Amirkabir University of Technology
(Tehran Polytechnic)
Vol. 45, No.1, Spring 2013, pp. 67 - 75

Amirkabir International Journal of Science & Research
(Modeling, Identification, Simulation & Control)
(AIJ - MISC)

A Stock Market Filtering Model Based on Minimum Spanning Tree in Financial Networks

A. Esfahanipour[1*], S. E. Zamanzadeh[2]

1- Assistant Professor, Department of Industrial Engineering and Management Systems, Amirkabir University of Technology
2- MSc Student, Department of Industrial Engineering and Management Systems, Amirkabir University of Technology

ABSTRACT

There have been several efforts in the literature to extract as much information as possible from the financial networks. Most of the research has been concerned about the hierarchical structures, clustering, topology and also the behavior of the market network; but not a notable work on the network filtration exists. This paper proposes a stock market filtering model using the correlation - based financial networks in which network nodes represent the potential stocks and network edges indicate the correlation coefficients of corresponding stock pairs. The model is capable of reducing the basic market size while keeping the diversification and risk - return expectations fairly constant. The novelty of this research is to develop a new market network filtering method which exploits Minimum Spanning Tree (MST) to reduce the number of network nodes (graph order) rather than the links (graph size). The proposed method chooses the nodes (stocks) based on dangling ends of the constructed MST. In order to verify our proposed model, we applied the model on data of three stock markets: New York Stock Exchange (NYSE), Germany Stock Exchange (DAX) and Toronto Stock Exchange (TSE). In conclusion, the numerical results showed that our proposed model can make a subset of the stock market in which its performance can imitate the whole market with a rather considerable reduction in size; as a result, we can have a diversified subset of the market compatible with that of the whole market. The performance of the model is confirmed by comparing the portfolio of the filtered market network with the whole market portfolio using the complement of Herfindahl Index as a measure of diversification.

KEYWORDS

Stock Market Filtering, Financial Networks, Minimum Spanning Tree (MST), Markowitz's Mean - Variance Method, Diversification.

* Corresponding Author, Email: esfahaa@aut.ac.ir

1- INTRODUCTION

Network and its application in the modeling problems - named graph theory - were coined by Euler (1736). [11] Readily, it became a popular tool in other fields such as mathematics, biology, sociology, transportation and recently in computer science, communication networks and finance. [14]

Networks have recently been applied to model the financial markets increasingly, as well as the economic relations. The essence of a network is dependent on the existence of a relation between two entities. This relation is presented by an edge connecting the network nodes as like the entities. In financial networks, the entities are stocks and the relations depict the correlation coefficients between stock pairs. [16]

Boginski et al. [4, 5] debated the concept of self - organized networks in finance and proposed independent sets in market network as diversified portfolios. Some other studies used network based methods to cluster the market network [19, 20, 22, 25]

Financial networks have assisted investors to describe the transactions among financial entities such as regional economics, banks and corporations [23]; also Likewise , they can be applied for market data analysis to model and investigate the market behavior. Since financial markets generate the huge amounts of data, networks have helped to get more reliable information from the financial markets. [10]

In order to obtain reliable information with less noise, it would be useful to eliminate unnecessary edges from the market network. Hence, to reduce the market data noise, we need an efficient method for the financial network. In the literature, there are some common methods which have been developed to make a network with as less noise as possible but performing as the whole data. In Random Matrix Theory (RMT) method, the real empirical correlation matrix is compared to a similar matrix with uncorrelated time series to recognize the noise in the matrix. Sharifi et al. [21] used the RMT method to remove the noise in the correlation matrix and then to optimize the portfolio more reliably.[8]

Many studies have been done on modeling the financial markets using correlation - based networks in which the market network has been filtered on edges (not nodes/stocks) to investigate the market behavior by developing hierarchical clustering methods. One method is to find the Planar Maximally Filtered Graph (PMFG). The PMFG is a network containing all the most correlated links that can be represented on a plane without crossing edges which is called planar graph. [25, 26]

Tumminello et al. [24] used the Minimum Spanning Tree (MST) of the market as a filtering tool on the linkages between stocks and constructed a hierarchical clustering method for the market network. Fieder (2014) introduced a new way to incorporate nonlinear dynamics and dependencies into the hierarchical networks to study financial markets using mutual information and its dynamical extension and finally showed that this approach leads to the different results than the correlation - based approach used in most studies, on the basis of 91 companies listed on the NYSE 100 between 2003 and 2013, using MST and PMFG. [12]

