University of Hormozgan

International Journal of Industrial Engineering & Management Science

journal homepage: www.ijiems.com

# Portfolio Selection by Optimizing Risk and Return Based on Complex Network Analysis (Case Study: Tehran Stock Exchange)

Zainabolhoda Heshmati Rafsanjani [a,*], Farideh Rahimnezhad [b]

[a] *Assistant Professor and Faculty Member of Department of Network Science and Technology, Faculty of New Sciences and Technologies, University of Tehran, Tehran, Iran*
[b] *PhD Candidate of Information Technology Engineering, Department of Network Science and Technology, Faculty of New Sciences and Technologies, University of Tehran, Tehran, Iran*

ARTICLE INFO

*Article history:*
Received: 2021-08-14
Received in revised form: 2021-11-20
Accepted: 2021-11-29

*Keywords:*
Complex Network
Networks
Portfolio Selection
Stock Market
Stock Return Correlation Network

ABSTRACT

Nowadays, complex networks are applied for analyzing a huge body of data. Since the stock market has large data that are constantly fluctuating, it is highly difficult to analyze these data and manage the stock purchase and sale for investors. In this research, complex network is applied to select stock portfolios in order to facilitate market analysis and decision making in business relationships, and reduce the risk of inaccurate decisions. For this purpose, Tehran Stock Exchange was selected and subsequently, the latest data were collected over six consecutive years. Afterwards, a stock return correlation network was developed. According to the community detection, cohort groups were identified and then, a stock was selected from each community by designing an optimization model from the network centralities, risk and returns. Finally, for checking the accuracy of the selected portfolio, the portfolio performance in two ways with and without risk was compared with the performance of the TEPIX index. Results of this study showed that complex networks played a very effective role in selecting stock portfolios with high returns and low risk by visualizing lots of stocks in one network picture and facilitate global characteristics analysis across the network.

## 1. Introduction

In general, complex networks are new tools for analyzing a huge set of data. Since the stock market has many data that are constantly fluctuating, it is difficult to analyze these data and manage the stock purchase and sale for investors. Thus, complex networks can be used in stock market analysis (Chaharsooghi & Rahimnezhad, 2016). Nowadays, according to the expansion of the application of complex networks in economics and finance, many researchers attempt to focus on different aspects of the market, create different types of networks and check aspects of the stock market. In addition to the financial area, complex networks in other fields such as computer science, biology, IT engineering and network science have been increasingly becoming important (Motter & Albert, 2012). A network is actually a graph of edges and nodes. A set of nodes is a one-dimensional set of points (in this research, a set of companies). In contrast, edges are two-dimensional sets of lines (in this study, collections of the correlation of stock returns). In terms of the number of nodes and communications, networks can be more complex and edges in various networks may be weighted and directed (Namaki et al., 2011).

In recent years, efforts have been made to facilitate the process of investment, and stock portfolio selection models have become a tool to improve decision-making processes. The stock portfolio theory was first developed by Markowitz (1952) and later, by his disciples Sharpe and Linder. From the beginning of the 1950s, the theory became the foundation for

[*] Corresponding author.
*E-mail address:* zheshmati@ut.ac.ir

further research. However, the complexity of financial markets has motivated scholars to undertake new research. In capital markets with a large number of stocks, the search space is highly widespread that makes it impossible to use mathematical models. Hence, meta-heuristics algorithms, such as genetics and ant colony algorithms, are critical (Aranha & Iba, 2009). Furthermore, the problem of each investor is to determine a set of stocks with maximum utility; this is equivalent to selecting the optimal stock portfolio. Before Markowitz's theory, investors were mainly focused on stock returns and in addition to maximizing returns, they wanted to minimize risk (Alexander, Sharpe & Baily, 1993). Therefore, multi-purpose decision-making models are able to apply risk-return preferences by allocating weight to the recommended and practicable adaptability of the investor (Lin & Liu, 2008). However, according to the very high volume of stock market data, there is a need to present the communications in the market simultaneously in a single image, so that one can observe types of relationships in the market at a glance and decide more quickly. Hence, complex networks are introduced as a new method for macro data analysis, especially in the financial field. Indeed, initial research has been focused more on network topological properties and less on data visualization (Sun, Tian & Yang, 2015). Therefore, given the major role of complex networks in the recent studies of financial markets, in this research, we attempt to focus on:

- Visualization and formation of network-based optimization models.
- Visualization lots of stocks in one picture.
- Facilitation one of the most important issues in the stock market called portfolio selection.
- Facilitation market analysis and decision making in business relationships for managers and investors in various industries.
- Reduction the risk of inaccurate decisions in stock portfolio selection
- Maximization the returns of stock portfolio even higher than the total index.
- Develop an optimization model based on centralities to increase the accuracy of stock selection from each community.

## 2. Literature Review

So far, many studies have been carried out in real-world modeling using complex networks. Such studies have been conducted in a variety of fields, including in business networks, World Trade Network and Board of Directors. All such studies have well shown that modeling in networks is a powerful tool for understanding behaviors in complex networks (Shariatmadari, 2014).

Mantegna (1999) was the first to build networks based on the stock price coefficient. He examined a stock portfolio for calculating Dow Jones Industrial Average index and the Standard and Poor' 500 index, and found a hierarchical order of stocks created by examining the daily time series of the logarithm of the stock price. Time series is an important method in analysis and predictive validity (Mohammed A.M, et al. 2021). Lee et al. (2007) showed that Korea's stock exchange network was a free-scale network. In another study, they found that market fluctuations were associated with the minimal mass of trees (Lee, Youn & Chang, 2012).

In addition to cost-efficiency networks, there are networks built on the basis of the relationships between investors and transactions in the stock market. For example, in a research of networks, Li and Wang (2007) used a variety of HSI fluctuations and extracted hidden patterns between fluctuations based on topological nodes.

Furthermore, complex networks based on stock indices such as stock prices and price fluctuations have attracted the attention of many researchers in physics and statistics; usually in these networks, stock is the node and the correlation between stock returns is the edge (Chen et al., 2015).

### 2.1. Stock Market Concepts

- Stock: Each stock represents the smallest unit of ownership.
- Investment: Investment is the purchase of assets or securities that over time will generate income and increase value for the investor.
- Company symbol: The companies listed on the stock exchange are each classified in their own industry and have a specific name.
- Closing price: The price of the last trade in a symbol today.
- Portfolio: It is a combination of stocks or other assets that an investor buys.
- Index: It is a number by which we examine the changes made in a phenomenon over a period of time.

