Journal of International Money, Banking and Finance
Vol. 3, No. 1, 2022, pp. 1-12

URL : www.arfjournals.com
https://DOI: 10.47509 /JIMBF.2022.v03i01.01

TOPOLOGICAL AND DYNAMICAL PROPERTIES OF THE NETWORK OF SHAREHOLDERS IN S&P 500 COMPANIES BASED ON GRAPH DATABASES

Aristotelis Kompothrekas[1], Basilis Boutsinas[1] and Konstantinos Kollias[2]

[1]MISBILAB, Dept. of Business Administration, University of Patras, Greece
[2]Dept. of Economics, Democritus University of Thrace, Greece
Corresponding Author E-mail: email:vutsinas@upatras.gr

Article History

Received : 08 January 2022
Revised : 11 February 2022
Accepted : 21 February 2022
Published : 21 June 2022

To cite this article:

Kompothrekas, A., Boutsinas, B., & Kollias, K. (2022). Topological and Dynamical Properties of the Network of Shareholders in S&P 500 Companies based on Graph Databases. *Journal of International Money, Banking and Finance*, Vol. 3, No. 1, 2022, pp. 1-12. https://DOI: 10.47509 /JIMBF.2022. v03i01.01

ABSTRACT

The interesting properties of scale-free and small-world networks recently observed have triggered the attention of the research community to the study of real growing complex networks. In scale-free networks, most vertices are sparsely connected, while a few vertices are intensively connected to many others, indicating a "preferential linking" during growing. In small-world networks, the average length of the shortest path between two randomly chosen nodes is small. In this paper, we study the topological and dynamical properties of the network of shareholders (NOS) in 11593 different companies. Based on Graph Databases, we calculate all the well-known in the literature topological and dynamical properties of a network along with centrality measures of nodes of NOS, which quantify the role that a node plays in the overall structure of NOS. We prove that NOS is both a scale-free and small-world network. An understanding of NOS helps in predicting the emergence of important new phenomena affecting portfolio management in general. Also, this work reveals the fact that graph databases could serve as an efficient tool for analyzing such network models for stock markets. To the best of the authors' knowledge, this is the first study calculating all the well-known in the literature topological and dynamical properties for Market Investments Networks, that is based on graph databases.

Keywords: scale-free networks, small-world networks, graph databases, network of shareholders, network models for stock markets

JEL classification codes: C63, D83, G11

1. INTRODUCTION

Recent empirical studies (see Dorogovtsev and Mendes, 2003 and Sen and Chakrabarti, 2014 for a survey) of real growing complex networks, like WWW (Albert, Jeong and

Barabási, 1999), showed that they are very compact networks, i.e., the average length of the shortest path between two randomly chosen nodes is small. It is 19 for WWW, which is of the order of the logarithm of its size. Due to the smallness of the average length of the shortest path, these networks are referred to as "small-world" networks.

Also, empirical studies for various complex networks showed that the distribution of the number of connections of a node has a fat-tailed form, i.e., a power-law distribution. Thus, the distribution has no natural scale, as for instance the Poisson or exponential distribution, due to the latter these networks are also referred to as "scale-free" networks.

The interesting topological properties of such scale-free networks have triggered the attention of the research community to the study of real growing complex networks. In scale-free networks most nodes are sparsely connected, while a few nodes are intensively connected to many others, indicating a preferential linking (Barabási and Albert, 1999) during growing. The latter means that while such a network grows, its new edges become preferentially attached to nodes with a high number of connections.

It is shown that apart from WWW many complex natural or artificial networks are scale-free networks, e.g. network of citations in the scientific literature (Redner, 1998), network of metabolic reactions (Jeong *et al.*, 2000), etc. Such networks have a unique stability against failures and random damage, which is very important especially for biological and communication networks. To destroy such networks, i.e., to decay to a set of small unconnected components, one has to remove almost all their nodes or edges (Albert, Jeong and Barabási, 2000). Recently, specialized analysis techniques are applied to social networks and communities for monitoring statistics and identifying influencers (graph analytics).

In order to investigate the "scale-free" and "small-world" properties of the NOS, we analyze the *connectivity*, *clustering coefficient*, *betweenness* and *chemical distance* properties of network topology along with the dynamic properties: *connectivity distribution* and *load distribution*.