In the most recent study, Birch et al. (2015) compared three methods of filtering including MST, asset graph, and PMFG using stock price returns of DAX 30 between 2001 - 2012 and specifically focused on two specific time periods: (a) a period of crisis (Oct. - Dec. 2008) and (b) a period of recovery (May - Aug. 2010) representing new insights about the growth dynamics of an economy. [2]

Generally, the filtering processes used in the literature are mostly applied on removing the noise from the market correlation matrix (market network) by filtering the linkages/edges to make the correlation matrix more reliable for portfolio optimization. Now, this paper proposes a new way of using MST as a tool for filtering the market indices.

The minimum - weight spanning tree problem (MSTP), one of the most typical and well - known problems of combinatorial optimization, is concerned with finding a spanning tree of an undirected connected graph, so that the summation of the weights of the selected edges becomes minimum. Because of the important applications of the MSTP in network problems, it has been for a long time the subject of research for scientists to develop more efficient algorithms. [13]

Bazlamacci and Hindi (2001) surveyed the existing algorithms for solving the MSTP and made a classification according to the network size and time performance as classical algorithms and modern algorithms. The modern algorithms, that have been developed recently, use the random search methods for solving the MSTP which causes to run more quickly than the classical algorithms. Therefore, the modern algorithms are more complex but more efficient for large scale problems than classical algorithms. In this research, we applied the Kruskal algorithm [15] as an available modern algorithm with less complexity to solve the MSTP.

In the present paper, focusing on financial networks, we developed a market network filtering model using correlation - based networks to filter the stock indices (nodes) rather than the linkages (edges). The main objective of our proposed model is to construct a reduced subset of stocks in the market resembling the behavior of the whole market. Therefore, the portfolio from the filtered market represents the portfolio of the whole market with remarkably similar performance. The most important advantage is the reduction in the cost and the risk of handling a big market data. Therefore, we can follow up a small group of stocks rather than the whole market.

In the next section, we will describe the proposed model, and then we will make a short look on mean - variance portfolio optimization model which is a tool for evaluating the performance of the filtering model. In section 3, we used data from three different stock market markets to evaluate the model.

2-THE PROPOSED STOCK MARKET FILTERING MODEL

In this section, our proposed stock market filtering model is described in detail. This model focuses on developing a filtering procedure to select stocks among the whole market and then running the mean - variance model for the selected subset to check the model efficiency.

As it is shown in Figure 1, the process of the proposed model can be described as follows.

i. Calculating cross - correlation matrices to set up the corresponding networks based on daily return
ii. Finding the MST of the market network using the Kruskal algorithm.
iii. Filtering the network by selecting vertices with degree one in the MST
iv. Running Markowitz's mean - variance portfolio optimization model for the selected stocks and compare with the portfolio from the whole market to compare the performance

In the following subsections, we describe the process of our proposed model thoroughly. Firstly, we use daily prices of the stock market to construct the market network. Then, we define the proposed filtering stock market model.

2- 1- CONSTRUCTING MARKET NETWORK

To calculate the cross - correlation matrices based on return we use the price matrix which contains time series of daily stock prices for each stock. We explain the calculation of network matrix as follows.

To obtain the return matrix, first of all, we calculate the time series of price return according to Eq. 1 and then the correlation coefficients between pairs of stocks by Eq. 2 which result in an $n \times n$ matrix called the cross - correlation matrix for return, as most of the references did [2 - 7, 10, 12, 16, 20, 22 - 27]. All of the correlation - based networks in literature are obtained from the return measure;

$$R_i(t) = \ln \frac{P_i(t)}{P_i(t-1)} \qquad \text{Eq. 1}$$

$$C_{ij} = \frac{E(R_i R_j) - E(R_i)E(R_j)}{\sqrt{Var(R_i) + Var(R_j)}} \qquad \text{Eq. 2}$$

where $P_i(t)$ is the price of asset i on date t.