### 2.2. Centrality

The position of each node in the community can be determined by connecting to other nodes, which is called centrality. If an industrial network is built on the basis of commercial flows between different industries, then, companies located in the center of the network are more exposed to systematic risks than other companies (Aobdia, Caskey & Ozel, 2014). Furthermore, researchers have shown that industries with higher centrality in the network are more productive (Ahern and Jarrad, 2014).

Nie et al. (2015) reviewed the model of the complex Chinese stock market. First, based on the correlation coefficient formula, the coefficient of correlation between the stocks was calculated and a quantitative analysis was performed to select the approximate threshold for determining the numerical value of the correlation. Then, they analyzed the structure of the stock market using a centralized analysis method such as betweenness and closeness centralities. A view of their network is shown in Figure 1.

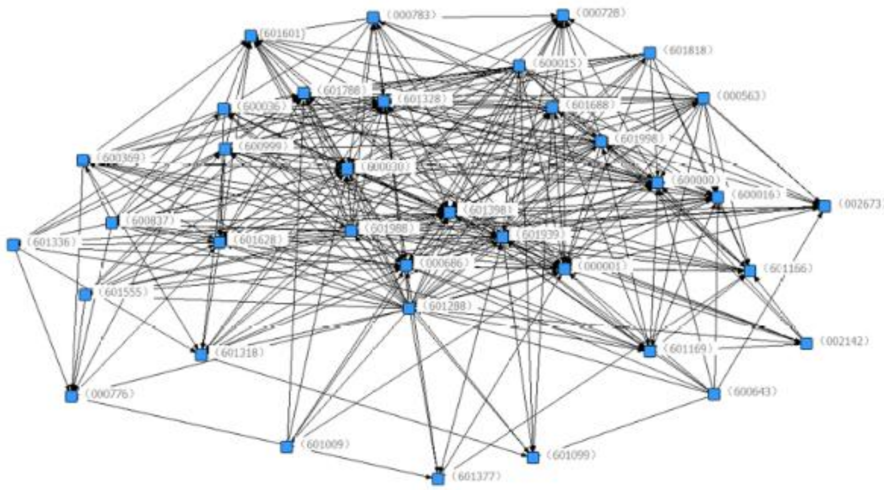


Fig. 1. The Chinese stock network (Nie et al., 2015)

### 2.3. Portfolio Selection

One of the most important issues in the stock market is how to select a portfolio. A few years ago, several researchers used the stock return correlation to identify the stock market as a complex network (Bonanno et al., 2004; Onnela et al., 2004b). Recently, Sun et al. (2015) reviewed portfolio management with network analysis and found that traditional portfolio management techniques using statistical properties showed the local behavior of stocks and could not generally display their behavior in the stock market. In another study, the Indian market stocks were clustered and the clusters were used to make a portfolio. The results showed that, via creating diversity in stock portfolios, the risk level was minimized (Nanda et al., 2010).

### 2.4. Networking Effects of Corporate Communications

In a research, the effect of communication between companies was investigated by building a directed network from the Chinese stock market. By analyzing the subnets of these companies, it became obvious that the effect of communications was quite different in one section from another. In addition, according to the centrality criterion, the ranking of the companies was also obtained and, finally, it was found that the type of communication between the companies directly affected the price fluctuations of the stocks (Gao, Zeng & Cia, 2015). Figure 2 shows the network of the companies.

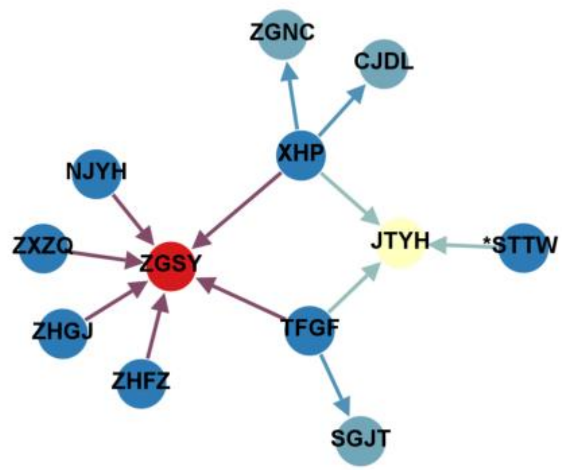


Fig. 2. The network of companies in the Chinese stock market (Gao, Zeng & Cia, 2015).

### 2.5. Effects of Network Structure on Stock Returns

Recent financial research has used network models to show the relationship between industries. Most of these networks are relatively simple, in which industries are considered as nodes and business relationships between industries are presented as edges. Chen et al. (2015) developed a network of Chinese industry stock market rebalancing during the period from 1994 to 2013, attempting to answer these questions:

- Does industrial relation in a network affect the industry's returns?
- Does the relationship of stocks in a network affect the return on stocks?

According to data from the period of twenty years, their findings showed that the stocks with less investment were more likely to be in the center of the network. They also used regression analysis in their research and found that the return on the industry was significantly and positively affected by the degree of interconnection between the industries on the network. Furthermore, stock returns were also significantly affected by stock positions in the network.

### 2.6. Effect of the Market Correlation Structure on Future Market Fluctuation Changes

A research found that previous changes in the market correlation structure were significantly associated with future changes in the market fluctuation. The researchers, applying correlation data filtering networks, created a new tool to predict the market fluctuation. They also introduced a new criterion called continuity of correlation structure, which expresses the rate of the market structure changes. The results of this study showed that this new method could be adaptable to the sudden changes in the market such as financial crises (Musmeci, Aste & Matteo, 2016).

### 2.7. Diffusion of the Financial Crisis

In another research, Sun et al. (2015) examined the American stock returns correlation network, in which nodes were stocks and edges were stocks return correlation. They first examined the community detection issue in this network and observed that there were different sections in the network and the stocks within each section had similar performance patterns. Then, they discussed the issue of the financial crisis diffusion in the network and observed that the financial crisis began with a negative-return company and gradually spread to other companies. After the visualization of the diffusion of the financial crisis, they observed some trends, as shown in Figure 3.

