Mugla Journal of Science and Technology, Vol 3, No 2, 2017, Pages 104-109

Mugla Journal of Science and Technology

A NEW TOPOLOGICAL MEASURE FOR THE COMMUNITIES OF STOCK MARKET NETWORKS

Ömer AKGÜLLER[1], Sinem ÖCAL[1], Mehmet Ali BALCI [1]*

[1]Department of Mathematics, Faculty of Science, Muğla Sıtkı Koçman University, 48000, Muğla, Turkey
oakguller@mu.edu.tr, sinemocal1992@gmail.com, mehmetalibalci@mu.edu.tr

Received: 31.10.2017, Accepted: 17.11.2017
*Corresponding author doi: 10.22531/muglajsci.348054

Abstract

Systems involving interactive actors such as food networks, scientific quotations, social networks, communications networks, the Internet and stock exchange networks have long been studied by many researchers under the concept of complex systems. Such systems are represented by weighted networks. The intensive connections and relationships between actors play a crucial role in forecasting or risk analysis. In this study; we propose a new approach to measure the hierarchical structure of the globally active stock market network. In this approach we propose, the relationship of 21 different world stock exchange markets to each other is determined by Pearson's correlations. Relevant stock network is based on a certain threshold value. At the same time, a new topological measure is used to characterize the interaction of the nodes of the graphical communities of the stock market, and this measure is examined for the time periods of 2008 global economic crisis.
Keywords: Correlation Networks, Network Construction, Minimum Spanning Tree, Topological Measure

BORSA AĞLARININ TOPLULUKLARI İÇİN YENİ BİR TOPOLOJİK ÖLÇÜM

Öz

Gıda ağları, bilimsel alıntılar, sosyal ağlar, haberleşme ağları, Internet ve borsa ağları gibi interaktif aktörleri içeren sistemler, karmaşık sistemlerin içeriği kapsamı altında pek çok araştırmacı tarafından incelenmiştir. Bu tür sistemler ağırlıklı ağlar tarafından temsil edilir. Aktörler arasındaki yoğun bağlantılar ve ilişkiler, tahmin veya risk analizinde önemli bir rol oynamaktadır. Bu çalışmada, aktif küresel borsa ağının hiyerarşik yapısını ölçmek için yeni bir yaklaşım önerilmiştir. Önerdiğimiz bu yaklaşımda, 21 farklı dünya borsa piyasalarının birbirleriyle ilişkisi Pearson ilişkileri tarafından belirlenmektedir. İlgili hisse senedi ağı belli bir eşik değerine dayanmaktadır. Aynı zamanda, borsa graf topluluklarının tepelerinin etkileşimini karakterize etmek için yeni bir topolojik ölçüm kullanılmaktadır ve bu ölçü 2008 yılı küresel ekonomik krizin zaman dilimleri için incelenmektedir.
Anahtar Kelimeler: Korelasyon Ağları, Ağ Yapılandırması, Minimum Geren Ağaç, Topolojik Ölçüm

1 Introduction

As individuals, we are a unit of different types of social networks and biochemical reactions as biological systems. Networks can be nested objects in the Euclidean space, such as electric power grids, the Internet, highways, public transport systems and artificial neural networks, or they can be described as the structure of acquaintance or partnership between individuals in an abstract space as complex systems. Complex systems are natural or social systems involving many nonlinear associative actors. The necessity of making sense of the phenomena in these systems has led many researchers to use the new models and use the complex system tools used in other branches. The most interesting feature of these systems is the existence of phenomena that cannot be obtained in a simple way or that cannot be clearly predicted from the structure of the system and from the individual interaction of the actors. Mathematical methods have shown that complex systems are effective in proving the existence of coexistence features such as noisy sampling effects, long-term relationships, determinism and flexibility in data evolution, scalability, and criticality [13,21].

Graph theory emerges as a powerful mathematical tool to represent complex systems. Graph theoretical approaches are efficient to determine several characteristics of the complex systems such as the long-term relationships [24], the noise in the data [7], the relationship between inevitability and flexibility in evolution [14], and the criticality [17].