The presence of a power-law connectivity distribution makes the NOS an example of scale-free networks (Barabási, Albert and Jeong, 1999), while an average length of the shortest path between pairs of nodes, which is sharply peaked around its average value, makes the NOS an example of small-world networks (Watts and Strogatz, 1998).

There is some related work in the literature analyzing some the topological properties of NOS, (mainly centrality measures, i.e. out-connectivity, betweenness and closeness), in order to investigate: interactions between companies (Yao, Evans and Christensen, 2019), the control and wealth across markets (Vitali, Glattfelder and Battiston, 2011; Glattfelder and Battiston, 2009), inter-firm interactions (Ohnishi, Takayasu and Takayasu, 2010), the shareholder behavior (H.Li *et al.*, 2016), the stock volatility (J.Li *et al.*, 2016), the trade size distribution (Jiang and Zhou, 2010), the shock transmission effects (D'Errico *et al.*, 2009), the stock market concentration (Rotundo and D'Arcangelis, 2014), and the portfolio optimization (Garlaschelli *et al.*, 2005).

In this paper, we calculate all the well-known in the literature topological and dynamical properties of a network, that are needed to prove the "scale-free" and "small-world" properties of NOS. We also calculate the centrality measures of nodes in NOS, which quantify the role that a node plays in the overall structure of a network. Thus, the proposed calculating methodology can improve our understanding of the interactive relationship between companies and shareholders and of developing of trading strategies.

The time complexity of calculating some of these properties is high. Thus, calculations are based on Graph Databases (Needham and Hodler, 2019). Graph Databases besides having an object-oriented thinking, they have a superior performance for querying related data either big or small. A native graph has the so-called index-free adjacency property, where each vertex maintains its neighbor vertices information only, and no global index about vertex connections exists. Graph databases can support the representation of rich and varied relationships between objects and the detection of patterns based upon these relationships (Bechberger and Perryman, 2019). Thus, Graph Databases constitute an excellent platform for queries on real time big data, since such a platform (graph compute engine) is a technology that enables computational algorithms to be run against large datasets (Robinson, Webber and Eifrem, 2015). To the best of the authors' knowledge, this is the first study calculating all the well-known in the literature topological and dynamical properties for Market Investments Networks, that is based on graph databases.

In the rest of the paper, first we define the NOS (Section 2) and then we calculate its properties (Section 3). In Section 4 we draw our perspectives and finally we conclude (Section 5).

2. BUILDING THE NOS

NOS is constructed by using data from S&P 500. The S&P 500, or just the S&P, is a stock market index that measures the stock performance of 500 large companies listed on stock exchanges in the United States. It is one of the most followed equity indices, and many consider it to be one of the best representations of the U.S. stock market. The S&P 500 is widely regarded as the best single gauge of large-cap U.S. equities. There is over USD 9.9 trillion indexed or bench-marked to the index, with indexed assets comprising approximately USD 3.4 trillion of this total. The index includes 500 leading companies and covers approximately 80% of available market capitalization.

In particular, NOS is constructed by using a snapshot of S&P 500 corresponding to March 2017. Data were collected by using the Bloomberg Professional service in a Bloomberg Terminal. Companies were grouped to seven categories: Energy, Financial, Information Technology, Health Care, Materials, Telecommunications Services, and

Utilities. Within the initial dataset, a different label could be assigned to a company as an investor and as a stock. Thus, we automatically detect and match the different labels.

After the preprocessing, NOS consists of 11593 nodes (N is the set of nodes) and 295989 edges (E is the set of edges). A node represents a company, and a directed edge is placed from a company-shareholder to each company-stock in which it has invested.

All experiments were run using the Python Programming Language and the Cypher Graph Query Language on an i7 processor with 8GB of memory. To characterize both the topological and the dynamical properties of the NOS, we used the nerworkx (Hagberg *et al.,* 2008) and powerlaw (Alstott *et al.,* 2014) packages. The Graph is formed by the Neo4j (Needham and Hodler, 2019) graph visualization capabilities.

3. TOPOLOGICAL PROPERTIES OF NOS

To characterize the topological properties of the NOS, we analyze the connectivity, the clustering coefficient, the chemical distance and the betweenness measures.