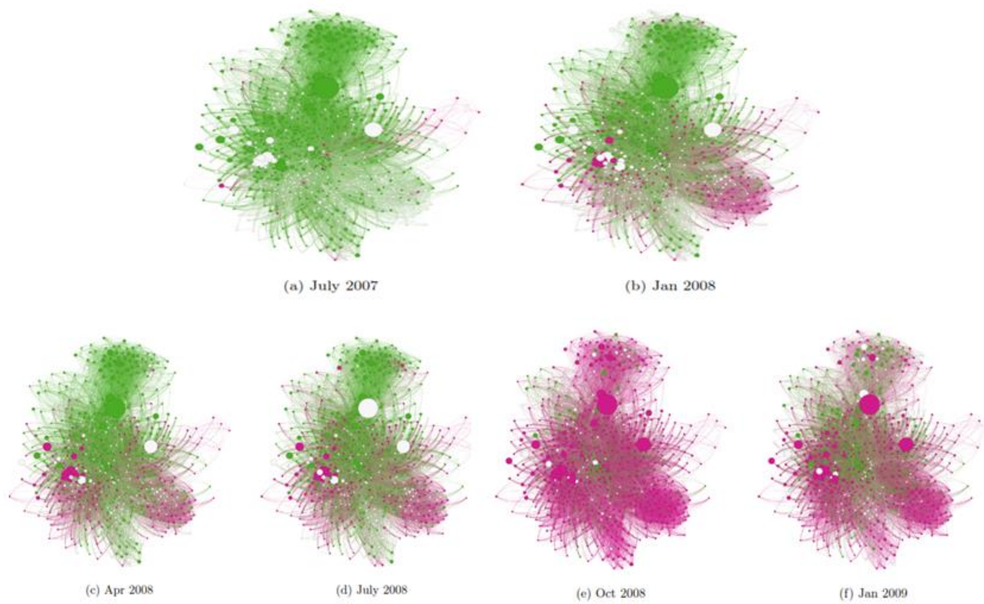


Fig. 3. Diffusion of the financial crisis (Sun, Tian & Yang, 2015)

About the financial crisis, in one of the recent research, Shi et al. (2020) Created a dynamic complex network of individual stocks, and showed the systemic risk of the market using the adjusted structural entropy. Their results represented the systemic risk of China's stock market as a whole showed a downward trend, and the periodic fluctuation of systemic risk had a long-term equilibrium relationship with the abnormal fluctuation in stock market.

## 2.8. Influential Nodes in Stock Market

In a research, Huang et al. (2021) ranked the influential nodes for Chinese stock market by complex network analysis approach and showed that the influential nodes were large-capitalization companies. The classification of all the research analyzing the stock market using complex networks is presented in a timely order in Table 1.

Table 1. Overview of the literature on stock market analysis

| Authors | Year | Methodology | Financial Market |
|---|---|---|---|
| Mantegna | 1999 | Portfolio selection by examining the daily time series of the stock price logarithm & construct a hierarchical order of stocks | United States |
| Albert, Jeong & Barabasi | 2000 | Scale free network & topology characteristic | - |
| Motter, Nishikawa & Lia | 2002 | Small world & network topology | - |
| Bonanno et al. | 2004 | Systematic risk assessment in financial markets | - |

| Authors | Year | Methodology | Financial Market |
|---|---|---|---|
| Onnela et al. | 2004 | Build a stock return correlation network | New York |
| Boginski et al. | 2005 | Develop a technical method for stock clustering and use the power distribution model | United States |
| Lee et al. | 2007 | Using scale free network | Korea |
| Li & Wang | 2007 | Extract hidden patterns among stock market fluctuations | Hong Kong |
| Huang, Zhuang & Yao | 2009 | Portfolio selection & network topology characteristic | China |
| Tabak et al. | 2010 | Relationships between different industries | Brazil |
| Nanda et al. | 2010 | Stock clustering based on data mining | India |
| Chi, Liu & Lau | 2010 | Build a stock return correlation network | United States |
| Bakker et al. | 2010 | Psychological factors affecting the investor trust network | North America |
| Namaki et al. | 2010 | Network topology characteristic | Tehran |
| Lee, Youn & Chang | 2012 | The least inclusive tree method in the stock market | Korea |
| Tumminello et al. | 2012 | Statistical evaluation of networks, validation of edges in a two-part system and determination of investor clusters | New York |
| Jiang et al. | 2013 | Abnormal business plans and stock utilization based on network topology | China |
| Aobdia, Caskey & Ozel | 2014 | Centrality & systematic risk | - |
| Ahern & Jarrad | 2014 | Maximize the efficiency of higher centralized industries | - |
| Shariatmadari et al. | 2014 | Network topology characteristic | Tehran |
| Sun, Tian & Yang | 2015 | Portfolio selection based on complex network analysis | United States |
| Gao, Zeng & Cia | 2015 | Building a directional network and ranking companies based on centrality criteria | China |
| Nie et al. | 2015 | Small world network & centrality based model | China |
| Huang et al. | 2015 | Consider different thresholds to compare network characteristics | China |
| Musmeci, Aste & Matteo | 2016 | Use correlation-based information filter networks and build a new tool for predicting market fluctuations | - |

| **Authors** | **Year** | **Methodology** | **Financial Market** |
|---|---|---|---|
| Li et al. | 2019 | Portfolio selection based on complex network | China & United States |
| Shi et al. | 2020 | The Evolution Characteristics of Systemic Risk in China's Stock Market Based on a Dynamic Complex Network | China |
| Huang et al. | 2021 | An empirical evaluation of the influential nodes for stock market network: Chinese A-shares case | China |

The issue of stock portfolio selection by optimizing risk and return has not been so far taken into consideration by the network in Iran, and all of these issues have been investigated and solved with old methods in the past including: stock price prediction and stock portfolio optimization using fuzzy neural networks and genetic algorithms (Zamani et al., 2013) as well as comparison of stock portfolio optimization with artificial neural networks (Monjemi et al., 2009), stock portfolio design using the Dematel method (Barzide et al., 2013), or optimization of stock portfolios using meta-heuristics methods (Talebi, 2010). Therefore, we intended to examine the issue of stock portfolio in this research by applying the analysis of Tehran Stock Exchange and considering the centrality criterion in the network and also by providing an optimization model and reviewing the final portfolio along with the combination of risk and return. Our research network was a set of companies with correlation of stock returns. Problem solving was performed by using the community detection method and optimization model. Finally, according to the results of the problem solving, the analysis of the problem was discussed based on the Tehran Stock Exchange index (TEPIX). In general, this issue attempts to show that general characteristics of a stock can be presented across the entire market against the traditional selection of portfolios that relies on statistical indices and only displays local behavior of stocks (Sun, Tian & Yang, 2015).

