Daroopaksh and AlborzDaroo factories have had a great influence on the market in the examined period, respectively. Also, 135 stocks were divided into 8 different clusters based on the behavioral similarity that is measured by meta-metrics.

An unexpected result was that two clusters with long-distance (non-similar behavior) were observed in the dendrogram of Fig. 2. The first cluster consists of the first, second, third and fourth clusters, and the second cluster comprises the fifth to eighth clusters. Given the stock of various clusters, pharmaceutical companies, mass production companies, food companies and banks were in the first cluster. These companies are in critical parts of the human needs, namely housing, health, food and (unrealistic) monetary economics, i.e., the banking sector. In the second cluster, there were the real sectors of the economics i.e. the companies of the metal and non-metallic minerals, especially cement, industrial machinery, technical engineering, petrochemicals, oil and leasing companies. Apparently, the clusters divided the market into two parts. In the periods and especially during the downturn of the industry, the first part is flourishing and the market is focused on this, and in a period when the real sector of the economy is active, the market focuses on the second cluster. On the other hand, the time during 2013-2018 can be divided into two parts. Another unexpected result was that three active companies in the field of medicine were introduced as the most influential companies in the orientation of shares of other companies by calculating the criteria for betweenness centrality and eccentricity.

The findings of this research have several contributions to the literature. First, a network was constructed whose vertices were shares of the companies admitted to the Tehran Stock Exchange and the weight of each edge represents the mutual information criterion of two vertices in two edges. Second, the minimum spanning tree of stocks of companies was calculated. According to the degree centrality, betweenness centrality and eccentricity, the influential stocks that were active in the field of medicine were identified. Thirdly, hierarchical clustering was done with respect to the minimum spanning tree and meta-metrics distances. The result of this hierarchical clustering is a classification of 135 stocks into 8 different clusters in which the shares have behavioral similarity in terms of price.

Finally, a number of important limitations need to be considered. First, in this paper, the combined behavior of stock companies has been investigated in recent years. If the period under study changes, the minimum spanning tree, affective stocks and stock clustering will also change. Secondly, if a larger time period is chosen to be examined, affective stocks might not be properly identified, and, on the other hand, it is more difficult to interpret the behavior of the indicators by attributing that behavior to the events of that period. Future researchers should focus their research attention on the network behavior of

Tehran Stock Exchange indicators, foreign currencies as well as the oil field.

This research has thrown up many questions in need of further investigation. Investors are encouraged to pay close attention to the network behavior of stock companies. By identifying the influential shares as well as the stock clustering structure, at any given time, the behavior of the stock of influential stocks can be monitored and concerning their price changes, one can predict the price behavior of other stocks related to that cluster.

References

Abbasian-Naghneh, S., Tehrani, R., & Tamimi, M. (2019). The Effect of JCPOA on the Network Behavior Analysis of Tehran Stock Exchange Indexes. *Advances in Mathematical Finance and Applications,* (Articles in Press).

Baghli, M. (2006). A model-free characterization of causality. *Economics Letters*, *91*(3), 380-388.

Bhattacharjee, B., Shafi, M., & Acharjee, A. (2017). Network mining based elucidation of the dynamics of cross-market clustering and connectedness in the Asian region: An MST and hierarchical clustering approach. *Journal of King Saud University-Computer and Information Sciences*.

Bollobas, B. (2012). Graph theory: an introductory course, (Vol. 63). *Springer Science & Business Media*.

Bonanno, G., Caldarelli, G., Lillo, F., & Mantegna, R. N. (2003). The topology of correlation-based minimal spanning trees in real and model markets. *Physical Review E*, *68*(4), 046130.

Bonanno, G., Caldarelli, G., Lillo, F., & Mantegna, R. N. (2003). The topology of correlation-based minimal spanning trees in real and model markets. *Physical Review E*, *68*(4), 046130.

Charness, G., Feri, F., Meléndez-Jiménez, M. A., & Sutter, M. (2014). Experimental games on networks: Underpinnings of behavior and equilibrium selection. *Econometrica*, *82*(5), 1615-1670.

Chi, K. T., Liu, J., & Lau, F. C. (2010). A network perspective of the stock market. *Journal of Empirical Finance*, *17*(4), 659-667.

Coletti, P. (2016). Comparing the minimum spanning trees of the Italian stock market using returns and volumes. *Physica A: Statistical Mechanics and its Applications*, *463*, 246-261.

Coronnello, C., Tumminello, M., Lillo, F., Micciche, S., & Mantegna, R. N. (2005). Sector identification in a set of stock return time series traded at the London Stock Exchange. *arXiv preprint cond-mat/0508122*.

Dabrowski, J., & Pułka, A. (1998). TH Cormen, CE Leiserson, and RL Rivest Introduction to Algorithms. McGraw-Hill, MIT Press, 1990. Discrete Approach to PWL Analog Modeling in VHDL Environment. *Analog Integrated Circuits & Signal Processing*, *16*(2), 91-99.

Eom, C., Oh, G., Jung, W. S., Jeong, H., & Kim, S. (2009). Topological properties of stock networks based on minimal spanning tree and random matrix theory in financial time series. *Physica A: Statistical Mechanics and its Applications*, *388*(6), 900-906.

Eryiğit, M., & Eryiğit, R. (2009). The network structure of cross-correlations among the world market indices. *Physica A: Statistical Mechanics and its Applications*, *388*(17), 3551-3562.

Frenzel, S., & Pompe, B. (2007). Partial mutual information for coupling analysis of multivariate time series. *Physical review letters*, *99*(20), 204101.

Forbes, K. J., & Rigobon, R. (2002). No contagion, only interdependence: measuring stock market comovements. *The journal of Finance*, *57*(5), 2223-2261.