The connectivity k_i of a node i is defined as the number of connections of this node with other nodes in the network ($k_i = |\{(i, j) , (i, j) + E\}|$), while the average connectivity k_{avg} is defined as the average of k_i over all nodes in the network ($k_{avg} = \Sigma_{i+N} k_i/|N|$). The in-connectivity ($kin_i = |\{(j, i) , (j, i) \in E\}|$) and out-connectivity ($kout_i = |\{(i, j) , (i, j) \in E\}|$) for directed networks are defined accordingly. Note that kin_i corresponds to portfolio diversification while the sum of incoming link weights corresponds to portfolio volume.

Since each edge contributes to the connectivity of two nodes, in the general we could have that $k_{avg}= 2|E|/|N| = 51.063$, where $|E|$ is the total number of edges and $|N|$ is the total number of nodes. For the NOS k_{avg} = 50.95, while kin_{avg} *H* + $kout_{avg}$ =25.532. Thus, in average each node has fifty connections, which is a small number compared with that of a fully connected network of the same size ($k_{avg} = |N| + 1$ H+ 10^4). Internet also exhibits the same behavior (Vázquez, Pastor-Satorras, and Vespignani, 2002). Therefore, the NOS are far from being a random graph.

The generalization of the connectivity of a node, i.e., the number of its nearest neighbors, is the number of its second nearest neighbors ($k2_i = |\{(s,t), (s,t) \in E \wedge (sa+i \vee (i,s)+E)\}|$), third ($k3_i$) and so on. The distance between a node and any of its m-nearest neighbors is m. In NOS $k2_{avg}$=28333.177 and $k3_{avg}$=1767.847. Also, the maximum connectivity is found to be 6273 in NOS.

The average connectivity gives information about the number of connections of any node but not about the overall structure of these connections, as the clustering coefficient. The clustering coefficient c_i (Watts and Strogatz 1998) of a node i ($c_i = e_i/(k_i(k_i + 1)/2)$) is defined as the ratio between the number of connections (e_i) among its k_i neighbors and the maximum possible connections among its k_i neighbors, i.e., $k_i(k_i + 1)/2$. The average clustering coefficient c_{avg} is defined as the average of c_i over all nodes in the network ($c_{avg} = \Sigma_{i+N} c_i/|N|$). The clustering coefficient represents the "cliquishness" of the neighborhood of a node, which is the extent of the mutual "acquaintance" of its neighbors.

Scale-free networks have much greater average clustering coefficient than that of the classical random graphs with the same total numbers of nodes and connections (Dorogovtsev and Mendes, 2003). For instance, the average clustering coefficient of Internet is 0.2, i.e., a relative difference about 300 from the corresponding classical random graph. In comparative results presented in (Dorogovtsev and Mendes, 2003) for 37 different fittings of data of different real networks, the values of average clustering coefficient are inside the band $0.03<c_{avg}<0.76$ much greater that the corresponding classical random graph.

The maximum value of c_{avg} is 1, corresponding to a fully connected network. For random graphs (Watts and Strogatz, 1998; Bollobás, 1985), which are constructed by connecting nodes at random with a fixed probability p, the clustering coefficient decreases with the network size N as $c_{avg\_random} = k_{avg} /|N|$. On the contrary, it remains constant for regular lattices.

For the NOS $c_{avg} = 0.143$, three orders of magnitudes larger than $c_{avg\_random} = k_{avg} /|N| \approx 10^{-4}$, corresponding to a random graph with the same number of nodes. Therefore, the NOS are far from being a random graph.

The average chemical distance d_{avg} (Vázquez, Pastor-Satorras, and Vespignani, 2002) is defined as the shortest path distance d_{ij} between two nodes i and j, averaged over every pair of nodes in the network. The diameter of a network is defined in the literature as either the average chemical distance or the length of the longest shortest path.

In comparative results presented in Dorogovtsev and Mendes, 2003, the values of average chemical distance (average shortest path distance) is inside the band $2.32 \leq d_{avg} \leq 16$ close to those for the corresponding classical random graph. Thus, the small-world property does not depend crucially on the scale-free property of a network.