Since we want to evaluate the market network by considering the degree of similarity among behavior of stocks, we need to define a similarity measure. The correlation coefficient of a pair of stocks cannot be used as a distance measure between the two stocks; because it does not satisfy the three axioms of an Euclidean metric as follows: (i) $d_{ii}=0$; (ii) $d_{ij} = d_{ji}$; (iii) $d_{ij} \le d_{ik} + d_{kj}$; [16, 6, 7] Hence, we apply a distance measure as of Eq. 3 as was suggested by Mantegna [16] and Bonanno et al. [6]to satisfy the above - mentioned three axioms of an Euclidean metric. The result of the transformation of coefficient correlation to the distance measure is an $n \times n$ symmetric matrix which is a corresponding mapping from interval of [- 1, 1] to the interval of [0, 2]. This transformed matrix is used representing the weights of the edges in the next section.

$$d_{ij} = \sqrt{2(1-c_{ij})} \qquad \text{Eq. 3}$$

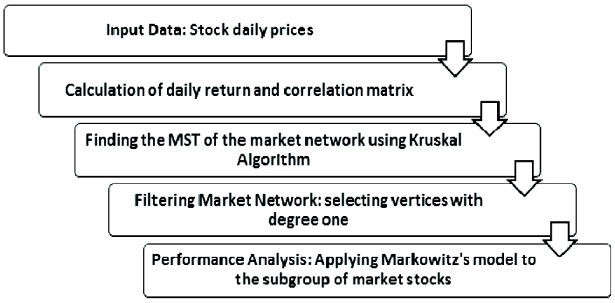

Figure 1: Schematic process of the proposed stock market filtering model

2- 2- FILTERING MARKET NETWORK USING MST

The second phase in the proposed model is to apply the correlation based network in the process of portfolio selection using the Minimum Spanning Tree Problem (MSTP). The novelty of the model is to define a method for filtering the MST of the network and selecting the assets of the portfolio according to the structure of the MST.

MST is a well - known tool in the network literature to filter redundant edges from networks. For instance, consider the tree $T(V, E')$ as a subset of graph $G(V, E)$ and $E' \subset E$. For each edge $e_{ij}=(v_i, v_j)$ in this tree, a corresponding weight w_{ij} has been defined so that $e_{ij} \in E$ and $v_i, v_j \in V$. The unique tree T for graph G is MST if and only if the value of the function $\sum_{e \in E'} w(e)$ is minimized.

As it was mentioned, because of less complexity, we apply the Kruskal algorithm to find the MST of the market network as it was used by Graham and Hell [14] for a network having 6358 nodes. The Kruskal algorithm is a greedy algorithm in which the edges are sorted according to the weights in an ascending order. The weights are the transformed values of distance matrix as described in the previous section. At first, edge e_1 with minimum weight is selected. Then edges e_2, e_3… e_{i+1} are selected, respectively if only one endpoint of the edge is in the same tree; otherwise, it is neglected. The time complexity of the algorithm is O (Elog(V)). The output of this algorithm is a tree with minimum sum of weights that having the same number of vertices (n) in the market network, but with (n - 1) edges.[1, 15] In the next section, we use this MST to select the portfolio.

The novelty of this study is in constructing a market subset with a size smaller than the market size but a more efficient one. To devise our selection model, we consider three topological properties in market network: (i) Important nodes, (ii) Edges, and (iii) Dangling ends. The important nodes should correspond to the companies that control or mediate the daily fluctuations of their neighborhoods. The edges are mediating the information (fluctuation) along the MST branches. Dangling ends are supposed to be less influenced by the market fluctuations and their corresponding symbols (or companies) do not necessarily follow the market indices. [27] Therefore, we look for the important nodes that are not influenced by market sudden changes and they are less correlated with the dominant stocks. This issue has been discussed by Vandewalle et al. [27] that leads to a meaningful topology in market network.

In a recent study, Boginski et al. [4] considered the elements of independent sets in a network as elements of a diversified portfolio. Independent sets are the group of stocks having the least or negative correlation. Similarly, in their recent research, they used the weighted market graph model to provide a new framework for selecting profitable diversified portfolios. [3]

With this idea in mind, we can select less correlated stocks to have more diversification and a portfolio with a smaller size. Therefore, in this study, we select dangling ends as the candidate stocks for filtering of the market.

