_____

Res. Assist. Yusuf Yargı BAYDİLLİ, PhD Candidate
E-mail: yusufbaydilli@karabuk.edu.tr
Assist. Professor Şafak BAYIR, PhD
E-mail: safakbayir@karabuk.edu.tr
Assist. Professor İlker TÜRKER, PhD
E-mail: iturker@karabuk.edu.tr
Karabuk University

A HIERARCHICAL VIEW OF A NATIONAL STOCK MARKET AS A COMPLEX NETWORK

Abstract. We created a financial network for Borsa Istanbul 100 Index (BIST–100) which forms of N=100 stocks that bargained during T=2 years (2011–2013). We analyzed the market via minimum spanning tree (MST) and hierarchical tree (HT) by using filtered correlation matrix. While using hierarchical methods in order to investigate factors that affecting grouping of stocks, we have taken account the other statistical and data mining methods to examine success of stock correlation network concept for portfolio optimization, risk management and crisis analysis. We observed that financial stocks, especially Banks, are central position of the network and control information flow. Besides the sectoral and sub-sectoral behavior, corporations play role at grouping of stocks. Finally, this technique provided important tips for determining risky stocks in market.

Keywords: Stock correlation network, stock market, financial network, correlation-based clustering, degree distribution, Istanbul Stock Exchange, minimum spanning tree, hierarchical tree.

JEL Classification: G11

1. Introduction

Stock market (or bourse) is a complex system where stocks and shares are bought and sold. One of the reasons of this complexity is that the time series of stock returns are unpredictable. Namely, stock returns can be described as random processes [1,2]. Likewise, other systems, stocks in this system are in various relationships. Analyzing these relationships help to understand current market dynamics [3,4], predict market movements [3] and determine major stocks giving direction to sectors and portfolio [5]. One of the methods, that are used to analyze this relationship, is stock correlation network technique.

This kind of network provides considerable economic information by arranging stocks as topological [1,5–7]. We are able to see cooperative behaviors in the stock market, by processing correlation among stocks [8] and obtaining sub-dominant ultra-metric distances through filtered correlation matrix which provide a

chance to look into structure via geometrical (through minimum spanning tree) and taxonomic (through hierarchical tree) [9].

Stock network analysis attracted researchers like investors especially after first work done by R. N. Mantegna [1]. This technique was used widely for stock exchanges, currency exchange rates, of world trade, commodities, of GDPs (Gross Domestic Products), corporations and world financial markets [10]. Most of these works dealt with major financial markets or other countries [2–19], besides, there are some works represent dynamics of Borsa Istanbul (formerly Istanbul Stock Exchange) [20,21]. However, according to our knowledge, there are no previous work studied such a large network for Borsa Istanbul 100 Index (most traded stocks) represent topological arrangement of stocks via *MST* and *HT* techniques. On the other hand, this study examines movements of stocks with several statistical and data mining techniques, which strengthen financial structure investigation concept for risk management, prediction and portfolio analysis.

2. Data

In this study, stock correlation network procedure applied to Borsa Istanbul 100 Index (BIST–100), which is main indicator of the market. Investigating data scale for *N=100* stocks was *T=506* days (2011–2013) and opening prices of stocks used as time series. This data scale was important to understand abilities of financial network techniques due to there were a decreasing period until *Feb-2012*. The reason of this fall can be shown effects of crisis periods that start with Great Recession (2008) leads to Global Recession (2010) which affect EU Markets, US credit rating downgrade, some important actions Turkish Central Bank made and sales of stocks by foreign investors because of panic [4,10,21–23].

In order to analyze factors affect grouping of stocks, we classified them according to *economical sectors*, *sub-sectors* and *indices* In addition, due to the composition of the BIST–100 index are reviewed and adjusted 4 times on a quarterly basis in a year, we created our data from the last 3-months period *(October–December of 2012)* to obtain how these stocks evolve through time from first period until last one. Investigated stocks can be seen in Table 1, which were classified according to asset code, name, sector and sub-sector. Classification and time series was provided by Bourse Istanbul Database [24].

3. Methodology

In this study, we used two main approaches; one investigates geometrical (MST) and the other one taxonomic (HT) of structure. We filtered correlation matrix with *ALCA* method due to MST is some kind of representation of *SLCA* and ALCA provides more information [18].

Let $P_i(t)$ be the opening price of a stock i at time t. Then, a similarity measure between the synchronous time series of a pair of stocks, i and j, is calculated by the following correlation coefficient formula

$$\rho_{i,j} = \frac{\langle r_i r_j \rangle - \langle r_i \rangle \langle r_j \rangle}{\sqrt{\left(\langle r_i^2 \rangle - \langle r_i \rangle^2\right)\left(\langle r_j^2 \rangle - \langle r_j \rangle^2\right)}} \tag{1}$$

r_i is the logarithmic return defined as $r_i = \ln P_i(t) - \ln P_i(t-\Delta t)$ means the change of $P_i(t)$ during the interval Δt. The correlation coefficient can vary between $[-1,1]$. If $\rho_{i,j} = 1(-1)$ means that two series completely correlated (anti-correlated), while $q_{i,j} = 0$ means they are uncorrelated. Cross-correlations among stocks form a $N \times N$ matrix named as the correlation matrix $\mathbf{C}$. Now, we can describe basics quantities of two main methods, used to extract meaningful economic information from correlation matrix.