For the NOS $d_{avg} = 2.943$ and its diameter (longest shortest path) is 4. If the NOS could be mapped into a two-dimensional regular lattice, we should observe $d_{avg+lattice} \approx |N|^{1/2} \approx 107 << 2.943$. Thus, the NOS, as the Internet for instance (Vázquez, Pastor-Satorras, and Vespignani, 2002), strikingly exhibits what is known as the "small-world" effect: in average one can go from one node to any other in the system passing through a very small number of intermediate nodes. This necessarily implies that besides the short local connections which contribute to the large clustering coefficient, there are some hubs and backbones which connect different regional networks, strongly decreasing the average chemical distance (Vázquez, Pastor-Satorras, and Vespignani, 2002).

Closeness centrality (Borgatti, 2005) is defined as $c(i)=1/\Sigma_{j \in N} d_{ij}$, where d_{ij} is the shortest path distance between nodes i and j. The maximum is reached by nodes connected to each of the others, while the minimum is reached by the first and last node of a chain network. The larger the closeness of a node, the shorter the distances to other nodes. For directed networks $c_{in}(i) = 1/\Sigma_{j\in N} d_{ij}$ and $c_{out}(i) = 1/\Sigma_{j\in N} d_{ij}$ are defined accordingly. We calculated the c(i), $c_{in}(i)$ and $c_{out}(i)$ for every $i \in N$. The results are shown in Fig. 1, presenting the number of nodes per closeness centrality interval. Intervals are produced by dividing the value to classes of equal length. The optimal histogram bin width has been calculated using Scott's rule (Scott, 1979). It is interesting that almost half of the companies are evaluated with a near to zero closeness centrality, while a great portion of the other half is positioned in the middle of the distribution.

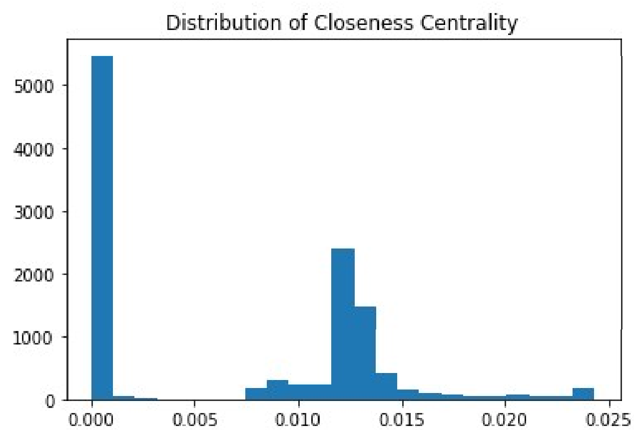

Figure 1: Closeness Centrality Distribution

The betweenness b_i (Newman, 2001) of a node i is defined as the total number of shortest paths between any two nodes in the network that pass through the node i, while the average betweenness b_{avg} is defined as the average of bi over all nodes in the network. The betweenness indicates whether or not a node is important in "traffic" on a network. Nodes with a larger betweenness are more "influenced".

For a star network, the betweenness takes its maximum value $|N|(|N| - 1)/2$ at the central node and its minimum value $|N| - 1$ at the vertices of the star. Mapping the Internet as an autonomous systems (AS) topology, i.e., a collection of subnetworks that are connected together, it is shown in Vázquez, Pastor-Satorras, and Vespignani, 2002 that betweenness for Internet is between $2|N|$ and $3|N|$, which is quite small in comparison with its maximum possible value $|N|(|N| - 1) \approx 10^7$.

For the NOS, $b_{avg} = 45{,}709 \approx 4|N|$ which is also quite small in comparison with its maximum possible value $|N|(|N| + 1) \approx 10^7$.

The flow betweenness generalizes the notion of betweenness for valued networks. All paths between nodes, not only *shortest paths*, are considered. The flow betweenness fb_i, e.g. in D'Errico *et al.*, (2009), of a node i is defined as the total number of maximum flow paths between any two nodes in the network that pass through the node i. Formally: $fb_i = \Sigma_{m<n} mf_{mn}(i)$, $m,n \neq i$, where $mf_{mn}(i)$ is the maximum flow from node *m* to node *n* passing through *i*. Thus, the flow betweenness measures the contribution of a node to all possible maximum flows. We calculated the fb_i, for every i+N, setting all flows to 1. The results are shown in Fig. 2, presenting the number of nodes per flow betweenness interval. Intervals are produced by dividing the value to classes of equal length. The great majority of the companies are equally evaluated and belong to the same distribution class having a near to zero flow betweenness.