## 3. Network Construction

In this section, a correlation network will be constructed between the stock returns in the Tehran market, and the network conceptual model, data collection method and network visualization method will be shown.

### 3.1. The Conceptual Network Model

The communication network is constructed as follows:

If $P_i(t)$ denotes the price of the stock i on day t, then, the return of the stock price i in a one-day period from t-1 to t will be in accordance with Equation (1) (Shariatmadari, 2014):

$$R_i(t) = \ln\left(\frac{P_i(t)}{P_i(t-1)}\right) \tag{1}$$

The correlation coefficient between the price of companies i and j is obtained from Equation (2) according to Pearson's correlation coefficient formula (Shariatmadari, 2014):

$$C_{ij} = \frac{\langle R_i R_j\rangle - \langle R_i\rangle\langle R_j\rangle}{\sqrt{\left\langle R_i^2 - \langle R_i\rangle^2\right\rangle\left\langle R_j^2 - \langle R_j\rangle^2\right\rangle}} \tag{2}$$

The element of $\langle R_i\rangle$ denotes the average return of stock i during N days and is calculated according to Equation (3):

$$\langle R_i\rangle = (1/N)\sum_{t=1}^{N} R_i(t) \tag{3}$$

The correlation coefficient determines the direction and intensity of the correlation between two random variables. The value of this coefficient varies between -1 and 1, meaning that 1 is a positive total correlation, 0 is no correlation and -1 is a complete negative correlation (Namaki, 2011).

Now, we use threshold constraints to draw the edge. In this way, we assign 0 to correlations smaller than the threshold and assign 1 to correlations larger than the threshold. Therefore, the adjacency matrix A is defined for the considered network in Formula (4):

$$A_{ij} = \begin{cases} 1 & c_{ij} \geq \theta, i \neq j \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

Then, we assume in the obtained network that the edge with $C_{ij}$ weight is plotted between two nodes if and only if the absolute value of their correlation coefficient is greater thanthe threshold $\theta$.

The positive correlation coefficient means that two companies have the same behavior over time and the value of $\theta$ determines the degree of this similarity. In fact, if the correlation coefficient is positive, the price increase of the two companies will be directly correlated and, if it is negative, the price increase will be inversely correlated.

### 3.2. Data Collection & Preprocessing

In order to collect data, a dataset extracted from the Rahavard Novin software, including the closing price of 142 companies, is used during the period from 03/26/2011 to 04/08/2017. Based on the 2010 research findings (Chi, Liu & Lau, 2010), a two-year period is enough to achieve a significant correlation coefficient in stock market data, and the network is strong and significant.

Pre-processing was also carried out on these initial data. To simplify the network drawing and analysis process, first companies with less than 100 days without transaction and then companies with less than 50 days without transaction were removed from the data set. So that data was prepared for applying in network analysis software, programming environments and numerical computing, and a list of nodes and edges in the Excel became ready to import to the Gephi software.

### 3.3. Network Visualization

In Figure 4, the network created in the Gephi software (ver. 0.9.1) can be observed for value $\theta$=0.1.

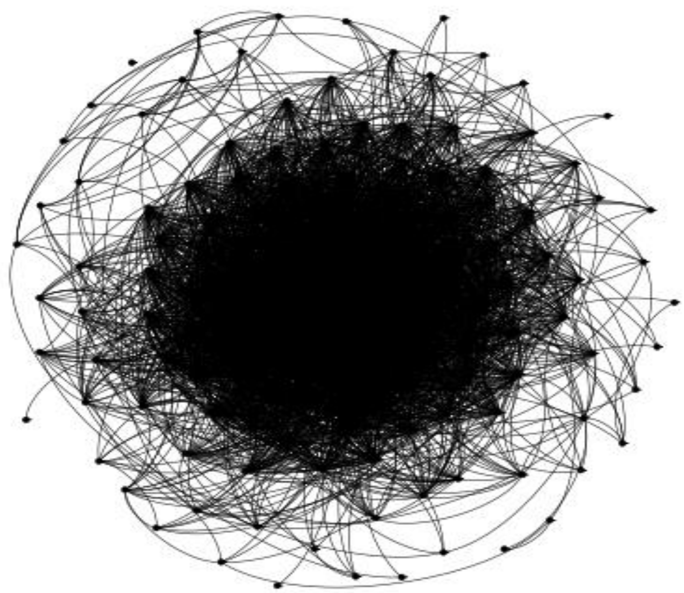

Fig. 4. The correlation network of stock returns based on $\theta$=0.1

In the following, for the network visualization, the Fruchterman Reingold layout algorithm was first used in the Gephi software, because this algorithm makes it easy to observe the correlation pattern of stocks across the market (Sun, Tian & Yang, 2015). Then, the Noverlap and Label Adjust layout algorithm were used to create a beautiful view without overlapping the nodes and with their labels being legible. Figure 5 shows the initial graph after applying the layout's algorithms.