In this study, we analyze the correlation network of globally operating stock exchange markets through the 2008 global economic crisis. To study granular structure of the networks, we focus our method to the Minimum Spanning Tree (MST) structures in the clusters of vertices. To construct the graph representation of the network, we determine edges by the correlation distance based on Pearson Correlation Coefficient of the logarithmic returns of the closure prices. Beside the well-known topological measures of MSTs, we propose a new topological measure based on the vertices in MST. In Section 2, we give basic definitions and theorems on graph theory. In Section 3, we present the methods we used. For the vertex clusters, we used graph communities with high modularity method. The details of the new topological measure we present are studied with certain bounds. The data and the network construction method are also presented in Section 3. In Section 4, we present detailed results that we obtain and in Section 5 we give the discussion.

2 Preliminaries

In this section, we give some preliminary definitions and theorems for the graphs. More details can be found in [5,11,25].

Graphs are the representation of a relation defined on discrete set of objects. A graph G is denoted by the tuple $G=(V,E)$ where $V=\{v_1,\dots,v_n\}$ is the set of vertices and E is the set of edges with the elements $e_k = (v_i, v_j)$. Throughout this study, we only consider finite simple and undirected graphs which means we assume $|V| = n$, the relation is symmetric, there is no $e_k \in E$ such that $e_k = (v_i, v_i)$.

An ordered sequence of the vertices and edges $v_0, e_1, v_1, \dots, v_{n-1} e_k v_n$ is called a walk, and the number of the edges in a walk is the length of that walk. If the vertices and edges of a walk are all distinct, then we this walk is called path. If there is at least one path between all vertices in a graph, then G is called a connected graph. If a path has the same vertex at the endpoints, this path is called a cycle. The cycle with minimum number of edges is called girth of a graph.

The degree of a vertex is the number of the adjacent vertices to that vertex and we denote it by $deg(v_i)$. If $deg(v_i) = 0$, then v_i is called an isolated vertex. If $deg(v_i) = N - 1$ for all $v_i \in V$, then $G = (V, E)$ is called complete graph.

For $V' \subset V$ and $E' \subset E$, the graph $G' = (V', E')$ is called a subgraph of $G = (V, E)$. If $V = V'$, then G' is called a spanning graph of G. A spanning tree with minimum edge weights is called minimum spanning tree (MST).

From the linear algebraic point of view, a graph $G = (V, E)$ can be represented with some matrices. A matrix $A_G = [a_{ij}]$ whose entities are

$$a_{ij} = \begin{cases} 1 & , \ if\ (v_i, v_j) \in E \\ 0 & \ \ otherwise \end{cases}$$

called the adjacency matrix of G. The diagonal matrix $D_G = diag[\deg(v_i)]$ is called the degree matrix of G. The matrix

$$L_G = D_G - A_G$$

is called the Laplacian matrix of G.

The spectrum of L_G tells us some structural characterizations of G as stated in the following theorems.

Theorem 2.1[8]Let L_G be the Laplacian matrix of G. The k multiplicity of the 0 eigenvalue of L_G is equal to the number of components of G.

Theorem 2.2 [8]Let L_G be the Laplacian matrix of G. Let $\lambda_1, \lambda_2, \dots, \lambda_{n-1}$ be the non-zero eigenvalues of L_G. Then, the number of distinct spanning trees of G is equal to

$$t(G) = \frac{1}{n} \prod_{i=1}^{n-1} \lambda_i.$$

3 Method

In this section, we present the method we use to analyze hierarchical structures of the communities of the stock market network of globally operating stock markets.

Graph communities are the cluster of vertices which are connected densely. There are several methods to determine the community structure of a network. These methods can be summarized as Minimum-cut method [18, 19], Hierarchical clustering [16], Girvan-Newman algorithm [20], High modularity [1], and Clique based methods [9,10].

In this study, we use the high modularity method to determine graph communities. High modularity is a maximization problem respect to

$$Q = \frac{1}{2m} \sum_{v_i, v_j} \left(a_{ij} - \frac{\deg(v_i)\deg(v_j)}{2\mathrm{m}} \right) s_i s_j \, , for\ m = |E|$$

modularity of the edges in a graph. This maximization problem is solved via the linear programming method proposed in [1].

3.1 Topological Measures

For the characterization of the MSTs in stock market networks, several topological measures are proposed. For the sake of simplicity, we denote vertices with their indices in the rest of the paper.

The mean correlation measure based on $N \times N$ correlation distance matrix $D = [d_{ij}]$ is defined as

$$L_{MCM} = \frac{2}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} d_{ij},$$

where N is the number of nodes in MST.