In previous section, we have used MST to decrease the number of relations between stocks to (n - 1). This means that among the all possible edges in market graph, the remaining n - 1 edges draw the most important relations and can demonstrate the entire market holistically. It is known that in a tree, there is only one path between every pair of stocks; the longer the path, the less correlated the stocks. Thus, to realize the diversification of the portfolio, we use the dangling ends of the tree. Therefore, in our proposed filtering market model, the selected stocks are those corresponding to the vertices with the degree of one as red nodes which are shown in Figure 2.

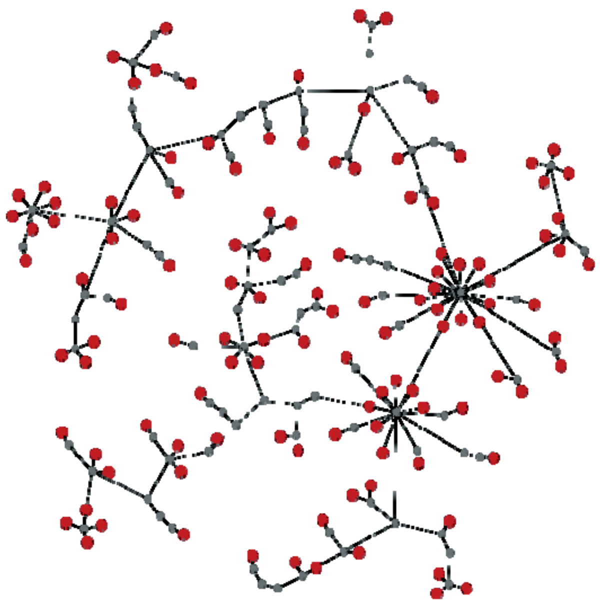

Figure 2: The MST for NYSE Market Network; dangling ends are colored (obtained from Gephi 0.8.2 software)

2- 3- PERFORMANCE ANALYSIS

To evaluate the performance of the market filtering model, we use the portfolio optimization for the selected stocks. Then, we can judge the effectiveness of the filtering model to avoid missing any necessary data.

Portfolio optimization is one of the most fascinating areas in finance that has attracted many studies. Many models have been developed to make an investment portfolio as efficient as possible. The most well - known method is Markowitz's model named as Modern Portfolio Theory (MPT) which is based on risk diversification and balancing the volatilities in the portfolio [17, 18]. The mean - variance method is a way to trade - off between portfolio return and portfolio risk using a diagram which is called efficient frontier. Depending on the model characteristics, there are several versions which can be applied in the proposed model. Various extensions of the Markowitz's model have been reported in Engels' master thesis report. [9]

According to the Markowitz's model, the mathematical programming model (1) must be solved to determine the portfolio weights of w_is. In this model, N is the number of risky assets, m_i is the mean return of asset i, σ_P is the standard deviation of the portfolio return which is considered as a risk measure that should be minimized, σ_{ij} is the covariance between returns of assets i and j.r_P is the portfolio required return which is determined by the investor.

Model (1):

$$Minimize\ \sigma_p^2 = \sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j \sigma_{ij}$$

$$Subject\ to:$$

$$\sum_{i=1}^{N} w_i m_i = r_p$$

$$\sum_{i=1}^{N} w_i = 1$$

The optimization problem consists of finding the vector w which minimizes σ_p for a given value of r_p. By changing values of r_p, we can have a diagram showing the risk - return curve for a risky asset which is called efficient frontier diagram. Having the efficient frontier, we can find the best portfolio according to the investor's preferences. The model (1) can be used in different settings as described in Engels' report [9]. In this paper, we use the tangency portfolio model which is the composition with maximum Sahrpe ratio - the most risk - efficient portfolio. (Engel, 2004)

To evaluate the degree of diversification (DI) of a portfolio, some measures of diversification have been developed. As it is mentioned by Woerheide and Persson [28], there is no profound reason for the use of a particular functional form. Therefore, we use the complement of the Herfindahl Index (HI) as the most widely used measure as follows.

$$DI = 1 - HI = 1 - \sum_{i=1}^{N} w_i^2 \qquad \text{Eq. 4}$$

According to the above relation, DI can takes the values in the range of [0, 1] in which zero represents no diversification while 1.0 would represent the maximum level of diversification.