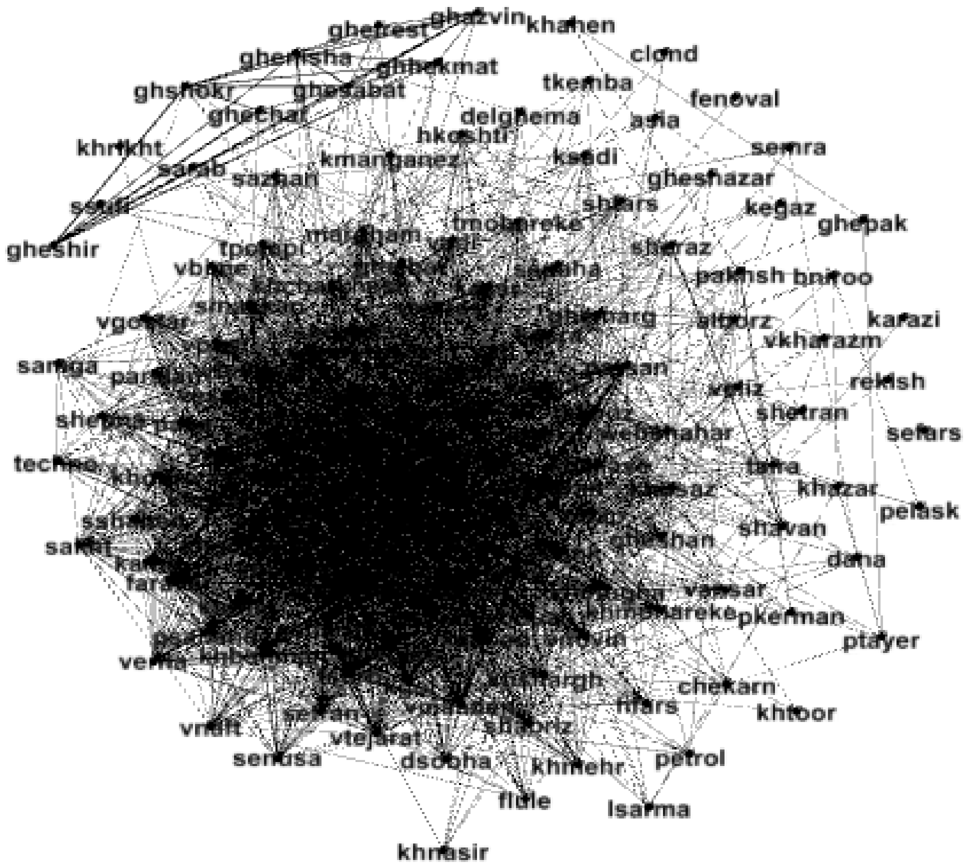


Fig. 5. The primary graph by applying Fruchterman Reingold, Noverlap and Label Adjust algorithms in Gephi

### 3.4. Community Detection

Community detection is one of the most popular methods for grouping nodes in the network according to the network structure. For community detection in this research, the Gephi software was used based on the Louvain algorithm. Then, using different colors in the network based on the modularity class, a colorful view of the network was presented, in which the same group nodes had the same color. Then, the betweenness centrality criterion was used to distinguish the size of the nodes. Thus, larger nodes had larger betweenness centrality. Furthermore, the edges were matched to their weight, so that the edges with more weight were darker. In the following, after the Louvain algorithm was applied, network is shown in Figure 6.

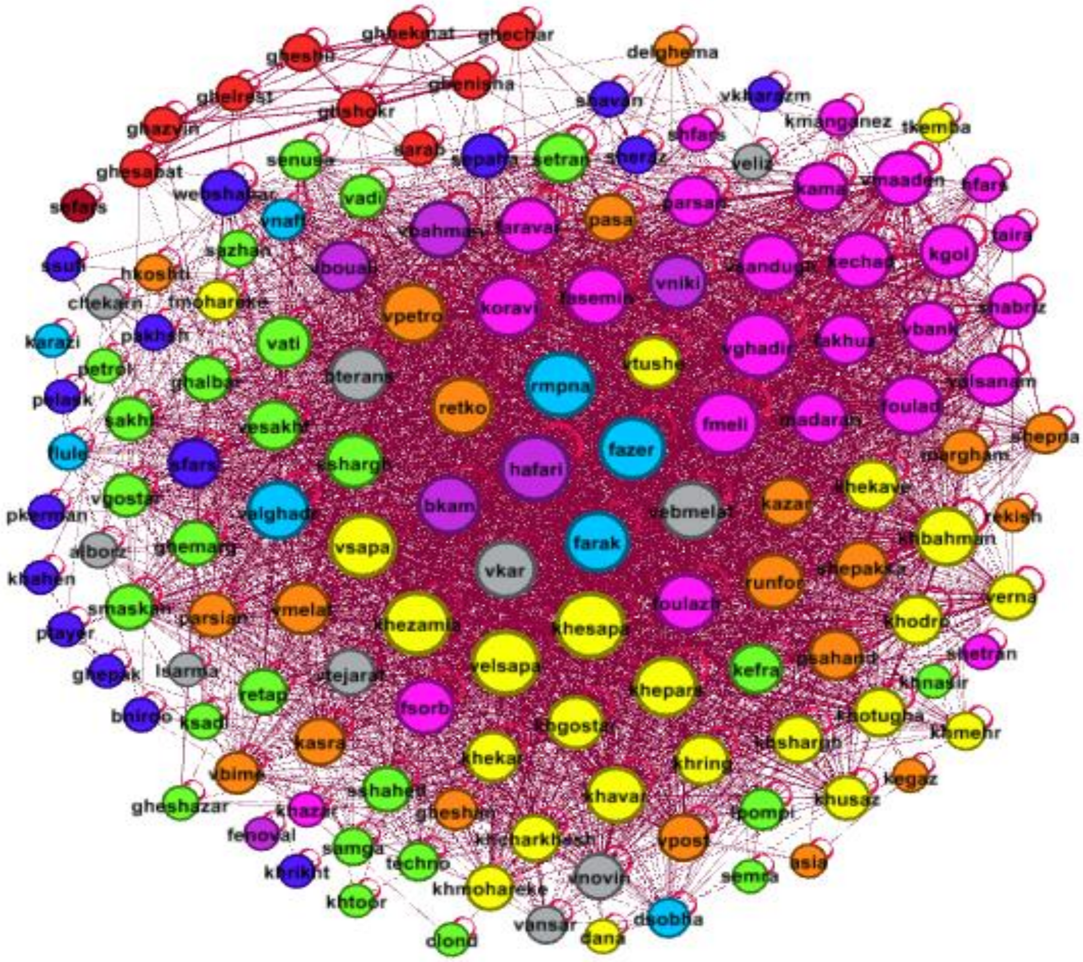


Fig. 6. The network visualization with 0.105 modularity and 10 communities

Table 2. Characteristics of the network

| Characteristics | Value |
|---|---|
| Number of nodes | 142 |
| Number of edges | 3457 |
| Number of communities | 10 |
| Layout | Fruchterman Reingold |
| Modularity | 0.105 |
| Average Degree | 24.345 |
| Network Diameter | 5 |
| Graph Density | 0.173 |
| Average Clustering Coefficient | 0.353 |
| Average Path Length | 1.737 |

### 3.5. The Mathematical Model

After the community was detected and different groups were identified, to create a portfolio, a stock had to be selected from each community, so that the portfolio diversity would be guaranteed in network and a diverse portfolio would be selected. The question now is which stock should be selected from each community. To answer this question, an optimization model was developed based on network centrality.