Gałązka, M. (2011). Characteristics of the Polish stock market correlations. *International review of financial analysis*, *20*(1), 1-5.

Gan, S. L., & Djauhari, M. A. (2015). New York Stock Exchange performance: evidence from the forest of multidimensional minimum spanning trees. *Journal of Statistical Mechanics: Theory and Experiment*, *2015*(12), P12005.

Ghanbari, A., Abbasian-Naghneh, S., & Hadavandi, E. (2011). An intelligent load forecasting expert system by the integration of ant colony optimization, genetic algorithms and fuzzy logic. In 2011 *IEEE symposium on computational intelligence and data mining* (CIDM) (pp. 246-251). IEEE.

Ghanbari, A., Hadavandi, E., & Abbasian-Naghneh, S. (2010). An intelligent ACO-SA approach for short term electricity load prediction. *In International Conference on Intelligent Computing* (pp. 623-633). Springer, Berlin, Heidelberg.

Heimo, T., Saramäki, J., Onnela, J. P., & Kaski, K. (2007). Spectral and network methods in the analysis of correlation matrices of stock returns. *Physica A: Statistical Mechanics and its Applications*, *383*(1), 147-151.

Huang, W. Q., Zhuang, X. T., & Yao, S. (2009). A network analysis of the Chinese stock market. *Physica A: Statistical Mechanics and its Applications*, *388*(14), 2956-2964.

Hughes, B. (2004). Trees and ultrametric spaces: a categorical equivalence. *Advances in Mathematics*, 189(1), 148-191.

Jahanshahloo, G. R., & Abbasian-Naghneh, S. (2011a). Data envelopment analysis with imprecise data. *Applied mathematical sciences*, *5*(61-64), 3089-3106.

Jahanshahloo, G. R., Zohrehbandian, M., & Abbasian-Naghneh, S. (2011b). Using interactive multiobjective methods to solve multiple attribute decision-making problems. *Australian Journal of Basic and Applied Sciences*, 5(9), 298-308.

Ji, Q., & Fan, Y. (2016). Evolution of the world crude oil market integration: A graph theory analysis. *Energy Economics*, 53, 90-100.

Kantar, E., Deviren, B., & Keskin, M. (2011). The hierarchical structure of Turkey's foreign trade. *Physica a: Statistical Mechanics and its Applications*, 390(20), 3454-3476.

Majapa, M., & Gossel, S. J. (2016). The topology of the South African stock market network across the 2008 financial crisis. *Physica a: Statistical Mechanics and its Applications*, 445, 35-47.

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

McDonald, M., Suleman, O., Williams, S., Howison, S., & Johnson, N. F. (2005). Detecting a currency’s dominance or dependence using foreign exchange network trees. *Physical Review E*, *72*(4), 046106.

Namaki, A., Raei, R., Asadi, N., & Hajihasani, A. (2019). Analysis of Iran Banking Sector by Multi-Layer Approach. Iranian Journal of Finance, 3(1), 73-89.

Namaki, A., Shirazi, A. H., Raei, R., & Jafari, G. R. (2011). Network analysis of a financial market based on genuine correlation and threshold method. Physica A: Statistical Mechanics and its Applications, 390(21-22), 3835-3841.

Nobi, A., Maeng, S. E., Ha, G. G., & Lee, J. W. (2014). Effects of the global financial crisis on network structure in a local stock market. *Physica a: Statistical Mechanics and its Applications*, 407, 135-143.

Onnela, J.-P., Chakraborti, A., Kaski, K., Kertesz, J., & Kanto, A. (2003). Dynamics of market correlations: Taxonomy and portfolio analysis. *Physical Review E*, 68(5), 056110.

Raei, R., Namaki, A., & Vahabi, H. (2019). Analysis of Collective Behavior of Iran Banking Sector by Random Matrix Theory. Iranian Journal of Finance, 3(4), 60-75.

Sabidussi, G. (1966). The centrality index of a graph. *Psychometrika*, 31(4), 581-603.

Sieczka, P., & Hołyst, J. A. (2009). Correlations in commodity markets. *Physica a: Statistical Mechanics and its Applications*, *388*(8), 1621-1630.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell* *system technical journal*, *27*(3), 379-423.

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

Tobin, J. (1969). A general equilibrium approach to monetary theory. *Journal of money, credit and banking*, *1*(1), 15-29.

Tumminello, M., Aste, T., Di Matteo, T., & Mantegna, R. N. (2005). A tool for filtering information in complex systems. *Proceedings of the National Academy of Sciences*, 102(30), 10421-10426.

Tumminello, M., Lillo, F., & Mantegna, R. N. (2010). Correlation, hierarchies, and networks in financial markets. *Journal of Economic Behavior & Organization*, 75(1), 40-58.

Wiliński, M., Sienkiewicz, A., Gubiec, T., Kutner, R., & Struzik, Z. (2013). Structural and topological phase transitions on the German Stock Exchange. *Physica a: Statistical Mechanics and its Applications*, 392(23), 5963-5973.

Yang, C., Zhu, X., Li, Q., Chen, Y., & Deng, Q. (2014). Research on the evolution of stock correlation based on maximal spanning trees. *Physica a: Statistical Mechanics and its Applications*, 415, 1-18.

You, T., Fiedor, P., & Hołda, A. (2015). Network analysis of the Shanghai stock exchange based on partial mutual information. *Journal of Risk and Financial Management*, 8(2), 266-284.

Bibliographic information of this paper for citing:

Abbasian-Naghneh, Salman; Tehrani, Reza & Tamimi, Mohammad (2020). Network Analysis of Tehran Stock Exchange using Minimum Spanning Tree and Hierarchical Clustering. *Iranian Journal of Finance*, 4(2), 1-18.