3- DATA AND RESULTS

This section describes application of the proposed model in three financial markets: New York Stock Exchange (NYSE), Germany Stock Exchange (DAX), and Toronto Stock Exchange (TSX). The first subsection illustrates our data collection. We apply our model on data from NYSE and then we conclude that efficient frontier which is obtained from the filtered market network is very closely matched with the efficient frontier that is obtained from whole market network. Based on this result, we apply our model on data from DAX and TSX for portfolio selection using only return measure. Then, we end up with our main results based on data of the three markets as follows. Our proposed method can finally select a portfolio that its risk - return performance can imitate that of the whole market with a significant reduction in market size while keeping a fairly similar diversification. Therefore, we used a diversification index to examine the portfolio diversification.

3- 1- DATA

To examine our developed model, we collect data from three financial markets. The input data is the daily stock price time series and they are used to calculate the series of return. The cross - correlation matrix results in a network. The Kruskal algorithm finds the MST and then the dangling ends of the MST will constitute the portfolio.

The input data are extracted from Google - Finance and Yahoo - Finance websites. The data from New York Stock Exchange (NYSE) includes stock prices of 2750 days for 210 companies which are listed in the S&P500 as Sharifi et al. [21] and Mantegna (1999) did. The NYSE price time series are for 2751 days in the range of October 2000 to August 2011 which forms a 210*2751 matrix as input for our model. This matrix is then used in MatLab software to run our model. The input data from Germany Stock Exchange (DAX) comprises stock prices of 83 companies for 2500 days and the input data for Toronto Stock Exchange (TSX) includes stock prices of 125 companies for 3055 days up to October 2012. The reasons for selection of these three stock markets are: i) in order to test our proposed model to more real data in different conditions; ii) to trade - off between the number of the stocks and the length of time series, which led to a smaller market size for DAX and TSX; iii) these three markets were among the top 10 stock markets that we had access to the data with the required conditions expressed above.

3- 2- RESULTS

The return time series calculated by Eq. 1 forms a 210*2750 matrix. The fluctuations of return for two sample companies are shown in Figure 3. We can see that the fluctuations are not similarly repeated through the time. Hence, some volatilities are due to the market (systematic) risk and some others are for the company.

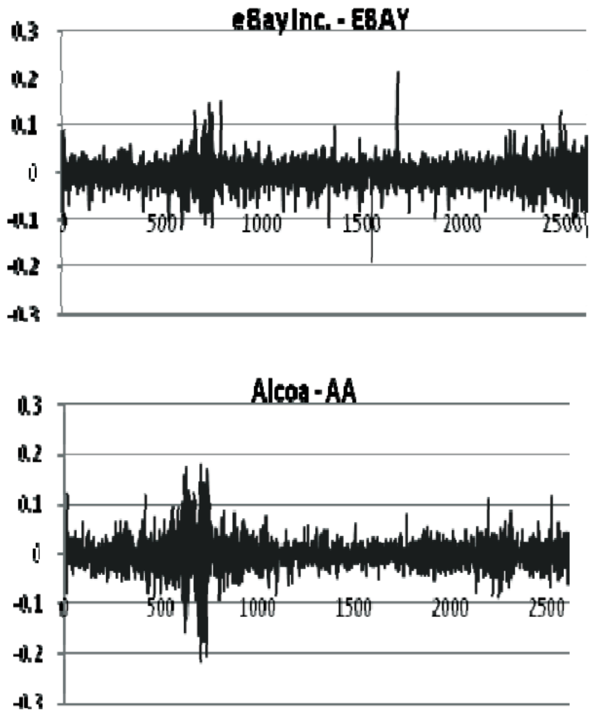

Figure 3: The fluctuations of return for companies Alcoa and eBay Inc from NYSE for 2750 days.

We used distance matrix as weights of edges in Kruskal algorithm to find the Minimum Spanning Tree (MST).In filtering process, we selected dangling ends as elements of the portfolio (Fig. 2). The number of stocks selected in this way was 119 out of 210 for NYSE. As we discussed earlier on performance analysis, we used the Markowitz's mean - variance method to draw the efficient frontier and find the tangency portfolio, which results in a portfolio with maximum Sharpe ratio.