Markowitz (1952) was the first to choose a model of optimization based on variance or standard deviation as a measure of stock market risk in the selection of stock portfolios, and used the model form as Formula (5):

The Mathematical Model

$$\begin{aligned} &\text{minimize} \quad \sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j \sigma_{ij} \\ &\text{subject to} \quad \begin{cases} \sum_{i=1}^{n} R_i x_i = R^* \\ \sum_{i=1}^{n} x_i = 1 \quad , x_i \geq 0 \end{cases} \end{aligned} \tag{5}$$

In this model, $R_i$ is a random profit to the Si stock, xi denotes the amount invested in the Si stock, R* is the optimal profit of the investor and $\sigma_{ij}$ is the standard deviation between the two stocks.

After Markowitz, many researchers have studied portfolio fluctuations and introduced various types of risk. In this research, considering that the optimization model is based on network concepts, we attempt to inspire the Markowitz model based on the concept of network centralities with the goal of optimizing risk and return simultaneously. In the following, we first present the concepts of centralities used in the model.

- Closeness centrality: In correlated networks, this parameter is equivalent to the inverse of the distance mean of each node from other nodes. The larger this number, the more relevant node would be located in the center of the network. Closeness centrality can be calculated according to Equation (6).

$$C_c(v) = \frac{|V| - 1}{\sum_{u \in V, u \neq v} d(u, v)} \tag{6}$$

In Equation (6), the distance is measured with d(u,v) (e.g., the number of passing edges in the shortest path between the two nodes or the sum of the weights of these edges in a weighted network).

- Betweenness centrality: The betweenness centrality value is high in nodes that most appear in the shortest path between the nodes of the network. The value of this quantity for the node v is calculated as Equation (7).

$$B_c(v) = \sum_{(u,w) \in V \times V, u \neq w, u \neq v, w \neq v} \frac{\sigma_{uw}(v)}{\sigma_{uw}} \tag{7}$$

In (7), $\sigma_{uw}$ is the number of shortest paths from the node u to the node w, and $\sigma_{uw}$(v) is the number of shortest paths between u and w passing through v.

- Eigenvector centrality: Eigenvector centrality measures the importance of the presence of a node in the network. To calculate the eigenvector centrality of v (xv), Equation (8) is used:

$$x_v = \frac{1}{\lambda} \sum_{t \in M(v)} x_t = \frac{1}{\lambda} \sum_{t \in G} a_{v,t} x_t \tag{8}$$

Where M (v) is a set of neighbors of the vertex v, $\lambda$ is a constant value and A is the adjacency matrix (Newman, 2004).

### 3.6. The Optimization Model

According to the importance of the concept of centrality in various studies, centrality is the main criterion for the formation of an optimization model. Three types of centralities including closeness, betweenness and eigenvector were calculated using the Gephi software. To determine the weight of each centrality in this model, the centrality of the average of the three types of centralities was calculated for all the nodes in each community according to Equation (9):

$$C_{ave} = \frac{1}{3}C_c + \frac{1}{3}C_b + \frac{1}{3}C_e \tag{9}$$

Then, the centralities of the node with the highest average centrality were selected as the model coefficients and the average of its return was selected as the model's return. The objective function of the optimization model was defined as maximizing the product of the Sharpe ratio and the linear composition of the three types of centralities, according to Equation (10):

$$\begin{aligned} &\max \quad SR_p(\alpha_1 C_c + \alpha_2 C_b + \alpha_3 C_e) \\ &SR_p = (\bar{r}_p - \bar{r}_f) / \sigma_p \end{aligned} \tag{10}$$

In the objective function,  is the symbol of the Sharpe ratio of the node with the highest average centrality and $C_c$, $C_b$ and $C_e$ are the symbols of closeness, betweenness and eigenvector centralities of the node with the highest average centrality, respectively. The coefficients $\alpha_1$, $\alpha_2$ and $\alpha_3$ are the decision variables of the model. Furthermore, the following restrictions are defined as the model constraints:

$$\alpha_1 + \alpha_2 + \alpha_3 = 1$$
$$\alpha_1, \alpha_2, \alpha_3 \geq 0$$

The first constraint states that the sum of the coefficients must be one, so that at least one centrality with a positive coefficient must be selected. The second constraint denotes that coefficients are non-negative.

The obtained optimization model is solved in MATLAB software to identify optimal alpha. Then, in each community, by putting the centralities and the return of each node in the objective function formula, the node that is the result of the objective function has the highest value, which is chosen as the superior node.

## 4. Results

Table 3 lists the selected stock of each community.

Table 3. The list of the selected stocks of each community after solving the optimization model

| C1 | C2 | C3 | C4 | C5 | C6 | C7 | C8 | C9 | C10 |
|---|---|---|---|---|---|---|---|---|---|
| foulad | vbahman | hafari | khesapa | retko | sshargh | webshahr | fazer | ghesabet | sefars |

### 4.1. Validation of the Selected Portfolio

To validate the selected portfolio, its performance was compared with the performance of the TEPIX index. Therefore, first, a brief explanation of the TEPIX index was considered; then, the results were compared and the resulting graphs were analyzed.

### 4.2. The TEPIX Index

The TEPIX index was calculated by Equation (11) according to the Laspeyres formula:

**In Equation (12):**

$$\mathrm{TEPIX}_t = \frac{\sum_{i=1}^{n} p_{it} q_{it}}{D_t} \times 100 \qquad (11)$$

$p_{it}$ :The price of the company i at time t
$q_{it}$ :The number of stocks released from company i at time t
$D_t$ :The base number at time t equal to the origin time is equal to $\sum p_{io} q_{io}$
$p_{io}$ :The price of company i at the time of origin
$q_{io}$ :The number of stocks released from company i at the origin time
n :The number of companies covered by the index

The TEPIX index includes all the companies admitted to the Exchange; if the company symbol is closed and is not traded for a while, the price of the last trade in the index is taken into account. As is clear from Equation (11), the number of stocks released by the companies is a measure of weighting in the index, which has led to the greater impact of large companies on the index (Namaki, 2010).