Another measure to study the property in the MST is normalized tree length is defined as

$$L_{NTL} = \frac{1}{N-1} \sum_{d_{ij} \in \Omega} d_{ij},$$

where Ω is the set of edges, and $N - 1$ denotes the number of edges present in the MST [15,23].

Characteristic path length is used to quantify the average minimal route between pairs of nodes. For an unweighted MST it is defined by

$$L_{CPL} = \frac{1}{N(N-1)} \sum_{i,j:i \neq j} l_{ij},$$

where l_{ij} is the number of edges in the shortest path between nodes i and j [6].

The mean occupation layer is the measurement of the change in the density of the MST. With the central node v_c whose level is taken as zero, the mean occupation layer is defined as

$$L_{MOL} = \frac{1}{N} \sum_{i=1}^{N} lev(v_i),$$

where $lev(v_i)$ denotes the level of node v_i with respect to v_c [22].

These measures are mainly depended on the edge weights. Beside the edge weight based measures, it is also possible to measure topological structure of MSTs with a method depended on the vertices:

Definition 3.1 The solitude number of a graph $G = (V, E)$ is defined as

$$S(G) = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} iso(G)}{2M},$$

where $|V| = N$, $|E| = M$, and $iso(G)$ is the number of isolated vertices at the subgraph $G' = (V - \{i, j\}, E')$.

The definition of the solitude number of a graph first proposed in [12] and several bounds for the different graph classes are studied.

In this study, we determine lower and upper bounds for the MSTs:

Lemma 3.2 Let $G = (V, E)$ be the graph with girth $g \leq 10$. Then there exists a spanning tree with at least $\frac{g-2}{2g-2}(N-2) + 2$ leaves. Besides, the number of leaves cannot exceed $\frac{7}{16}N + \frac{1}{2}$.

For the proof of Lemma 3.2 and detailed discussion we refer readers to [4].

Theorem 3.3 Let $G' = (V, E')$ be the MST of a graph $G = (V, E)$ with girth $g \leq 10$. Then

$$\frac{g-2}{2g-2}(N-2) + 2 \leq S(G') \leq \frac{7}{16}N + \frac{1}{2}.$$

Proof.In MSTs, the solitude number is directly equal to the number of leaves. That is, when the junction vertex removed, we directly add $S(G')$ with $\frac{l_i}{2M}$ up, where l_i is the number of leaves adjacent to the junction i. Hence the bounds are directly obtained by using Lemma 3.2. This concludes the proof. □

3.2 Data and Network Construction

In this study, a financial network of globally operating stock markets are modelled by a simple undirected graph G = (V, E), where V is the set of stock markets and E are the edges determined by the Pearson Correlation amongst the markets of Holland (AEX), All Ordinaries (AORD), Austria (ATX), Belgium (BFX), India (BSESN), Brazil (BVSP), France (FCHI), Germany (GDAXI), USA (GSPC) and(GSPTSE), Hong Kong (HIS), Indonesia (JKSE), Malaysia (KLSE), South Korea (KS11), Argentina(MERV), Mexica (MXX), Japan (N225), New Zeeland (NZ50), Spain (SMSI), Singapore (STI), and Taiwan (TWII).

The data we used is obtained from the daily logarithmic return of the closure price of each market between the dates from 01.01.2005 to 31.12.2014 as including the global economic crisis. To show the efficiency, we divide scale of our analysis into three subintervals of pre-crisis 01.01.2005-31.12.2007, crisis 01.01.2008-31.12.2011, and post-crisis 01.01.2012-31.12.2014.

For the daily closure price Cl_i of the i-th stock exchange market, the daily logarithmic return R_i is calculated as

$$R_i = \log Cl_{i+1} - \log Cl_i.$$

The relation between the logarithmic return of the closure prices can be determined by the Pearson Correlation Coefficient

$$\rho_{ij} = \frac{< R_i R_j > - < R_i >< R_j >}{\sqrt{(< R_i^2 > - < R_i >^2)(< R_j^2 > - < R_j >^2)}},$$

where $< .. >$ is a temporal average performed on the trading days. It is well known that ρ_{ij} varies between -1 and 1, that is, $\rho_{ij} = -1$indicates the maximum negative correlation while $\rho_{ij} = 1$ indicates the maximum positive correlation. To avoid negative weights on edges, we introduce a distance based on ρ_{ij} by $d_{Corr}(i,j) = \sqrt{2(1-\rho_{ij})}\big/2$.