Table 1. BIST–100 stocks in October–December 2012

CODE	COMPANY	INDEX	SECTOR	CODE	COMPANY	INDEX	SECTOR	CODE	COMPANY	INDEX	SECTOR
AEFES	ANADOLU EFES	XUSIN	XGIDA	FENER	FENERBAHÇE FUTBOL	XUHIZ	XSPOR	NETAS	NETAŞ TELEKOM.	XUTEK	XBLSM
AFYON	AFYON ÇİMENTO	XUSIN	XTAST	FROTO	FORD OTOSAN	XUSIN	XMESY	NTHOL	NET HOLDİNG	XUMAL	XHOLD
AKBNK	AKBANK	XUMAL	XBANK	GARAN	GARANTİ BANKASI	XUMAL	XBANK	NTTUR	NET TURİZM	XUHIZ	XTRZM
AKENR	AK ENERJI	XUHIZ	XELKT	GLYHO	GLOBAL YAT. HOLD.	XUMAL	XHOLD	OTKAR	OTOKAR	XUSIN	XMESY
AKSA	AKSA	XUSIN	XKMYA	GOLDS	GOLDAŞ KUYUMCULUK	XUMAL	XHOLD	PETKM	PETKİM	XUSIN	XKMYA
AKSEN	AKSA ENERJI	XUHIZ	XELKT	GOLTS	GÖLTAŞ ÇİMENTO	XUSIN	XTAST	PRKME	PARK ELEK. MADEN.	XUSIN	XMADN
ALARK	ALARKO HOLDING	XUMAL	XHOLD	GOODY	GOOD-YEAR	XUSIN	XKMYA	RHEAG	RHEA GİRİŞİM	XUMAL	XHOLD
ANHYT	ANADOLU HAYAT EMEK.	XUMAL	XSGRT	GOZDE	GÖZDE GİRİŞİM	XUMAL	XHOLD	SAFGY	SAF GMYO	XUMAL	XGMYO
ANSGR	ANADOLU SİGORTA	XUMAL	XSGRT	GSDHO	GSD HOLDİNG	XUMAL	XHOLD	SAHOL	SABANCI HOLDİNG	XUMAL	XHOLD
ARCLK	ARÇELİK	XUSIN	XMESY	GSRAY	GALATASARAY SPORTİF	XUHIZ	XSPOR	SASA	SASA POLYESTER	XUSIN	XKMYA
ASELS	ASELSAN	XUTEK	XSVNM	GUBRF	GÜBRE FABRİK.	XUSIN	XKMYA	SISE	ŞİŞE CAM	XUMAL	XHOLD
ASYAB	ASYA KATILIM BANKASI	XUMAL	XBANK	HALKB	T. HALK BANKASI	XUMAL	XBANK	SKBNK	ŞEKERBANK	XUMAL	XBANK
AYGAZ	AYGAZ	XUSIN	XKMYA	HURGZ	HÜRRİYET GZT.	XUSIN	XKAGT	SNGYO	SİNPAŞ GMYO	XUMAL	XGMYO
BAGFS	BAGFAŞ	XUSIN	XKMYA	IHEVA	İHLAS EV ALETLERİ	XUSIN	XMESY	TAVHL	TAV HAVALİMANLARI	XUMAL	XHOLD
BANVT	BANVİT	XUSIN	XGIDA	IHLAS	İHLAS HOLDING	XUMAL	XHOLD	TCELL	TURKCELL	XUHIZ	XILTM
BIMAS	BİM MAĞAZALAR	XUHIZ	XTCRT	IPEKE	İPEK DOĞAL ENERJI	XUSIN	XMADN	THYAO	TÜRK HAVA YOLLARI	XUHIZ	XULAS
BIZIM	BİZİM MAĞAZALARI	XUHIZ	XTCRT	ISCTR	İŞ BANKASI (C)	XUMAL	XBANK	TIRE	MONDİ TİRE KUTSAN	XUSIN	XKAGT
BJKAS	BEŞİKTAŞ FUTBOL YAT.	XUHIZ	XSPOR	ISFIN	İŞ FIN.KİR.	XUMAL	XFINK	TKFEN	TEKFEN HOLDİNG	XUMAL	XHOLD
BOYNR	BOYNER MAĞAZACILIK	XUHIZ	XTCRT	ISGYO	İŞ GMYO	XUMAL	XGMYO	TOASO	TOFAŞ OTO. FAB.	XUSIN	XMESY
BRISA	BRİSA	XUSIN	XKMYA	ISYHO	IŞIKLAR YAT. HOLD.	XUMAL	XHOLD	TRCAS	TURCAS PETROL	XUSIN	XKMYA
BRSAN	BORUSAN MANNESMANN	XUSIN	XMANA	ITTFH	İTTİFAK HOLDİNG	XUMAL	XHOLD	TRGYO	TORUNLAR GMYO	XUMAL	XGMYO
CCOLA	COCA COLA İÇECEK	XUSIN	XGIDA	IZMDC	İZMİR DEMİR ÇELİK	XUSIN	XMANA	TRKCM	TRAKYA CAM	XUSIN	XTAST
CEMAS	CEMAŞ DÖKÜM	XUSIN	XMANA	KARSN	KARSAN OTOMOTİV	XUSIN	XMESY	TSKB	T.S.K.B.	XUMAL	XBANK
DEVA	DEVA HOLDİNG	XUSIN	XKMYA	KARTN	KARTONSAN	XUSIN	XKAGT	TSPOR	TRABZONSPOR SPORTİF	XUHIZ	XSPOR
DOAS	DOĞUŞ OTOMOTİV	XUHIZ	XTCRT	KCHOL	KOÇ HOLDİNG	XUMAL	XHOLD	TTKOM	TÜRK TELEKOM	XUHIZ	XILTM
DOHOL	DOĞAN HOLDING	XUMAL	XHOLD	KLGYO	KİLER GMYO	XUMAL	XGMYO	TTRAK	TÜRK TRAKTÖR	XUSIN	XMESY
DYHOL	DOĞAN YAYIN HOL.	XUMAL	XHOLD	KONYA	KONYA ÇİMENTO	XUSIN	XTAST	TUPRS	TÜPRAŞ	XUSIN	XKMYA
ECILC	ECZACIBAŞI İLAÇ	XUMAL	XHOLD	KOZAA	KOZA MADENCİLİK	XUSIN	XMADN	ULKER	ÜLKER BİSKÜVI	XUSIN	XGIDA
ECZYT	ECZACIBAŞI YATIRIM	XUMAL	XHOLD	KOZAL	KOZA ALTIN	XUSIN	XMADN	VAKBN	VAKIFLAR BANKASI	XUMAL	XBANK
EGEEN	EGE ENDÜSTRİ	XUSIN	XMESY	KRDMD	KARDEMİR (D)	XUSIN	XMANA	VESTL	VESTEL	XUSIN	XMESY
EGGUB	EGE GÜBRE	XUSIN	XKMYA	METRO	METRO HOLDİNG	XUMAL	XHOLD	YKBNK	YAPI VE KREDİ BANK.	XUMAL	XBANK
EKGYO	EMLAK KONUT GMYO	XUMAL	XGMYO	MGROS	MIGROS TİCARET	XUHIZ	XTCRT	ZOREN	ZORLU ENERJİ	XUHIZ	XELKT
ENKAI	ENKA İNŞAAT	XUHIZ	XINSA	MNDRS	MENDERES TEKSTİL	XUSIN	XTEKS				
EREGL	EREĞLİ DEMİR ÇELIK	XUSIN	XMANA	MUTLU	MUTLU AKÜ	XUSIN	XMESY				

3.1 MST Analysis

The MST is a graph, which connects all the nodes (n) via shortest weighted links $(n-1)$ without loops. MST provides us the *sub-dominant ultra-metric distances* among stocks in a direct way to analyze clustering of stocks. This method is known as nearest neighbor single linkage cluster analysis in multivariate analysis [9].

Correlation coefficients matrix elements cannot be used directly as weights due to they do not fulfill the three axioms that define a Euclidean metric. Mantegna provided a function, which transforms cross-correlation among stocks to metric distances [1]. This function is

$$d_{ij} = \sqrt{2\left(1 - \rho_{ij}\right)}, \qquad 0 \le d_{ij} \le 2 \tag{2}$$

where small values mean strong relation between the pair of stocks. If we apply this function to correlation matrix **C**, we obtain a distance matrix **D** that is used to construct MST.

3.2 HT Analysis

The HT is a way to extract economical information to investigate factors affecting elements in system by characterizing complex networks. This tree ranks the data with sub-dominant ultra-metric distances and orders them according to their importance in structure. In this study, in order to obtain advantages of this kind of classification, we used a hierarchical clustering procedure applied correlation matrix, which define similarities between elements. Filtering procedure applied matrix mirrors information stored into hierarchical tree completely. We used in this study ALCA filtering procedure, which was discussed in detail in Tumminello et al.'s work [18].

3.3 Probing Financial Structure
3.3.1 Correlation Coefficients Matrix

Investigated correlation coefficients matrix (**C**) forms a $N \, x \, N$ square. In order to have a sequence, which shows links among stocks without multi-edges and loops and forms $N\left(N-1\right)/2$ edges, we use triangular of this matrix except the diagonal. This matrix completely characterizes correlation coefficients between stocks. In addition, we calculated *probability density function* and *statistical moments*.

The first two moments deal with location, variability, also, aspect of the distribution of the numbers (positions of individual numbers). On the other hand, the third and fourth moments are interested in information about shape of distribution.