In this research, the TEPIX index value was collected from the website tse.ir during the period from 03/26/2011 to 04/08/2017, and its performance was calculated for all the working days. After calculating the return of the selected portfolio performance during the mentioned time period, the results were compared with the total index performance. Figure 7 shows the results of this comparison.

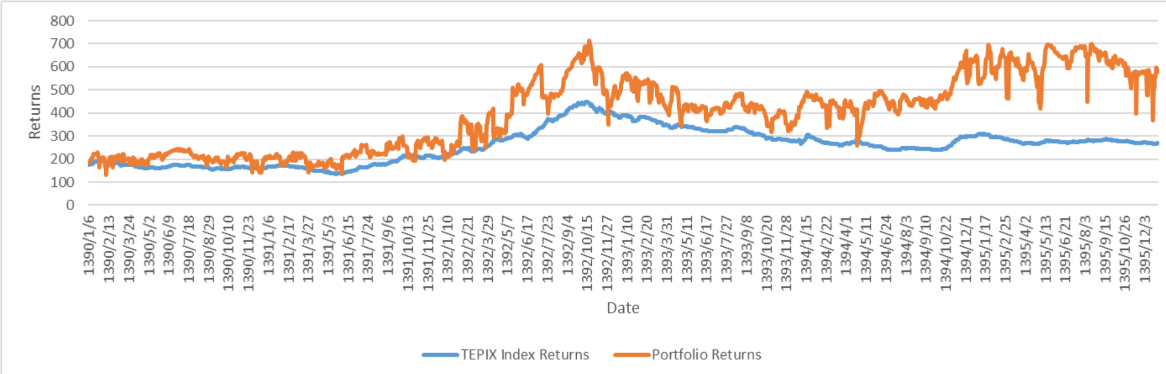


Fig. 7. Comparison of the selected portfolio returns with the TEPIX returns (portfolio model with risk)

As shown in Figure 7, the portfolio return chart was consistent with the TEPIX index return chart and over all the working days, had a higher return than the TEPIX index. Therefore, it can be mentioned that the portfolio selected in this research had good validity and accuracy. Also, given the greater volatility of the portfolio than the index, it can be said to be a high-risk portfolio.

For a more detailed analysis of the optimization model, we remove the risk parameter from the objective function and again compare the model output with the TEPIX index. As shown in Figure 8, in this case the portfolio also has a higher return than the index and, due to risk elimination, has less volatility and is almost a risk-free portfolio.

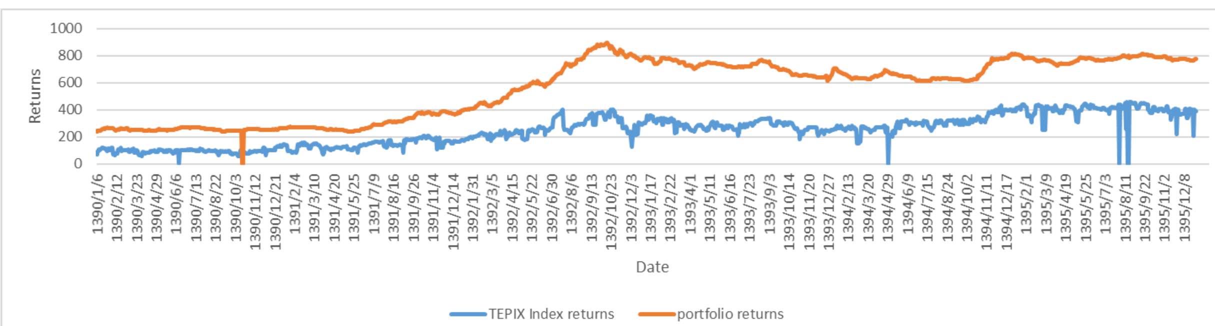


Fig. 8. Comparison of the selected portfolio returns with the TEPIX index returns (Portfolio model without risk)

## 5. Conclusion

Today, one of the most attractive ways to analyze and visualize big data is to use complex networks. Due to the large fluctuations of the stock market, decision making and analysis will be much more difficult for managers and financial experts. Accordingly, Tehran Stock Exchange was considered and, after creating a stock return correlation network, by using community detection, the stocks of the same group were analyzed. Finally, stocks with the same industry were placed into the same group and the community of the Louvain algorithm, which was based on modularity optimization, was completely logical and scientific. Furthermore, by designing a network-based optimization model with the goal of optimizing risk and returns, a stock was selected from each community. Then, by analyzing the performance of the network and selected portfolio with and without risk, it was determined that the portfolio efficiency was higher than the TEPIX index efficiency on all the days of the years and completely complied with it. Therefore, it can be stated that complex networks had very high ability to select stock portfolios. In summary, Contributions in the presented research included creating an optimization model based on the network centralities and optimizing risk and return simultaneously and paying attention to global characteristics of a stock across the network. In the paper of sun et al. (2015) they constructed an optimization model with centralities concept and just return. In our research, we considered risk too for analyzing this model in Tehran stock exchange more complete.

## References


[1] Ahern, K.R. & Harford, J. (2014). The importance of industry links in merger waves. The Journal of Finance, 69(2), pp.527-576.

[2] Albert, R., Jeong, H. & Barabasi, A.-L. (2000). Nature 406.378.

[3] Alexander, G.J., Sharpe, W.F. & Bailey, J.V. (1993). Fundamentals of investments (Doctoral dissertation, Univerza v Mariboru, Ekonomsko-poslovna fakulteta).

[4] Aobdia, D., Caskey, J. & Ozel, N.B. (2014). Inter-industry network structure and the cross-predictability of earnings and stock returns. Review of Accounting Studies, 19(3), pp.1191-1224.

[5] Aranha, C. & Iba, H. (2009). The memetic tree-based genetic algorithm and its application to portfolio optimization. Memetic Computing, 1(2), pp.139-151.

[6] Bakker, L., Hare, W., Khosravi, H. & Ramadanovic, B. (2010). A social network model of investment behaviour in the stock market. Physica A: Statistical Mechanics and its Applications, 389(6), pp.1223-1229.