Since ρ_{ij} varies between -1 and 1, it is straightforward to see that $d_{Corr}(i,j)$ varies 0 and 1.

For the network construction, we follow the threshold method for correlation networks presented in [2,3].

With the empirically chosen threshold value ThV, we form edges by following the formation rule

$$(i,j) \in E \;\boldsymbol{iff}\; d_{Corr}(i,j) \leq ThV.$$

Initially we start by a complete graph by choosing $ThV = 1$. Then, we decrease ThV by $1/h$, and repeat the formation rule. At certain point between 0 and 1, there is a ThV such that graph becomes with two components. That is, we choose the empirical threshold value which makes graph with one component and optimally many edges. To control the number of connected components we use Theorem 2.1. For the computational complexity, we refer [3] and for the disparity measure of the networks constructed via this threshold value, we refer [2].

In Figures 1-3, we present the matrices of d_{Corr} values of pre-crisis, crisis, and post-crisis periods with vertices numbered as the alphabetical order of stock exchange market tickers. Ticks in the each axes represents the vertex number.

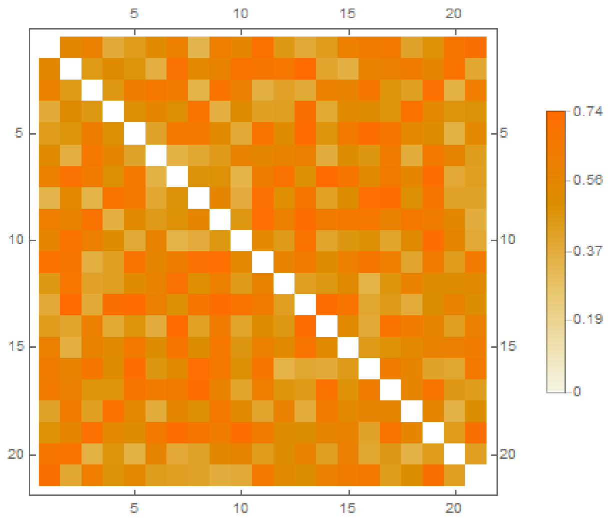

Figure 1. The matrix of $\boldsymbol{d_{Corr}}$ values in pre-crisis period

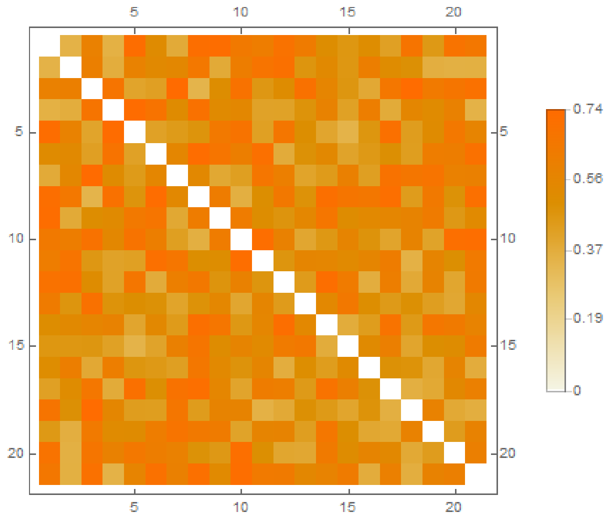

Figure 2. The matrix of $\boldsymbol{d_{Corr}}$ values in crisis period

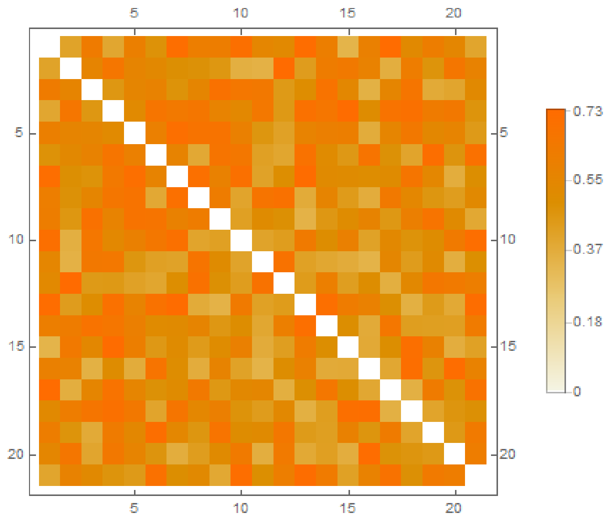

Figure 3. The matrix of $\boldsymbol{d_{Corr}}$ values in post-crisis period.