We can compare the performance of the portfolio from filtered market with the portfolio from whole market in Fig. 5. The line "All Symbols (210)" represents the efficient frontier for whole market portfolio. The most obvious outcome of Fig. 5 is that using MST to filter the market has no effect on the diversification and performance of the whole market portfolio; because the portfolios from the filtered network imitate the whole market behavior ideally.

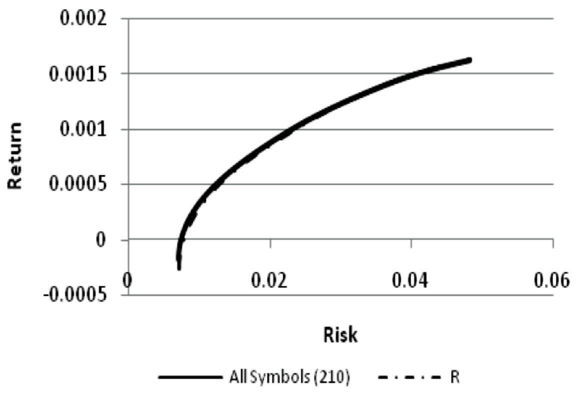

Figure 4: Efficient frontiers obtained from optimization of the whole market (all 210 symbols) and filtered return (R) matrix using data from NYSE

Finding the tangency portfolio for the efficient frontier, proceeds to the combination of the portfolio and the weights of the assets. We used MatLab for the weights of stocks in two portfolios are shown in Table 1.The table shows the share of 13 stocks in the tangency portfolio obtained from running Markowitz's portfolio optimization model on the selected stocks (119 stocks) and the whole market (210 stocks), respectively. The similarity in the number of portfolio stocks is casual, and it may not happen in another market.

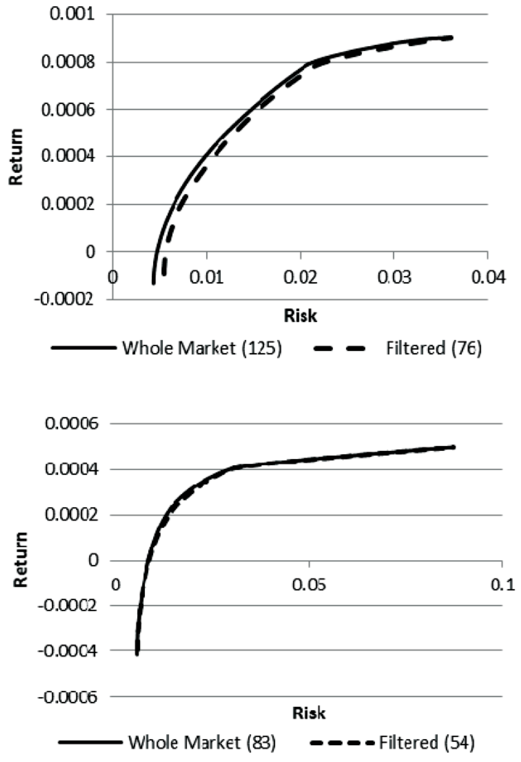

Figure 5: Efficient frontiers obtained using filtered matrix and the whole market: (1st) Toronto Stock Exchange, (2nd) Germany Stock Exchange

We can see that the effective size of the portfolio has not been changed after the filtering process. 9 out of 13 stocks are also the same for both portfolios. The aforementioned results show that filtering the whole market with the proposed method does not change the whole market portfolio performance remarkably.

To confirm the results, we also tested the portfolio using data from two other stock markets: Germany and Canada. In Figure 5, the efficient frontiers derived from the mean - variance method are drawn for these two markets using the whole market data and the filtered one. It can be seen that the efficient frontier for the whole market (125 stocks for TSX and 83 stocks for DAX) can coincide the efficient frontier of the filtered market (76 selected stocks for TSX and 54 for DAX). This means that we can get the whole market risk - return outcome from the selected stocks.

To verify the claim that the proposed model can keep the diversification of the portfolios, we used the complement of the Herfindahl Index. The obtained measures for the NYSE are 86.6 and 87.8 for the portfolio of the filtered market and the whole market, respectively. This index has been also calculated for DAX and TSX which are shown in Table 2.This table shows a less than 10% negative change in DI for different markets. DI is directly influenced by the total number of stocks, in which diversification increases by the effective size. Although the percent of change are is in line for three markets, the differences between the reduction in DI and portfolio size are mainly due to the differences in size of the market data and also the topology of the market structures.