3.3.2 Properties of Network

After conversation of correlation matrix into distance matrix and constructing MST, we calculated *normalized tree length* and *degree distribution k* with *power-law fitting* in order to see strength of clusters via distances among stocks. P_k is description of relative frequencies of nodes that have different degrees k. Degree distributions of MST networks show Power-law behavior [5] and it is written as $P_k \, \alpha \, k^{-\gamma}$ where, $-\gamma$ is the *power-law exponent* that is calculated with power-law fitting method (log-log plot of P_k and k). Other measured parameters are; *average clustering coefficient, diameter, average degree, average path length, node betweenness, edge betweenness* and *closeness*.

4. Results

In this section of paper, we introduced results in three sections. First two sections consist of statistical and empirical results of data set provide market condition, movements of stocks and effects of downfall periods. On the other hand, last section offers analyzing ability of stock correlation network concept in portfolio optimization.

4.1 Distribution of Correlation Coefficients

Probability distribution of correlation coefficients was plotted in Figure 1 in order to figure out evolution of market, and *six-month time windows* (t=126 trading days) were created and calculated the statistical moments of each to capture how the correlation coefficients evolve through time. Finally, these data combined and distributions of *2-years* were analyzed in Figure 2.

Correlation coefficients of BIST–100 change in $0.1 < \rho_{ij} < 0.6$ and all of them are larger than -0.1. This distribution similar to *"normal"* and indicates *"up and down in the same direction"* means prices of most stocks in BIST–100 fluctuate towards the same direction [16]. Besides, there is an increase observed in *mean of correlation coefficients* between *May 2011* and *February 2012*.
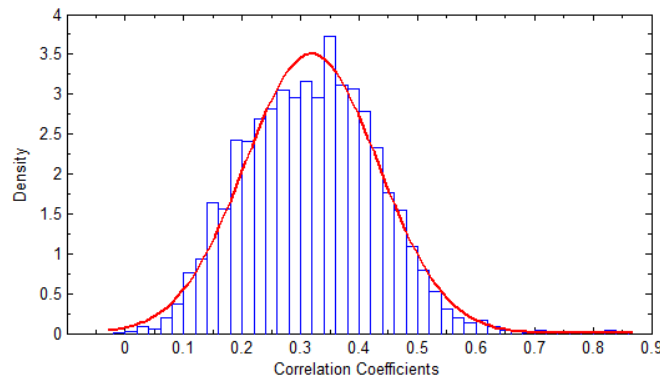
Figure 1. Distribution of correlation coefficients

The *skewness* decreased toward zero, implying that the distribution of the correlations becomes more Gaussian. According to our knowledge, the high values of mean correlation coefficient and the low values of skewness are can be defines as a sign of crisis period [3,7,13,15,16,20]. Mean distances among stocks in MSTs, in other term, normalized tree length, is shown in Figure 3. As well as statistical moments, during the crisis period *(second six-month period)* the tree shrank considerably which means that the distances between the stocks rapidly decreased [13].

4.2 Constructing Financial Structures

Owing to grouping of stocks was not perfect at the branch level, the term *cluster* is described as a subset of a branch [5]. A *complete cluster* contains all the companies in same sector, in practice; however, clusters were mostly incomplete in our networks. In addition, all trees had banks cluster in center of tree. High-degree nodes of banks clusters changed but this fact did not. After the crisis period, hub nodes of clusters became more meaningful, also. Before analyzing economical states of stocks in tree, we computed network properties.

Average clustering coefficient of the network is 0.17. This value shows that mostly nodes of the network pass 1 or 2 triangle through it and show meaningful characteristic in clusters. The *diameter* value calculated from network shows that the distances between nodes are small and they are connected highly correlated. *Average degree* value pointed that this network has some high-degree nodes, but degrees of mostly nodes are approximately 2. Finally, by computing *average path length* value, it was shown that nodes of network locate closely. These properties showed us BIST–100 stock correlation network display "small world" and "scale-free" characteristic [16,17].
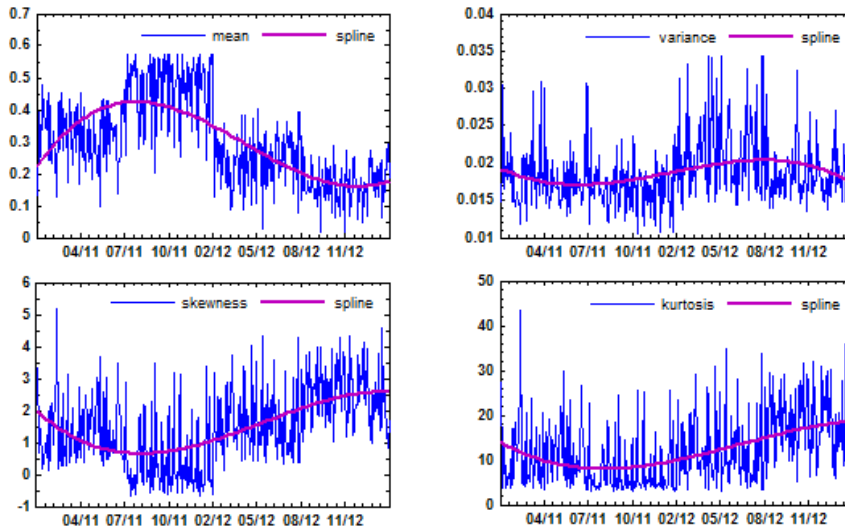
Figure 2. Statistical moments of correlation coefficients

Node properties of network provided us to obtain that some nodes had high-degree compared to others, such as *12, 14, 28, 37, 51, 59, 81, 85,* and *99*. These nodes, especially, *51* acted like a hub node and gave direction to clusters and namely to network. *In-degrees* are incoming edges to nodes and *out-degrees* are outgoing. some of high-degree nodes had high in-degrees. This property provided to understand significance of these nodes in network for connecting other nodes and clusters. Then, *node betweenness* and *closeness centrality* measures helped to
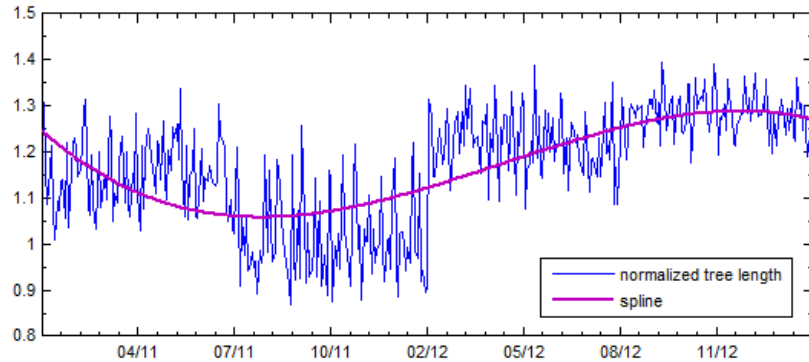
_____

Figure 3. Normalized tree length of MSTs

analyze role of nodes in network. Nodes with a high betweenness centrality were interesting, because, they control information flow in a network. It was seen that high-degree nodes have high-value of betweenness and closeness. It can be said that high valued nodes have more important roles than others have.

In addition, *closeness centrality* is a value shows how fast information spreads from the node to other nodes in the network. As well as node betweenness centrality, high-degree nodes have higher values. This means nodes that have higher closeness measure are more capable than others at connecting nodes and have more important role in clusters than other nodes in network.