4 Results

E For $h = 10000$, the threshold values are obtained as 0.6971 for pre-crisis period, 0.6933 for crisis period, and 0.6939 for post-crisis period. Resulting networks and community structures on them are presented in Figures 4-6.

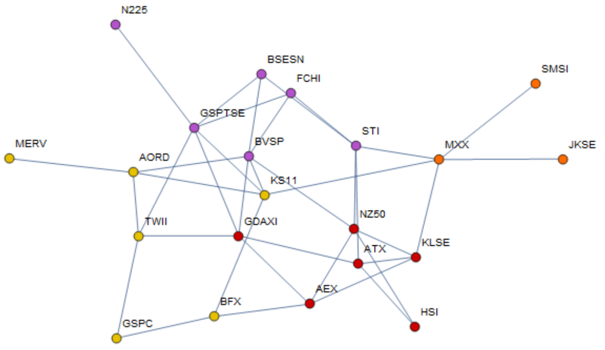

Figure 4. The resulting network in pre-crisis period. The graph communities are represented with different colors.

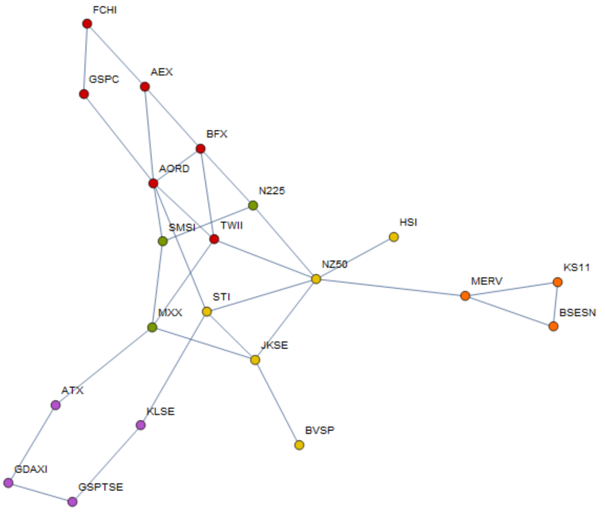

Figure 5. The resulting network in crisis period. The graph communities are represented with different colors.

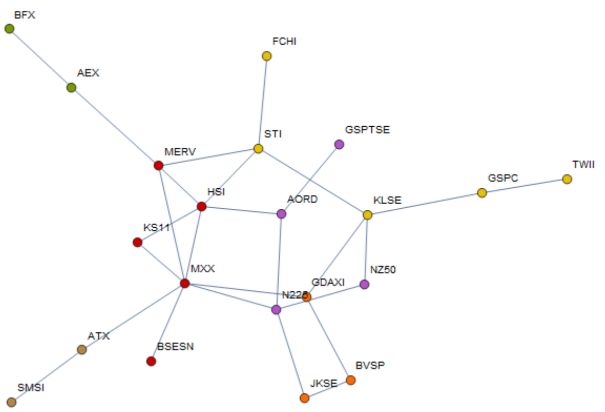

Figure 6. The resulting network in post-crisis period. The graph communities are represented with different colors.

In the pre-crisis period the network has 4 communities, in the crisis period the network has 5 communities and in the post-crisis period the network has 6 communities.

To determine hierarchical structure in communities, we construct complete graphs with edge lengths determined by d_{Corr}. The MSTs in each community represent the hierarchies. In Figures 7-9, we represent MSTs in each community.