[7] Barzideh, F., Taghvifard, M. & Zamanian, F. (2013). Analytic Network Process, DEMATEL, TOPSIS, Portfolio. Empirical Studies in Financial Accounting, 10(39), pp.105-124.

[8] Boginski, V., Butenko, S. & Pardalos, P.M. (2005). Statistical analysis of financial networks. Computational statistics & data analysis, 48(2), pp.431-443.

[9] Bonanno, G., Caldarelli, G., Lillo, F., Micciche, S., Vandewalle, N. & Mantegna, R.N. (2004). Networks of equities in financial markets. The European Physical Journal B-Condensed Matter and Complex Systems, 38(2), pp.363-371.

[10] Chaharsooghi, K. & Rahimnezhad, F. (2016). Application of complex networks in stock market analysis. In Proceedings of Innovation in IS/IT Management with BI Approuch conference, Tehran, Iran.

[11] Chen, K., Luo, P., Sun, B. & Wang, H. (2015). Which stocks are profitable? A network method to investigate the effects of network structure on stock returns. Physica A: Statistical Mechanics and its Applications, 436, pp.224-235.

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

[13] Gao, Y.C., Zeng, Y. & Cai, S.M. (2015). Influence network in the Chinese stock market. Journal of Statistical Mechanics: Theory and Experiment, 2015(3), p.P03017.

[14] Huang, C., Wen, S., Li, M., Wen, F. and Yang, X., 2021. An empirical evaluation of the influential nodes for stock market network: Chinese A-shares case. Finance Research Letters, 38, p.101517.

[15] 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), pp.2956-2964.

[16] Jiang, Z.Q., Xie, W.J., Xiong, X., Zhang, W., Zhang, Y.J. & Zhou, W.X. (2013). Trading networks, abnormal motifs and stock manipulation. Quantitative Finance Letters, 1(1), pp.1-8.

[17] Lee, J., Youn, J. & Chang, W. (2012). Intraday volatility and network topological properties in the Korean stock market. Physica A: Statistical mechanics and its Applications, 391(4), pp.1354-1360.

[18] Lee, K.E., Lee, J.W. & Hong, B.H. (2007). Complex networks in a stock market. Computer physics communications, 177(1), p.186.

[19] Li, P. & Wang, B.H. (2007). Extracting hidden fluctuation patterns of Hang Seng stock index from network topologies. Physica A: Statistical Mechanics and its Applications, 378(2), pp.519-526.

[20] Li, Y., Jiang, X.F., Tian, Y., Li, S.P. & Zheng, B., 2019. Portfolio optimization based on network topology. Physica A: Statistical Mechanics and its Applications, 515, pp.671-681.

[21] Lin, C.C. & Liu, Y.T. (2008). Genetic algorithms for portfolio selection problems with minimum transaction lots. European Journal of Operational Research, 185(1), pp.393-404.

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

[23] Markowitz, H., (1952). Portfolio selection. The journal of finance, 7(1), pp.77-91.

[24] Mohammed, A. M. (2021). Analysis and predictive validity of Kelantan River flow using RQA and Time Series Analysis. Kuwait Journal of Science, 48(1).

[25] Monajemi, A., Abzari, M. & Rayati, A. (2009). Pridicting of Stock Prices in the Stock Market Using Fuzzy Neural Network and Genetic Algorithms and Comparison with Artificial Neural Network. Journal of Economics. DOI: 10.22055/jqe.2009.10697.

[26] Motter, A.E. & Albert, R. (2012). Networks in motion. arXiv preprint arXiv:1206.2369.

[27] Motter, A.E., Nishikawa, T. & Lai, Y.C. (2002). Range-based attack on links in scale-free networks: Are long-range links responsible for the small-world phenomenon? Physical Review E, 66(6), p.065103.

[28] Musmeci, N., Aste, T. & Di Matteo, T. (2016). What does past correlation structure tell us about the future? An answer from network filtering. arXiv preprint arXiv:1605.08908.

[29] Namaki, A. (2010). Analysis of Tehran Stock Exchange using complex networks and randomized matrix theory. Thesis of Master of Science. Faculty of management. Tehran University.

[30] 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), pp. 3835-3841.

[31] Nanda, S.R., Mahanty, B. & Tiwari, M.K. (2010). Clustering Indian stock market data for portfolio management. Expert Systems with Applications, 37(12), pp.8793-8798.

[32] Newman, M.E. (2004). Analysis of weighted networks. Physical review E, 70(5), p.056131.

[33] Nie, H., Zhang, Y.B., Chen, J.R. & Xin-Guang, L. (2015). Research of Chinese Stock Market'Complex Network Structure. International Journal of Economics and Finance, 7(5), p.116.

[34] Onnela, J.P., Kaski, K. & Kertész, J. (2004). Clustering and information in correlation based financial networks.

The European Physical Journal B-Condensed Matter and Complex Systems, 38(2), pp.353-362.
[35] Shariatmadari, F. (2014). Pridicting of stock price fluctuations in Tehran Stock Exchange using social network analysis. Thesis of Master of Science. Faculty of Computer Engineering. Tehran University.
[36] Shariatmadari, F. & Asadpour, M. (2014). Social network analysis of Tehran Stock Exchange. International Management Conference. COI: ICOM01_0404.
[37] Shi, Y., Zheng, Y., Guo, K., Jin, Z. and Huang, Z., 2020. The Evolution Characteristics of Systemic Risk in China's Stock Market Based on a Dynamic Complex Network. Entropy, 22(6), p.614.
[38] Sun, W., Tian, C., & Yang, G. (2015). Network Analysis of the Stock Market.
[39] 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), pp.3240-3249.
[40] Talebi, A. (2010). Selection and optimization of the stock portfolio by using meta-heuristic and comparing it with the portfolio of experts and newcomers in the Tehran Stock Exchange. Thesis of Master of Science. Faculty of Industrial Engineering. Shahroud University of Technology.
[41] Tumminello, M., Lillo, F., Piilo, J. & Mantegna, R.N. (2012). Identification of clusters of investors from their real trading activity in a financial market. New Journal of Physics, 14(1), p.013041.
[42] Zamani, M., Afsar, A., Saghafinezhad, S. & Bayat, A. (2014). Expert system Stock price prediction and portfolio optimization using fuzzy neural networks, fuzzy modeling and genetic algorithm. Journal of Financial Engineering and portfolio Management, 21(5), pp.107-130.