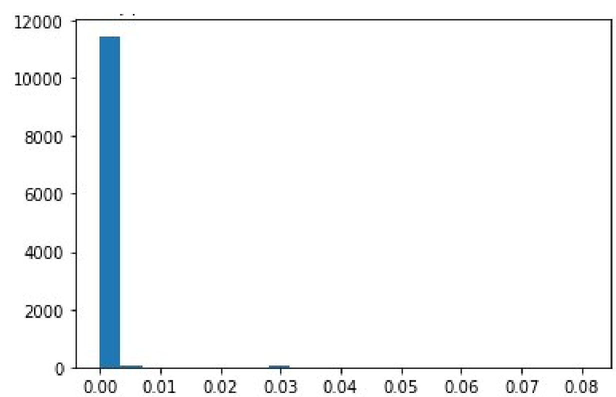

Figure 2: Distribution of Flow Betweenness

The eigenvector centrality (Borgatti, 2005) is defined as the principal eigenvector of the adjacency matrix defining the network. The eigenvector centrality a node is proportional to the sum of the eigenvector centralities of its neighbors. We calculated the ev_i, for all nodes. The results are shown in Fig. 3, presenting the number of nodes per eigenvector centrality interval. Intervals are of equal length here too. The eigenvector centrality distribution shows a power law behavior where the first three classes dominate (80%) the rest to the right of the long tail. By excluding the last two classes, the distribution follows a power law with an exponent γ =1.75

Networks consist of separate connected components. The statistics of the connected components are related to percolation problems, as the size of the percolating cluster (the giant connected component-GCC) and the distribution of the sizes of these connected components (Dorogovtsev and Mendes, 2003). The relative size of the GCC indicates the stability of a network. A network is not a unit organism if the GCC is absent.

Scale-free networks contain a complex set of interpenetrating giant components. Excluding the GCC, the rest of such network includes finite-size disconnected components (see Dorogovtsev and Mendes, 2003). The analysis presented in Broder *et al.,* 2000 reveals

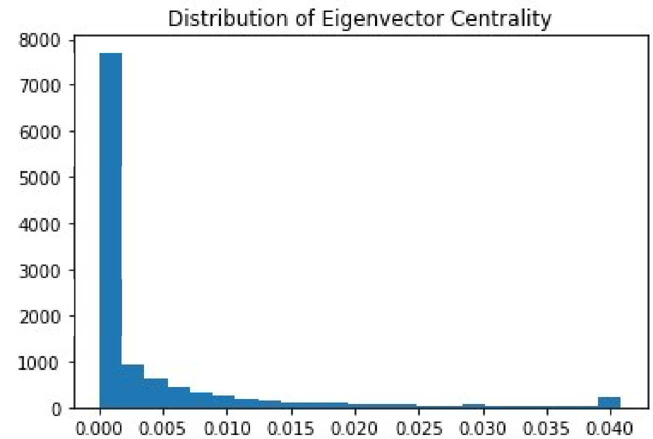

Figure 3: Eigenvector Centrality Distribution

that most (over 90%) of the approximately 203 million nodes in the subset of WWW used during tests, form a single connected component if links are treated as undirected edges. This connected web breaks naturally into four pieces. The first piece is a central core, all of whose pages can reach one another along directed links — this "giant strongly connected component" (SCC) is at the heart of the WWW. The second and third pieces are called IN and OUT. IN consists of pages that can reach the SCC, but cannot be reached from it - possibly new sites that people have not yet discovered and linked to. OUT consists of pages that are accessible from the SCC, but do not link back to it, such as corporate websites that contain only internal links. Finally, the TENDRILS contain pages that cannot reach the SCC, and cannot be reached from the SCC. Perhaps the most surprising fact is that the size of the SCC is relatively small, i.e., it comprises about 56 million pages. Each of the other three sets contain about 44 million pages. Thus, all four sets have roughly the same size.

In NOS all nodes form a connected component (GCC) in the case of undirected connections. For the directed graph, there are 2 strongly connected components (SCC) with 22 members, and 1 strongly connected component with 50 members. Only a small part of NOS is consisted of a strongly connected component. They are set of members that own directly and/or indirectly shares in every other member (Vitali, Glattfelder and Battiston, 2011). The IN and OUT strongly connected components are both consisted of the same set of 50 companies.