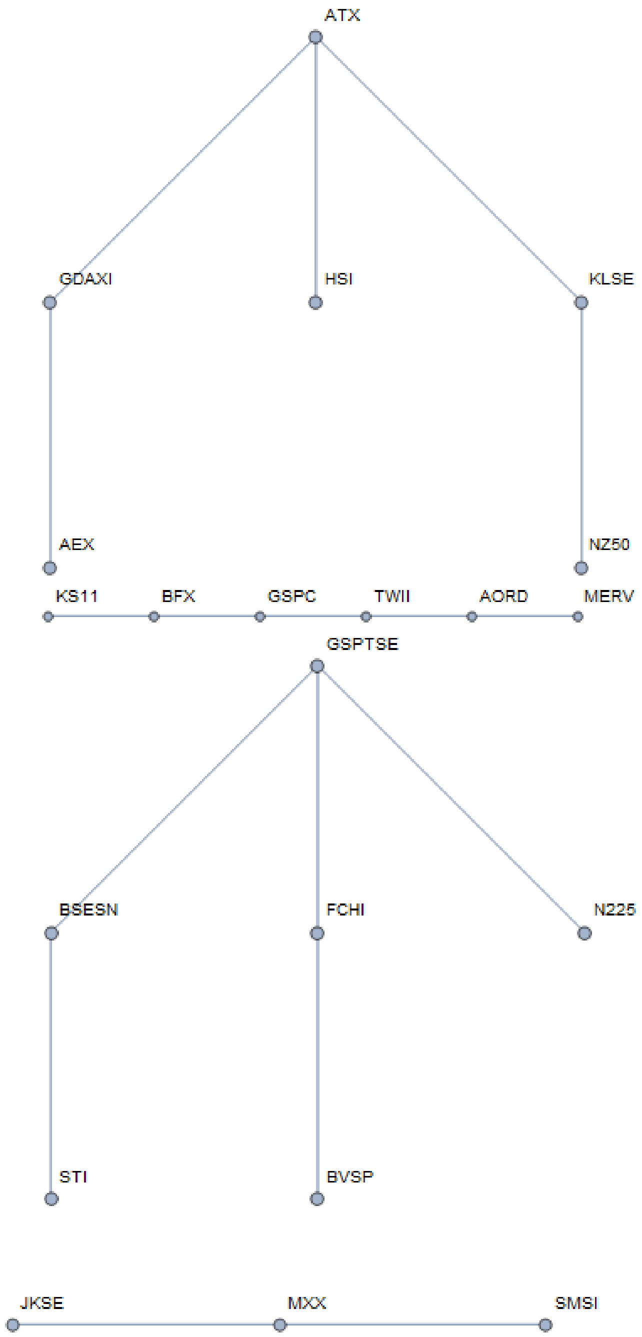

Figure 7. MSTs of communities in pre-crisis period

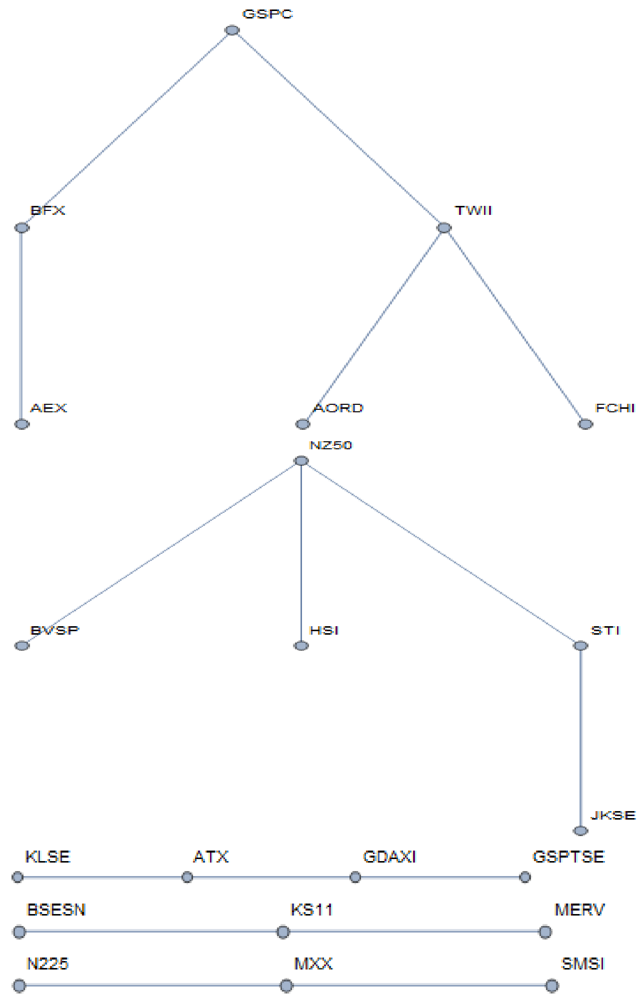

Figure 8. MSTs of communities in crisis period

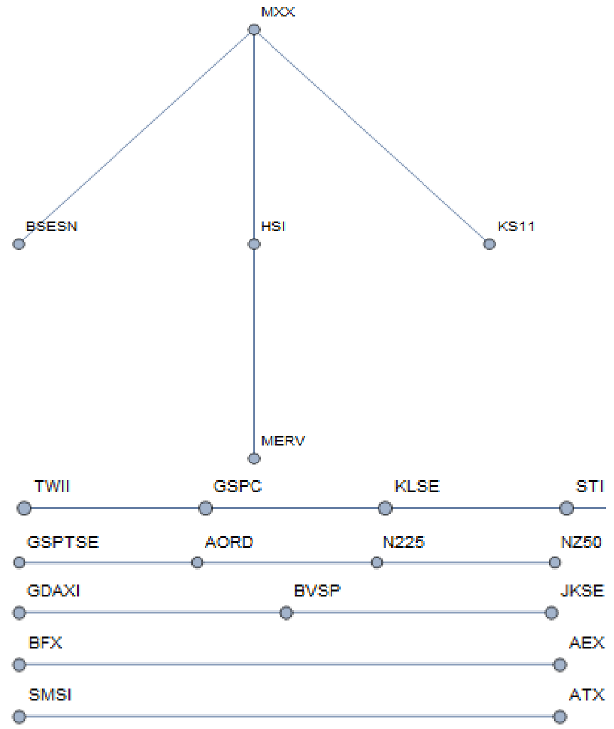

Figure 9. MSTs of communities in post-crisis period

Structural changes in the hierarchies in communities can be seen in the figures. However, these structural changes are not enough to characterize the impacts of the global economic crisis. Therefore, the computational results of the topological measures are presented in Table 1-3.

Table 1. Topological measures for pre-crisis period

	Communities				
Measure	1	2	3	4	Mean
L_{MCM}	0.22	0.23	0.23	0.46	0.286
L_{NTL}	1.37	1.38	1.37	1.38	1.375
L_{CPL}	1.92	3.76	5.56	28.8	10.01
$S(G')$	0.93	0.8	0.93	1	0.916

Table 2. Topological measures for crisis period

	Communities					
Measure	1	2	3	4	5	Mean
L_{MCM}	0.23	0.27	0.34	0.46	0.46	0.352
L_{NTL}	1.37	1.37	1.37	1.37	1.38	1.373
L_{CPL}	1.92	4.39	8.48	17.8	18.62	10.24
$S(G')$	0.933	1	1	1	1	0.986

Table 3. Topological measures for post-crisis period

	Communities						
Measure	1	2	3	4	5	6	Mean
L_{MCM}	0.27	0.27	0.34	0.46	0.69	0.69	0.45
L_{NTL}	1.37	1.38	1.37	1.38	1.39	1.38	1.378
L_{CPL}	1.57	3.07	6.28	13.4	40.1	41.2	17.7
$S(G')$	1	0.9	1	1	0	0	0.65

5 Discussion

Correlation network of stock exchange markets emerge as a powerful concept to study financial actors. In this study, the relationship of 21 different world stock exchange markets to each other is determined by Pearson's correlations. Then, we analyze hierarchies in the communities of each network through pre-crisis to post-crisis periods.

The economic crisis that took place in 2008 has caused radical changes in country behavior. The relationship between the country's stock exchanges is changing with the crisis period. For example; The ATX Austrian stock exchange is in direct contact with three stock exchanges, namely GDAXI, HSI and KLSE stock exchanges, and indirectly AEX, NZ50 stock exchanges. During the crisis period, ATX is in direct contact with the GDAXI and is indirectly involved with the GSPTSE and the KLSE stock exchanges, in total, with 3 exchanges. In the post-crisis period, it is only in involved with SMSI stock exchange. Therefore, Austria is among the countries most affected by the crisis.

The topological changes during the crisis are also studied in this paper. Beside the edge based measures such as L_{MCM}, L_{NTL}, L_{CPL}, the new measure we present in this paper has a significant change in post-crisis era to indicate the structural changes in the communities. Hence, we may conclude that the solitude measure is an efficient tool to characterize topological changes.