Betweenness is a measure of the number of times an edge occurs on a geodesic. In addition, it is the total amount of flow it carries counting flow between all pairs of nodes using this edge. Shortly, edge betweenness is a way to show important edges. In MST, most high-betweenness valued edge is *51-99* or *99-51* edge according to graph. Deeply, we observed that edges which connecting hub nodes had higher values than other edges.

In order to get more information from the MST, the *distribution of the degree k*, which is the number of links that are connected to each node of the network, was examined. After that, the distribution for MST was plotted with the function in Equation (8). The log-log plot of the topological distribution is shown in Figure 4.

In a scale-free network, degree distribution shows a power-law characteristic. This kind of network has some high-degree, but mostly, low-degree nodes. From the point of price fluctuation, most of the stocks are at the same level, but, some ones have a *"higher status"* and *"stronger market influence power"* [16].

The degree distribution for the average data MST showed a power-law behavior with exponent $(\gamma = 2.68 \pm 0.1)$. Power-law behaviors with similar exponents were observed in previous studies, also, such as, Vandewalle et al., Bonanno et al., Onnela et al., Garas and Argyrakis, Huang et al., Tse et al. and L.S. Junior [5,8,10,11,13,16,17].

When power-law fitting procedure was applied on other periods, it was seen that power-law exponent value is decreased to $(\gamma = 2.11 \pm 0.1)$, in falling period. A similar situation was observed in Onnela et al.'s research that investigated Black Monday (1987) period [5].
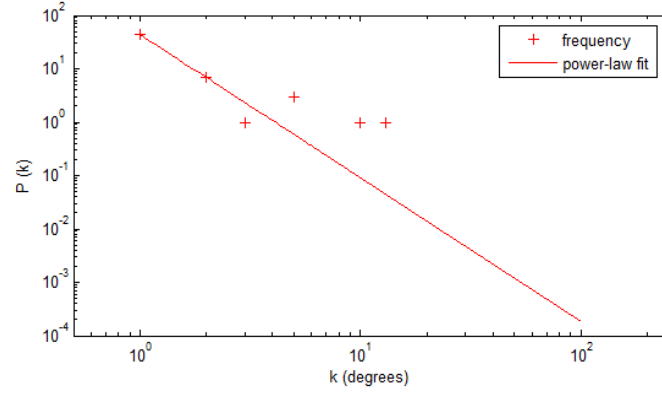

Figure 4. Degree distribution-k of MST

4.2.1 Probing MST

First, we colored the MST according to their *sectors* in Figure 5. Generally, stocks in same sector were in the same clusters or connected to each other in direct relationship. This behavior showed that sectoral information of stocks is one of the
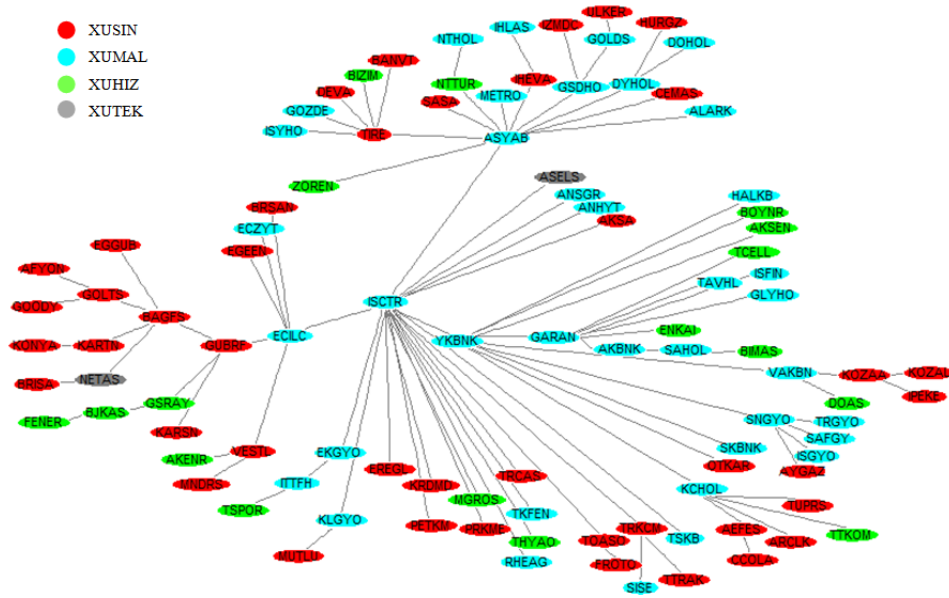
Figure 5. View of MST (sector) for 2011–2013

ways used in analyzing classification of stocks. Most important info gained from this graph is that *financial* stocks were central nodes of network mean they give direction to MST, so, BIST–100.

Thereafter, grouping of stocks would be investigated as four big clusters to have more detail information about states of stocks in MST. The MST includes *sub-sector* information of stocks can be seen in Figure 6.

Central Cluster: Central zone of MST consists of three high-degree (hub) *XBANK (Banks)* nodes: *ISCTR, ASYAB* and *YKBNK*. It is clearly seen in graph that node *51 (ISCTR)* is root of MST and reference stock for other clusters. It is placed in the central position of *XBANK (Banks)* cluster, too. Besides, Node *99 (YKBNK)* is in a considerable condition according to its degree and it has strong relationships with other Bank nodes. Banks cluster consists of *nine* nodes and it is a complete cluster. When weights of edges connecting nodes were investigated, it was seen that distances are short, means: Banks are in strong contact.
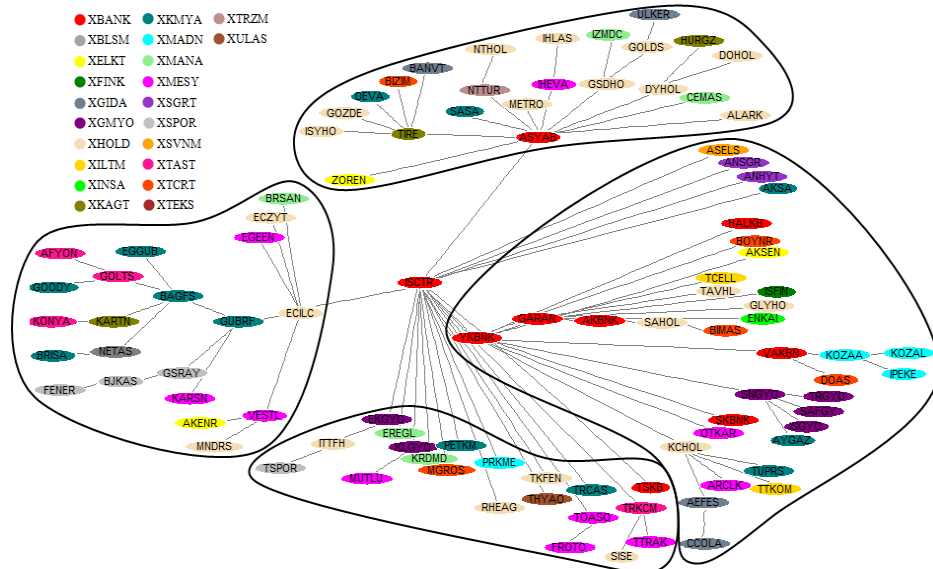

Figure 6. View of MST (sub-sector) for 2011–2013

North Cluster: This cluster holds stocks of three sectors, *financial*, *service* and *industrial*. Nodes in these cluster spread via second high-degree node *12 (ASYAB)*. When studied this cluster deeper, it was seen that most of stocks are in *XHOLD (Holdings)* sub-sector. Otherwise, some unexpected situations were sighted in this cluster. Some assets connected to their main company, such as, *70 (NTHOL)* and *71 (NTTUR)*. These stocks are sub-assets of *NET Holding*. Besides, *IHLAS Holding* and its sub-assets *48 (IHEVA)* and *49 (IHLAS)* were connected. Similar condition was observed for *DOGAN Holding*, also. *26 (DOHOL)* and *47 (HURGZ)* were connected to *27 (DYHOL)*. This case showed that these stocks ignored their sub-sectoral act and set up a relation according to their corporative

state. On the other hand, these stocks connected to *12 (ASYAB)* that was *"hub"* of North Cluster means that connections between sub-assets and main company stocks are strong.