4. DYNAMICAL PROPERTIES OF NOS

To characterize the dynamical properties of the NOS, we analyze the connectivity and the betweenness probability distributions. We show that the NOS structure is reflected in non-trivial scale-free connectivity and betweenness correlation functions.

The NOS has a fat-tailed distribution of connections. Thus, connectivity displays a power-law distribution with an exponent $\gamma \approx 2.019$. The power law in the connectivity distribution $P(k) = k^{-\gamma}$ characterizes the scale-free networks, however the exponent γ is not universal and depends on the detail of network structure. The exponent γ for various networks is in the range $2 < \gamma \leq 3$ (Goh *et al.*, 2002; Vázquez, Pastor-Satorras, and Vespignani, 2002; Redner, 1998).

In NOS, both the in-connectivity and out-connectivity also follow a power-law distribution with an exponent $\gamma \approx 2.01$ (setting parameter xmin to 28 after several tests) and $\gamma \approx 1.99$ (setting parameter xmin to 3 after several tests), respectively. Similar results have been obtained also for the WWW (Broder *et al.*, 2000; Barabási and Albert, 1999), where $\gamma \approx 2.1$ and $\gamma \approx 2.72$ respectively. Thus, we observe that these connectivity, in-connectivity and out-connectivity distributions exhibit an algebraic behavior and are characterized by scaling exponents which are stationary in time.

Also, in NOS the betweenness distribution follows a power law with an exponent g » 2.815 (setting parameter xmin to 1 after several tests). Since the topological feature of a network can be grasped more transparently using betweenness distribution, the latter is extensively investigated in the literature. In Goh *et al.*, 2002, scale-free networks can be classified into only two classes, based on numerical measurements of the exponent g of betweenness distribution. For the first class, the exponent is $\gamma \approx 2.2$ and for the second class the exponent is $\gamma \approx 2.0$. It is shown that such different universal behaviors in the betweenness distribution originate from different generic topological features of networks, which in turn impose different behaviors. For instance, the resilience of networks under an attack is different for each class: the networks of the second class are much more vulnerable to the attack than those for the first class.

The clustering coefficient distribution is also an interest quantity. A definitive evidence for a power-law behavior of clustering coefficient distribution is not described in the literature. In Vázquez, Pastor-Satorras, and Vespignani, 2002, the distribution of the clustering coefficient as a function of the node connectivity (c_{k_avg}) is investigated for the Internet. In the case of c_{k_avg} the local clustering coefficient of each node c_i is averaged over all nodes with the same connectivity k. Also, in this case, measurements yield a power-law behavior with exponent g = 0.75 ± 0.03, extending over three orders of magnitudes. This implies that nodes with a small number of connections have larger local clustering coefficients than those with a large connectivity. In NOS the c_{k_avg} distribution also follows a power law with an exponent $\gamma \approx 2.5$ (xmin 4). Finally, note that the distribution of the sizes of SCCs also obeys a power law with an exponent $\gamma \approx 2.5$ (Broder *et al.*, 2000). This is true also for the NOS, where the latter distribution follows a power law with an exponent $\gamma \approx 1.21$.

5. DISCUSSION

Topological and dynamical properties of such real growing complex networks have important direct consequences. An impressive property is their stability against failures and random damage. A network can be destroyed by transforming it to a set of small unconnected clusters. Thus, to destroy a scale-free network one has to remove at random almost all its vertices or edges (Albert *et al.,* 2000; Cohen *et al.,* 2000). Such resilience to failures is obviously necessary for biological, e.g. for epidemic spreading (Pastor-Satorras and Vespignani, 2001), and communication networks, e.g. for network breakdown (Cohen *et al.,* 2000). This property of NOS could help in investigating the propagation of economic shocks. For example, NOS is robust against companies and shareholders variations and the market volatility could be predicted (j.Li *et al.,* 2016).

At the same time, another impressive property of scale-free networks is the absence of an epidemic threshold (Pastor-Satorras and Vespignani, 2001). Thus, “diseases” may easily spread within them. NOS is a compact network; thus, a large amount of its members is expected to be “infected” by a significant change of the economic environment. For example, this property of NOS could help in investigating how control tends to be dispersed among shareholders (Glattfelder and Battiston, 2009).