Table 1

Portfolio weights obtained from Return matrix and the whole market using NYSE data

	Filtered Market		Whole Market	
No.	Symbol	Weight	Symbol	Weight
1	GE	0.002697	C	0.079791
2	BMY	0.013422	PFE	0.040992
3	MRK	0.052804	CSCO	0.108154
4	EMC	0.059811	EMC	0.039937
5	NWSA	0.054918	NWSA	0.007195
6	S	0.102736	S	0.088205
7	TWX	0.089996	TWX	0.083837
8	AES	0.030387	AES	0.02975

	Filtered Market		Whole Market	
No.	Symbol	Weight	Symbol	Weight
9	AIG	0.261234	AIG	0.248261
10	BRCM	0.069819	BRCM	0.017111
11	DUK	0.110687	DUK	0.058038
12	ETFC	0.142646	ETFC	0.09988
13	CNP	0.008842	FITB	0.098849
Sum of the weights		1		1
Diversification (DI)		86.6		87.8

Table 2

Diversification Index (DI) and size for the portfolios of the Whole Market and the Filtered Market

Market	Whole Market Portfolio			Filtered Market Portfolio				
	No. of Stocks	DI	Size	No. of Stocks	% change	DI	% change	Size
NYSE	210	87.8	13	119	- 43%	86.6	- 1.36%	13
DAX	83	77.3	6	54	- 35%	72.5	- 6.20%	5
TSX	125	88.1	15	76	- 39%	80.8	- 8.29%	10

4- CONCLUSION

Considering the objectives of the research, in this paper, we used the correlation based financial networks to build the MST of the market. Then we selected the dangling ends of the MST as our filtering benchmark.

According to the results of running the proposed filtering model on assets in New York Stock Exchange (NYSE), Germany Stock Exchange (DAX) and Toronto Stock Exchange (TSX), the coincidence of the risk - return diagrams showed that the performance of the portfolio from the selected stocks imitates the portfolio of the whole market very closely. This means that we can use a subset of the whole stock market as a typical representative keeping the same characteristics. This result was confirmed when the Diversification Index calculated for the optimized portfolios (both the whole market and the filtered one) was fairly close.

For future research, the authors suggest using the model as a means of building a market index to trace a market subset instead of the whole market index. The filtering method can also be enriched theoretically to work more effectively on reducing the number of selected stocks and to be used as a portfolio selection method similar to methods as introduced in [29, 30].

5- REFERENCES

[1] Bazlamacci, Cuneyt F., and Khalil S. Hindi, "Minimum - weight spanning tree algorithms: A survey and empirical study, "Computers & Operations Research, 28, pp. 767 - 785, 2001.

[2] Birch J., Pantelous A. A., Soramake K, "Analysis of Correlation Based Networks Representing DAX 30 Stock Price Returns", Computational Economics, DOI: 10.1007/s10614 - 015 - 9481 - z, 2015.

[3] Boginski V., Butenko S., Shirokikh O., Trukhanov S., Lafuente J. G, "A Network - based data mining approach to portfolio selection via weighted clique relaxations", Annual Operations Research, 216, pp. 23 - 34, 2014.

[4] Boginski, V., Butenko S. and Pardalos, P.M, "Mining market data: A network approach.", Computers & Operations Research, 33, pp. 3171 - 3184, 2006.

[5] Boginski, V., Butenko, S. and Pardalos, P.M, "Statistical analysis of Financial networks.", Computational Statistics & Data Analysis, 48, pp. 431 - 443, 2005.

[6] Bonanno, Giovanni, Guido Caldarelli, Fabrizio Lillo, and Rosario N. Mantegna, "Topology of correlation based minimal spanning trees in real and model markets", Phys. Rev. E, 68, 2003.

[7] Bonanno, Giovanni, Guido Caldarelli, Fabrizio Lillo, S. Micciche, N. Vandewalle, and Rosario N. Mantegna, "Networks of equities in financial markets", Eur. Phys. J. B, pp. 363 - 371, 2004.