West Cluster: This cluster has all kind of sectors: *financial*, *industrial*, *service* and *technology*. On the other hand, this cluster can be named as Industrial Cluster because of that most stocks in cluster are industrial stocks. Node *28 (ECILC)* is reference stock for this cluster. As well as North Cluster, *28 (ECILC)* and *29 (ECZYT)*, which are sub-assets of *ECZACIBASI Holding*, had direct connection. In addition, most of *XKMYA (Chemical, Petroleum and Plastic)* stocks and *XTAST (Non-metal Mineral Products)* stocks were found in this cluster. Except *92 (TSPOR)*, all *XSPOR (Sports)* stocks existed in this cluster and connected directly. This means *18 (BJKAS)*, *35 (FENER)* and *44 (GSRAY)* moved together and were affected from each other, briefly.

East Cluster: As well as West Cluster, this cluster had all sectors, either. Interesting thing is *8 (ANHYT)* and *9 (ANSGR)*, which are in *XSGRT (Insurance)* sub-sector, had direct connection to root node *51 (ISCTR)* that is main company of them. The other part of cluster spread via another financial sector node *99 (YKBNK)* and this means East Cluster was directed by financial nodes like West Cluster and North Cluster. Another attractive observation in this cluster is attaching of sub-assets of *SABANCI Holding*, *3 (AKBNK)* and *77 (SAHOL)*. Only two beverage – *XGIDA (Food, Beverage)* – company of BIST–100, *1 (AEFES)* and *22 (CCOLA)* existed in this cluster and had direct connection.

Besides, most of *XGMYO (Real Estate Investment Trusts)* stocks, *53 (ISGYO)*, *76 (SAFGY)*, *81 (SNGYO)*, *89 (TRGYO)* were in this cluster and formed a sub-cluster, except *13 (AYGAZ)*. Furthermore, three of four *XMADN (Mining)* stocks existed in here and formed a complete cluster. Interestingly, all stocks in this sub-cluster, *50 (IPEKE)*, *62 (KOZAA)* and *63 (KOZAL)* are sub-assets of *KOZA-IPEK Holding*. Moreover, *10 (ARCLK)*, *13 (AYGAZ)*, *59 (KCHOL)*, *72 (OTKAR)*, *95 (TUPRS)* and *99 (YKBNK),* which are sub-assets of *KOC Holding* existed in this cluster. These observations showed that, corporative condition determines grouping of stocks as well as sub-sectoral classification.

South Cluster: Nodes in this cluster have direct connection with the root node of MST. This cluster had all type of sectors excluding *XUTEK (Technology)*, too. Remarkable facts in this cluster are that two auto-manufacturing stocks, – *XMESY (Metal Products, Machinery)* – which are sub-assets of *KOC Holding*, *36 (FROTO)* and *87 (TOASO)* and also, two glass-product companies *79 (SISE)* and *90 (TRKCM)* connected by one same edge. Although, *SISE* was in *XHOLD (Holdings)* sub-sector and *TRKCM* was *XTAST (Non-metal Mineral Products)*, they ignored their sub-sectoral behavior and product manufacturing defined states in cluster.

Finally, in order to see movement of stocks in BIST–100 according to *average daily trading volume* and *liquidity*, stocks were colored using their indices. It is a fact that *XU030*, *XU050* and *XU100* are defined by these quantities and it is a

very effective way to grouping stocks. Figure 7 is a demonstration of BIST–100 with indices.

It was seen that central nodes consisted of *XU030* index stocks in MST; these nodes are liquid and have higher trading volumes than others have. When it was thought that central nodes were financial stocks, especially Banks, it can be said that these stocks are in very important place at BIST–100. As expected, the other *"hub"* nodes, which spread nodes to clusters, in other words, connects to root node, were situated in *XU050* index. On the other hand, *XU100* stocks were existed in side parts of MST. As a result, it was obvious that the most liquid stocks are laid in the center of the tree and become the reference for many other stocks that are relatively more illiquid [3].
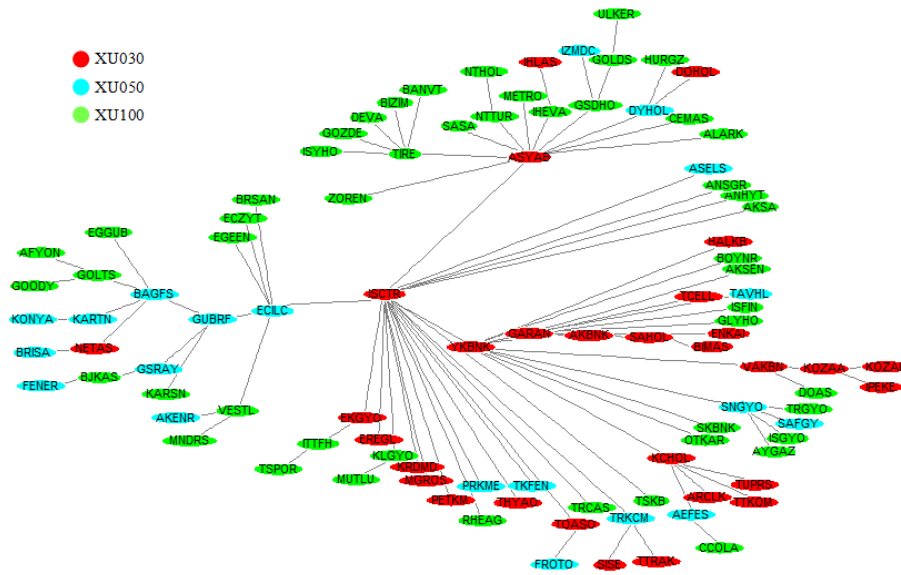
Figure 7. Classification of stocks by indices

4.2.2 Probing HT

Constructed HT following *ALCA* procedure seen in Figure 8 was colored if least two stocks in same sector existed next to each other. *XBANK (red)* stocks formed a cluster and they had shortest distances among them, also, their cluster was the biggest cluster that has most of one kind stock. This means Bank nodes are in strong contact and move together. *XHOLD (cyan)* stocks spread over HT as well as MST and two *XSGRT (teal)* stocks, *8 (ANHYT)* and *9 (ANSGR)* formed a cluster. Unlike MST, two *XKMYA (olive)* stocks formed a section; *73 (PETKM)* and *88 (TRCAS)*, which are sub-assets of *PETKIM-PETROKIMYA Holding*, specified a connection that was not noticed in MST. Also, unlike MST, *32 (EKGYO)*, *53 (ISGYO)* and *81 (SNGYO)* generated a *XGMYO (maroon)* cluster. Moreover, *XTAST (purple)* stocks; *2 (AFYON)* and *40 (GOLTS)*, another *XKMYA (olive)* and manure stocks; *14 (BAGFS)* and *31 (EGGUB)*, *XGIDA (magenta)*

stocks; *1 (AEFES)* and *22 (CCOLA)* generated clusters. Finally, *XHOLD* stocks formed clusters that are spread to different areas of HT: *39 (GOLDS)* and *43 (GSDHO)*; *59 (KCHOL)* and *77 (SAHOL)*; *26 (DOHOL)* and *27 (DYHOL)*; *54 (ISYHO)* and *65 (METRO)*; *49 (IHLAS)* and *70 (NTHOL)*.