Obviously, an understanding of NOS helps in predicting the emergence of important new phenomena affecting portfolio management in general.

Thus, preferential linking is justified by the fact that investors are seeking to invest in high liquidity companies (liquid stocks - a measure of liquidity, though not the best, is the number of institutional investors which attracts) and high free-float (free float - percentage of the company available to the investing public). As a result, one can explain the fact that “the world is in control of a few important shareholders” (Vitali, Glattfelder and Battiston, 2011).

The “small-world” effect is justified by the fact that most of the sites belong to investors, because the same investment companies mainly invest in “strong” stocks such as those included in the S&P 500.

Clustering Coefficient is justified by focusing mainly on investment companies. It is true that each one of them builds around most S&P 500 companies.

Betweenness is justified by the fact that we have few nodes in the network (S&P 500 companies) where we would expect many shortest paths to pass (investors).

Also, cross-shareholdings are represented as strongly connected components (SCCs) in graph theory. They are sub-network structures where companies own each other directly or indirectly through a chain of links. It means that they form in the graph cycles and are all reachable by every other firm in the SCC. Institutions, such as the antitrust regulator, which must guarantee competition in the markets, pay attention in this kind of ownership relations. The companies, as well, set up cross-shareholdings for coping with possible takeovers, monitoring and strategies reducing market competition.

6. CONCLUSION

We proposed a network description (NOS) of a stock market (assets traded in and corresponding shareholders). We proved that NOS is both a scale-free and small-world network, since NOS is characterized by power-law distributions and a small average length of the shortest path between pairs of nodes.

We calculated all the well-known in the literature topological and dynamical properties of NOS, as well as the centrality measures of nodes in NOS. These properties are the subject of fundamental financial issues, such as portfolio optimization, shareholder behavior, stock volatility, shock transmission effects, corporate control, etc.

Since the time complexity of calculating some of these properties is high, the calculations are based on Graph Databases. This work reveals the fact that graph databases could serve as an efficient tool for analyzing such network models for stock markets. Such analytics is path analysis that focuses on the relationships between the nodes, connectivity analysis that focuses on the weight of the edges between nodes, community analysis that focuses on the interactions between nodes and centrality analysis that focuses on the relevancy of each node.

Acknowledgments

The authors wish to thank both Jacqueline Delliou for data transformation/preparation and Christomanolis Konstantopoulos for his invaluable suggestions and discussions.

References

Van der Geer, J., Hanraads, J. A. J., and Lupton R. A. (2000). The art of writing a scientific article. *Journal of Scientific Communications* 163, 51-59

Albert, R., Jeong, H., and Barabási, A. L. (1999). Diameter of world-wide web, *Nature*, 401(6749), 130-131.

Albert, R., Jeong, H., and Barabási, A. L. (2000). “Error and attack tolerance of complex networks”, *Nature*, 406, 378-382.

Alstott, J., Bullmore, E., and Plenz, D. (2014). “Powerlaw: A Python Package for Analysis of Heavy-Tailed Distributions”, Rapallo, F. (Ed.), *PLoS ONE, 9.1: e85777. Crossref. Web*.

Barabási, A. L., and Albert, R. (1999). “Emergence of scaling in random networks”, *Science*, 286(5439), 509-512.

Barabási, A. L., Albert, R., and Jeong, H. (1999). “Mean-field theory for scale-free random networks”, *Physica A*, 272(1), 173-187.

Bechberger, D., and Perryman, J. (2019). Graph Databases in Action, Manning Publications, New York.

Broder, A., Kumar, R., Maghoul, F., Raghavan, P., Rajagopalan, S., Stata, R., Tomkins, A., and Wiener, J. (2000). “Graph structure in the web”, *Computer Networks: The International Journal of Computer and Telecommunications Networking*, 33(1-6), 309-320.

Bollobás, B. (1985). Random Graphs, Academic Press, London.

Borgatti S. (2005). “Centrality and network flow”, *Social Network*, 27(1), 55-71.

Cohen, R., Erez, K., ben-Avraham D., and Havlin, S. (2000). "Resilience of the Internet to random breakdowns", *Phys. Rev. Lett.*, 85(21), 4626-4628.