[8] Daly, J., M. Crane, and H. J. Ruskin, "Random matrix theory filters in portfolio optimisation: A stability and risk assessment", Physica A, 387, pp. 4248 - 4260, 2008.

[9] Engels, Marnix, "Master Thesis: Portfolio Optimization: Beyond Markowitz", Universiteit Leiden, Rabobank, 2004.

[10] Eom, C., Oh, G., Jung, W. S., Jeong, H., Kim, S, "Topological properties of stock networks based on minimal spanning tree and random matrix theory in financial time series", Physica A, 388, pp. 900 - 906, 2009.

[11] Euler, L, "Solutio problematis ad geometriam situs pertinentis"m Commetarii Academiae

Scientiarum Imperialis Petropolitanae", pp. 128 - 140, 1736.

[12] Fiedor P, "Networks in financial markets based on the mutual information rate", Physical Review E 89(5), 2014.

[13] Gabow, H. N., Z. Galil, T. H. Spencer, and R. E. Tarjan, "Efficient algorithms for finding minimum spanning trees in undirected and directed graphs", Combinatorica, pp. 109 - 122, 1986.

[14] Graham, R. L., and Pavol Hell, "On the history of minimum spanning tree", Annals of the History of Computing 7(1), pp. 43 - 57, 1985.

[15] Kruskal, Joseph B, "On the Shortest Spanning Subtree of a Graph and the Traveling Salesman Problem", Proceedings of the American Mathematical Society, 1956.

[16] Mantegna, R. N, "Hierarchical structure in financial markets", Eur. Phys. J. B, pp. 193 - 197, 1999.

[17] Markowitz, H, "Portfolio Selection", Journal of Finance, 7(1), pp. 77 - 91, 1952.

[18] Markowitz, H, "Portfolio Selection: Efficient Diversification of Investment", New York: Wiley, 1959.

[19] Onnela, J. P., A. Chakraborti, and K. Kaski, "Dynamics of market correlations: taxonomy and portfolio analysis", Phys. Rev E, 68, no. 056110, 2003.

[20] Onnela, J. P., K. Kaski, and J. Kertesz, "Clustering and information in correlation based financial networks", Eur. Phys. J. B, pp. 353 - 362, 2004.

[21] Sharifi, S., M. Crane, A. Shamaie, and H. Ruskin, "Random matrix theory for portfolio optimization:a stability approach", Physica A, 335, pp. 629 - 643, 2004.

[22] Tola, V., F. Lillo, M. Gallegati, and R. M. Mantegna, "Cluster analysis for portfolio optimization", Economic Dynamics & Control, 32 pp. 235 - 258, 2008.

[23] Tse, Chi K., Jing Liu, and Francis C. M. Lau, "A network perspective of the stock market", Journal of Empirical Finance, 17, pp. 659 - 667, 2010.

[24] Tumminello, M., F. Lillo, and R. N. Mantegna, "Correlation, hierarchies, and networks in financial markets", Journal of Economic Behavior & Organization, 75, pp. 40 - 58, 2010.

[25] Tumminello, M., F. Lillo, and R. N. Mantegna, "Hierarchically nested factor model from multivariate data", Europhysics Letters, (2007).

[26] Tumminello, M., T. Di Matteo, T. Aste, and R. M. Mantegna, "Correlation based networks of equity returns sampled at different time horizons", Eur. Phys. J. B, pp. 209 - 217, 2007.

[27] Vandewalle, N., F. Brisbois, and X. Tordoir, "Non - random topology of stock markets", Quantitative Finance, 1, pp. 372 - 374, 2001.

[28] Woerheide W., Persson D, "An Index of Portfolio Diversification, "Financial Services Review, 2 (2), pp 73 - 85, 1993.

[29] Albadvi, A., Chaharsooghi, S.K. and Esfahanipour, A, "Decision Making in Stock Trading: An application of PROMETHEE", European Journal of Operational Research, 177, pp. 673 - 683, 2007.

[30] Chaharsooghi, S.K. Albadvi, A., and Esfahanipour, A, "Portfolio Selection in Stock Exchange through Industries and Comanies Ranking (Written in Persian) ", Amirkabir, 17 (65 - B), pp. 21 - 29, 2007.