This kind of analysis captured some important points that MST was not, also. In contrast MST, two *XILTM (navy) – Telecommunications* – stocks *83 (TCELL)* and *93 (TTKOM)* created a cluster. These stocks placed in different clusters in MST. Moreover, it was seen that all-four *XSPORT (lime)* stocks existed in same cluster. This means ALCA HT analysis is skillful as well as MST to observe sub-sectoral conditions. Parallel to MST, ALCA HT pointed three *XMADN* (yellow) stocks, which are sub-assets of *KOZA-IPEK Holding*, *50 (IPEKE)*, *62 (KOZAA)* and *63 (KOZAL)*. As more example for this situation, *8 (ANHYT)* and *9 (ANSGR)*, which are sub-assets of *51 (ISCTR)*, were existed next to each other. Besides, *DOGAN Holding* sub-assets *26 (DOHOL)*, *27 (DYHOL)* and *47 (HURGZ)* formed a cluster. It can be said that HT also inspects sub-asset behavior in BIST–100. However, MST has more ability to investigate cooperative states.
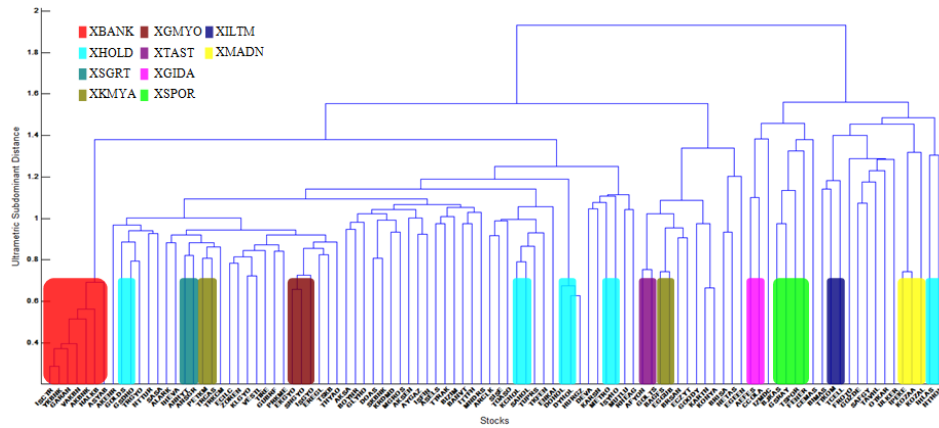

Figure 8. Sub-sectoral view of HT

When HT, seen in Figure 9, was colored according to sectoral information, it was seen that sectoral information is an important agent determines act of stocks. Financial stocks had high-relation in own sector due to *XUMAL* stocks had the shortest distances in each cluster formed in HT. Moreover, while *XUSIN* stocks were investigating, it was seen that generally, these stocks moved together and they had second shortest distances after financial ones. Although, *XUHIZ* stocks moved together, it was seen that they had longest distances between them and the other sector stocks in clusters. On the other hand, *XUTEK* stocks act insignificant in HT.

In Figure 10, in order to probe liquidness of stocks, ALCA HT was divided into four clusters according to branching of HT to examine stocks in detail by taking advantage of their indices. It was seen that same indices had strong relations

but some exceptions. As first cluster were probed, it was clearly seen that all stocks in this cluster consisted of *XU030 (most liquid)* assets and they were in central point of BIST–100 because of having shortest distances. When second cluster were investigated, it was seen that liquid stocks split from illiquid ones and moved together. Third cluster of HT was formed of mostly illiquid stocks moved related. This proved that same kind of stocks had strong connection according to index or liquidness as well as sectors. Thus, fourth cluster kept stocks that had longest distances. Although, liquid and illiquid stocks existed in this cluster, they were separated and were grouped according to their indices, so, liquidness.
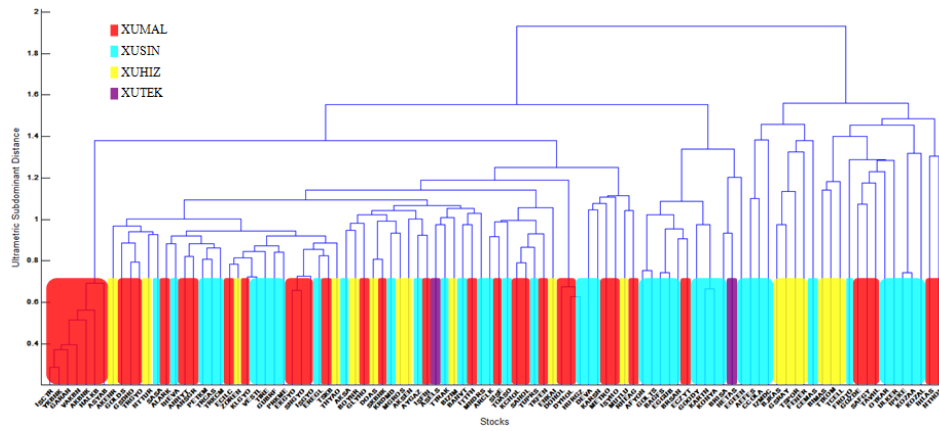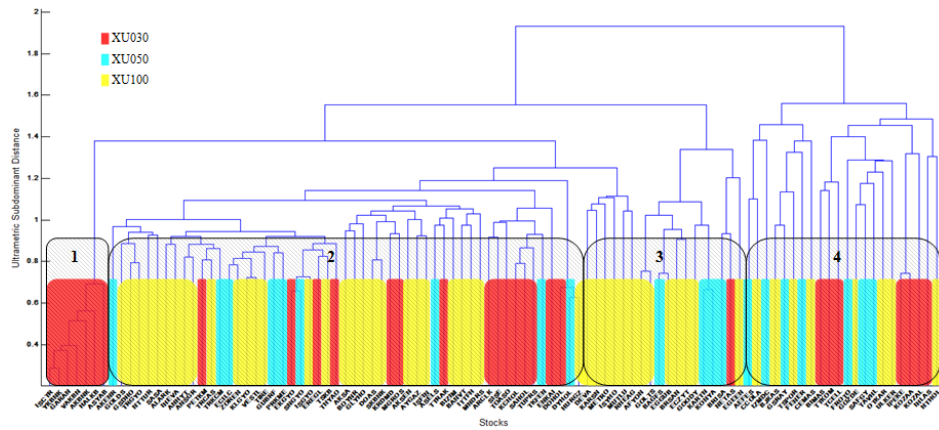

Figure 9. Sectoral view of HT

Figure 10. Index view of HT

4.3 Portfolio Optimization

After completing MST and HT analysis, we indented to notice effectiveness of *"Stock Correlation Network"* for risk management of market by paralleling with *"Modern Portfolio Theory"*; also, with taking advantage of this kind of comparing,

it would be provided to see accuracy of MST and HT for using as a investment guide or not.