6 Acknowledgement

This paper has been granted by the Mugla Sitki Kocman University Research Projects Coordination Office. Project Grant Number: 17/093 and title Boundary Curvatures of Graph Communities.

7 References

[1] Agarwal, G. and Kempe, D., “Modularity-maximizing graph communities via mathematical programming”. The European Physical Journal B, 66, 3, 409-418, 2008.
[2] Balci, M. A. “Fractional virus epidemic model on financial networks”, Open Mathematics, 14, 1, 1074-1086 2016.
[3] Balci, M. A. “Hierarchies in Communities of Borsa Istanbul Stock Exchange”, Hacettepe Journal of Mathematics and Statistics, In Press, DOI:10.15672/hjms.201614520777.
[4] Bankevich, A. V. “Bounds of the number of leaves of spanning trees in graphs without triangles”. Journal of Mathematical Sciences, 184, 5, 557-563, 2012.
[5] Bollobas, B. Graph theory: an introductory course. Vol. 63. Springer Science & Business Media, 2012.
[6] Centola, D. and van de Rijt, A. “Choosing your network: Social preferences in an online health community”. Social science & medicine, 125, 19-31, 2015.
[7] Chan, T. F., S. Osher, and Shen,J. "The Digital TV Filter and Nonlinear Denoising." IEEE Transactions on Image Processing, 10, 2, 231-241, 2001.
[8] Chug, F. R. K. Spectral Graph Theory, American Mathematical Society (1997).
[9] Evans, T.S., “Clique graphs and overlapping communities”, Journal of Statistical Mechanics: Theory and Experiment, 12, P12037, 2009.
[10] Everett, M.G. and Borgatti, S.P., “Analyzing clique overlap”, Connections, 21, 1, 49-61., 1998.
[11] Gross, J. L., and Yellen J., eds. Handbook of graph theory. CRC press, 2004.
[12] Guler, H., Dundar, P., and Balci, M. A.“Solitude Number at Graphs”, IJ Pure and Applied Mathematics, 66, 3, 355-364, 2011.
[13] Hatpoğlu, V . (2017). Applcaton of a New Quanttatve Approach to Stock Markets: Mnmum Spannng Tree. Alphanumerc Journal, 5 (1), 163-169. DOI: 10.17093/alphanumerc.323988
[14] Hur, S-W. and Lillis, J., "Relaxation and Clustering in a Local Search Framework: Application to Linear Placement.", Proceedings - Design Automation Conference, 1999, 360-366.
[15] Jang, W., Lee, J., and Chang, W. “Currency crises and the evolution of foreign exchange market: evidence from minimum spanning tree”, Physica A, 390, 707–718, 2011.
[16] Lancichinetti, A., Santo F., and Kertész, J. "Detecting the overlapping and hierarchical community structure in complex networks." New Journal of Physics 11, no. 3, 033015, 2009.
[17] Morel, B. and Ramanujam, R., "Through the Looking Glass of Complexity: The Dynamicsof Organizations as Adaptive and Evolving Systems." Organization Science, 10 , 3, 278-293, 1999.
[18] Newman, M.E.,”Detecting community structure in networks.” The European PhysicalJournal B-Condensed Matter and Complex Systems, 38,2, 321-330, 2004.
[19] Newman, M.E.,”Fast algorithm for detecting community structure in networks.” Physicalreview E, 69, 6 066133, 2004.
[20] Newman, M.E. and Girvan, M., “Finding and evaluating community structure in networks”. Physical review E, 69, 2, p.026113, 2004.
[21] Newman, M. E. “The structure and function of complex networks”, SIAM review, 45, 2, 167-256, 2003.
[22] Onnela, J.-P., Chakraborti, A., and Kaski, K. “Dynamic asset trees and portfolio analysis”, The European Physical Journal B, 30, 285–288, 2002.
[23] Onnela, J.-P., Chakraborti, A. and Kaski, K. “Dynamics of market correlations: taxonomy and portfolio analysis”, Physical Review E, 68, 056110, 2003.
[24] Ozkanlar, A. and Clark, A. E., "ChemNetworks: A Complex Network Analysis Tool for Chemical Systems." Journal of Computational Chemistry 35, 6, 495-505, 2014.
[25] West, D. B.,Introduction to graph theory. Vol. 2. Upper Saddle River: Prentice hall, 2001.