D'Errico, M., Grassi, R., Stefani, S., and Torriero, A. (2009). "Shareholding Networks and Centrality: An Application to the Italian Financial Market", Naimzada, A.K., Stefani, S., Torriero, A. (Ed.s), Networks, Topology and Dynamics, Lecture Notes in Economics and Mathematical Systems 613, Springer, Berlin, Heidelberg.

Dorogovtsev, S.N., and Mendes, J.F.F. (2003), Evolution of Networks: From Biological Nets to the Internet and WWW, Oxford University Press, Oxford.

Garlaschelli, D., Battiston, S., Castri, M.V., Servedio, D.P., and Caldarelli, G. (2005). "The scale-free topology of market investments", *Physica A*, 350, 491-499.

Glattfelder, J.B., and Battiston, S. (2009). "Backbone of complex networks of corporations: The flow of control", *Phys. Rev. E*, 80, 036104.

Goh, K.-I., Oh, E.S., Jeong, H., Kahng, B., and Kim, D. (2002). "Classification of scale free networks", *Proceedings of the National Academy of Science*, 99(20), 12583-12588.

Hagberg, A., Swart, P., and Chult D. (2008). Exploring network structure, dynamics, and function using NetworkX", in Varoquaux, G., Vaught, T., Millman, J. (Ed.s.), Proceedings of the 7th Python in Science Conference (SciPy2008), Pasadena CA USA, 11–15.

Jeong, H., Tombor, B., Albert, R., Oltvai, Z.N., and Barabási, A.-L. (2000). "The large-scale organization of metabolic networks", *Nature*, 407, 651-654.

Jiang, Z-Q, Zhou, W-X. (2010). "Complex stock trading network among investors", *Physica A*, 389(21), 4929-4941.

Li, H., An, H., Huang, J., Huang, X., Mou, S., and Shi, Y. (2016). "The evolutionary stability of shareholders' co-holding behavior for China's listed energy companies based on associated maximal connected sub-graphs of derivative holding-based networks", *Applied Energy*, 162, 1601–1607.

Li, J., Ren, D., Feng, X., and Zhang, Y. (2016). "Network of listed companies based on common shareholders and the prediction of market volatility", *Physica A*, 462, 508-521.

Needham, M., and Hodler, A.E. (2019). Graph Algorithms: Practical Examples in Apache Spark and Neo4j Specs, O'Reilly Media, Sebastopol California.

Newman, M.E.J. (2001). "Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality", *Phys. Rev. E*, 64, 016132.

Ohnishi, T., Takayasu, H., and Takayasu, M. (2010). "Network motifs in an inter-firm network", *Journal of Economic Interaction and Coordination*, 5, 171–180.

Pastor-Satorras R., and Vespignani, A. (2001). "Epidemic spreading in scale-free networks", *Phys. Rev. Lett.*, 86(14), 3200-3203.

Redner, S. (1998). "How popular is your paper? An empirical study of the citation distribution", *European Physical Journal B*, 4(2), 131–134.

Robinson, I., Webber, J., and Eifrem, E. (2015). Graph Databases: New Opportunities for Connected Data, O'Reilly Media, Sebastopol California.

Rotundo, G., and D'Arcangelis, A.M. (2014). "Network of companies: an analysis of market concentration in the Italian stock market", *Quality and Quantity*, 48, 1893–1910.

Sen, P., and Chakrabarti, B.K. (2014). Sociophysics: An Introduction, Oxford University Press, Oxford.

Scott, D.W. (1979). "On optimal and data-based histograms", *Biometricka*, **66**(3), 605-610.

Vázquez A., Pastor-Satorras, R., and Vespignani, A. (2002). "Large-scale topological and dynamical properties of Internet", *Physical Review E*, **65**(6), id. 066130.

Vitali, S., Glattfelder, J., and Battiston, S. (2011). "The Network of Global Corporate Control", *PLoS ONE*, **6**(10), e25995.

Watts, D.J., and Strogatz, S.H. (1998). "Collective dynamics of 'small-world' networks", *Nature*, **393**, 440–442.

Yao, Q., Evans, TS., and Christensen, K. (2019). "How the network properties of shareholders vary with investor type and country", *PLoS ONE*, **14**(8), e0220965