It is one of the most important, influential economic theories dealing with finance, and investment suggests that it is possible to construct an *"efficient frontier"* of optimal portfolios, offering *the maximum possible expected return* for a given level of risk. It suggests that it is not enough to look at the expected risk and return of one particular stock. By investing in more than one stock, an investor can reap the benefits of diversification, particularly a reduction in the riskiness of the portfolio [25]. Figure 11 is *mean-variance-efficient frontier graphs* for six-month periods of BIST–100.

As the figure investigated, it was clearly seen that *mean-variance-efficient frontier* graphs adapted with *statistical moments* graph and *normalized tree length* graph. While these graphs were investigating, it had been observed that second and third period of data scale had been affected negatively because of crisis. Likewise, MPT showed that risk of these periods were greater than the other periods. Hence, MST concept is a *good-guide for risk management*.
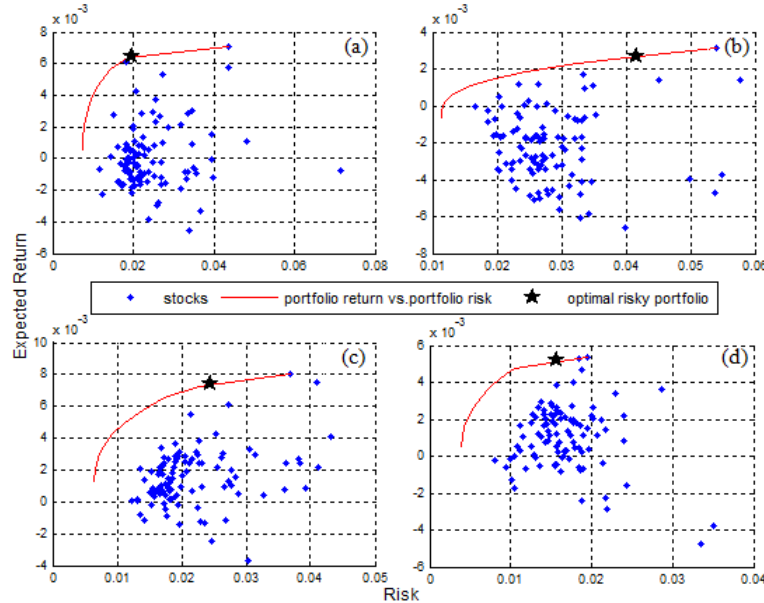

Figure 11. Efficient frontiers of six-month periods of BIST-100; a) January–June 2011, b) July–December 2011, c) January–June 2012 and d) July–December 2012

Following, MPT was investigated by plotting *mean-variance-efficient frontier* graph for all data scale. Plotted graph showed that BIST–100 stocks existed mostly in same risk rate area, except some ones. Well, which stocks moved out of common risk area and which stocks were more riskless. Yet more important,

these acts and states can be guessed from MST or/and HT? Figure 12 was helped to answer these questions.

Firstly, to examine risks of stocks in detail, stocks, which are out of range, must be determined. Standard deviation $\left(\sigma = 0.0059\right)$ and mean $\left(\mu = 0.0235\right)$ of risks were computed to observe extraordinary stocks and their volatilities, which were spotted by using plot.

In first place, when stocks, which are more riskless, were investigated, it was seen that these stocks were mostly financial stocks and took place closely to *"hub"* stocks. More generally, *ALARK*, *NTHOL*, *BIMAS*, *ECILC* and *ISGYO* had direct connection with bank cluster. Besides, it was seen that *ECILC* had direct connection hub stock *ISCTR*. When more risky stocks were investigated, it was seen that these stocks existed far away from hub stocks, especially *ISCTR*. On the other hand, interestingly, most of these stocks existed in West Cluster, which consisted of mostly *XKMYA* and *XTAST*, so, *XUSIN* stocks. In addition, these stocks had behaved meaningless according to their sectoral and/or sub-sectoral states and had more variety, as can be seen previous parts. Furthermore, it was seen that all sport stocks, which were *XUHIZ* stocks in this cluster, were riskier. Although, *TSPOR* was in South Cluster, it was risky, too. It was only sport stock in South Cluster and it behaved differently from the other stocks in own sub-sector. On the other hand, *CEMAS* and *GOZDE*, which were in North Cluster, had more risk rates than others have. When these stocks were probed, it was seen that these stocks had close relation with hub stocks. However, while ALCA was investigated, it was seen that *ultra-metric sub-dominant distances* of these stocks are much and close to *XSPOR* stocks, which were more risky.
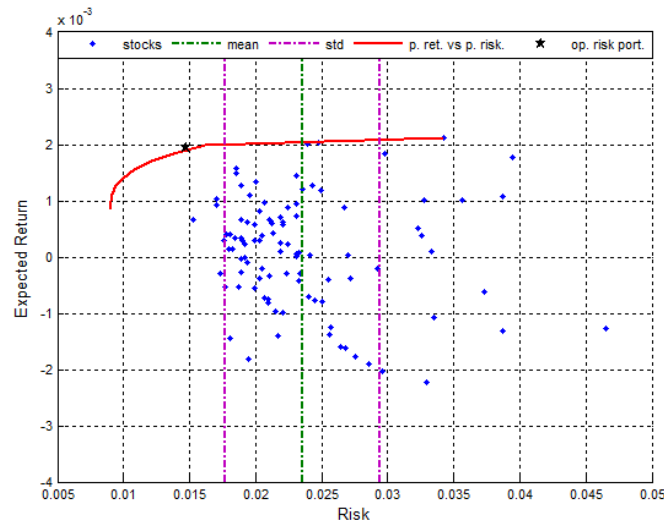

Figure 12. Efficient frontier graph of BIST-100

Last two risky stocks stated in East Cluster. One of them, *SAFGY* had longer edge-weight than the other *XGMYO* stocks in own sub-cluster. Otherwise, it was observed that it was far away from the other sub-cluster stocks in ALCA and connected with *GOZDE*, which was another risky stock. The other risky stock, *IPEKE*, was situated in a *XMADN* sub-cluster, which connected directly to hubs and all nodes consisted of sub-assets of same holding. However, this stock had long sub-dominant distance in ALCA. Farther, all stocks in this cluster formed a cluster, which had long sub-dominant distances in ALCA, too. To prove effectiveness of ALCA for risk management, it was taken look at risk rates of the other stocks in this cluster and seen that its risk rates were more than mean value.

5. Summary and Conclusions

Correlation based networks can be obtained from financial markets by investigating time series. *"Filtering procedure"* applied correlation matrix is created by the returns of a portfolio of financial assets provided to obtain distance matrix which selects a topological space for the stocks traded in a market. Therefore, in this study, it was showed how to associate a correlation matrix with a hierarchical tree and correlation based trees or graphs.

Investigating data scale for *N=100* stocks was *T=506* days (2011–2013) and opening prices of stocks used as time series. *Distribution of correlation coefficients, statistical moments* and *normalized tree length* of matrixes suggested us effects of crisis periods in portfolio and moving prices of most stocks on the same way. We investigated some network properties and *distribution of the degree k*. These calculations and *power-law* behavior seen at average MST $(\gamma = 2.68 \pm 0.1)$ provided us to understand that BIST–100 network shows a *scale-free* characteristic; some of nodes are *"hub"*, which has *"higher status"* and *"stronger market influence power"*. Then, *power-law degree distribution* applied to the other periods of scale and seen that in crisis period, it decreases to *2.0* level. This observation showed that again MST concept is a good *risk-management* concept.

Next, sectoral states of MST was investigated and seen that financial stocks formed in center of network and control nodes behaviors. After that, *sub-sectoral* behavior was investigated by dividing network to four general clusters. Stock *51 (ISBNK)* existed in central position *(root)* of network and had strong relation with other *XBANK* stocks, which generated complete cluster whose all stocks have short distances and act as *"hub"* nodes. This means banks are in strong contact. The other *high-degree* stocks were *12 (ASYAB)*, *28 (ECILC)* and *99 (YKBNK)* that form other clusters. Briefly, financial stocks have important role in network because of the other nodes and clusters spreading via these nodes. Generally, stocks generated clusters according to their *sectoral* and *sub-sectoral* information. However, sometimes, stocks ignored *sub-sectoral* states and acted according to their *corporative* situations. In addition, in some conditions, stocks moved according to *product manufacturing factor. Average daily trading volume* and *liquidity*

investigated by grouping stocks according to *indices* and obtained that *XU030* stocks in the center, *XU050* and *XU100* stocks spread to sides of MST in order. Therefore, it was seen that the most liquids exist in central position and become *reference* for the other stocks.

Again, distance matrix was used to construct Hierarchical Tree (HT). *ALCA* filtering procedure used to clean matrix to create dendrogram because of its stability. Firstly, *sub-sectoral* states were investigated and seen that all bank nodes form cluster in shortest *ultra-metric sub-dominant distances*. On the other hand, *ALCA* captured two Telecommunications stocks together. This means that HT is successful as well as MST and shows MST cannot. However, although, HT concept shows *sub-asset* acting, it was observed that MST is more skillful for this task. While a sectoral study was being applied to HT, seen that *XUMAL* stocks had shortest distances among them. Furthermore, *XUSIN* stocks had short distances means moved together. This clarified those financial stocks in strong connection and control information flow. *Liquidness* was also investigated by dividing HT into four clusters and it was observed most *liquid* stocks *(XU030)* had *shortest distances* and more liquid stocks separated from *illiquid* ones. So, stocks moves according to *indices* and *liquidness* situations as well as *sectors* and *sub-sectors*.

Lastly, *Modern Portfolio Optimization* used to analyze effectiveness of *financial network technique*. *Mean-variance-efficient frontier graphs* were showed that risk rate of portfolio increased as well as *normalized tree length*, in second period. Thus, MST concept is very successful for *risk management*. MPT for average MST was investigated to define risky stocks. It was clearly seen that *riskless stocks* took place in *central points* of MST or close to *"hubs"*. *Risky stocks* were generally existed far from *"hubs"*, especially the root, *51 (ISBNK)*. In addition, stocks, which ignored *sectoral* and *sub-sectoral* behavior, had more risk than others have. Although, some stocks were closely to *"hubs"*, they were risky. At this point, HT was used to analyze them and observed that these stocks have long *sub-dominant distances* and form clusters own risky ones. With this last part, it was showed that MST and HT could be used as a *portfolio risk management tool*, briefly.

REFERENCES

[1] **R.N. Mantegna, R.N. et al. (1999),** *Hierarchical Structure in Financial Markets*; Eur. Phys. J. B. 11; 193–197;
[2] **G. Bonanno, et al. (2000),** *Taxonomy of Stock Market Indices;* Phys. Rev. E. 62, 7615–7618;
[3] **H. Situngkir, et al. (2005),** *On Stock Market Dynamics through Ultrametricity of Minimum Spanning Tree;* Bandung Fe Institute, Indonesia;
[4] **E. Kantar, et al. (2011),** *Investigation of Major International and Turkish Companies via Hierarchical Methods and Bootstrap Approach*; Eur. Phys. J. B. 84, 339–350;
[5] **J.P. Onnela, et al. (2003),** *Kanto, Dynamics of Market Correlations: Taxonomy and Portfolio Analysis;* Phys. Rev. E. 68, 6110–6122;

Yusuf Yargı Baydilli, Şafak Bayir, İlker Türker

[6] **G. Bonanno, et al. (2004),** *Networks of Equities in Financial Markets*; Eur. Phys. J. B. 38, 363–371;

[7] **R. Coelho, et al. (2007),** *The Evolution of Interdependence in World Equity Markets: Evidence from Minimum Spanning Trees;* Phys. A. 376, 455–466;

[8] **N. Vandewalle, et al. (2001),** *Non-random Topology of Stock Markets;* Quant. Finance. 1 ; 372–374;

[9] **G. Bonanno, et al. (2001),** *High-frequency Cross-correlation in a Set of Stocks;* Quant. Finance. 1 , 96–104;

[10] **L.S. Junior (2012),** *Pruning a Minimum Spanning Tree*; Phys. A. 391, 2678–2711;

[11] **G. Bonanno, et al. (2003),** *Topology of Correlation-based Minimal Spanning Trees in Real and Model Markets;* Phys. Rev. E. 68 , 1–4;

[12] **S. Miccichè, et al. (2003),** *Degree Stability of a Minimum Spanning Tree of Price Return and Volatility*; Phys. A. 324 , 66–73;

[13] **A. Garas, et al. (2007),** *Correlation Study of the Athens Stock Exchange;* Phys. A. 380 , 399–410;

[14] **T. Heimo, et al. (2007),** *Spectral and Network Methods in the Analysis of Correlation Matrices of Stock Returns;* Phys. Stat. Mech. Its Appl. 383 , 147–151;

[15] **M. Eryiğit, et al. (2009),** *Network Structure of Cross-correlations among the World Market Indices;* Phys. A. 388; 3551–3562;

[16] W.Q. Huang, et al., A network analysis of the Chinese Stock Market, Phys. A. 388 (2009) 2956–2964.

[17] **C.K. Tse, et al. (2010),** *A Network Perspective of the Stock Market;* J. Empir. Finance. 17 , 659–667;

[18] **M. Tumminello, et al. (2010),** *Correlation, Hierarchies and Networks in Financial Markets;* J. Econ. Behav. Organ. 75, 40–58;

[19] **E. Kantar, et al. (2011),** *Hierarchical Structure of Turkey's Foreign Trade;* Phys. A. 390, 3454–3476;

[20] **S. Çukur, et al. (2007),** *Cross Correlations in an Emerging Market Financial Data;* Phys. A. 376 , 555–564;

[21] **E. Kantar, et al. (2012),** *Analysis of the Effects of the Global Financial Crisis on the Turkish Economy, Using Hierarchical Methods*; Phys. A. 391, 2342–2352;

[22] **Associated Press (2013),** *Turkish Stock Market Drops Sharply;* http://finance.yahoo.com/news/Turkish-stock-market-drops-apf-1945962593.html (accessed July 30, 2013);

[23] **G. Wearden (2013),** *Debt Crisis Sends Financial Markets into Turmoil –* Monday 8 August 2011, The Guardian. (2013); http://www.theguardian.com/business /2011/aug/08/stock-market-turmoil-us-downgrade-eurozone-crisis (accessed July 30, 2013);

[24] Bourse Istanbul, Constituent Equities of Indices, (2013). http://borsaistanbul.com/en/data/data/equity-market-data/index-data (accessed June 10, 2013);

[25] *B. McClure, Modern Portfolio Theory: Why It's Still Hip;* Investopedia. (2013) http://www.investopedia.com/articles/06/mpt.asp (accessed July 12, 2013).