Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 5 (2011) 442–449

Procedia Computer Science

The 2nd International Conference on Ambient Systems, Networks and Technologies (ANT)

Identifying influential stock indices from global stock markets: A social network analysis approach

Ram Babu Roy, Uttam Kumar Sarkar*

Indian Institute of Management Calcutta, Kolkata, 700 104, India

Abstract

We have proposed a method to rank the stock indices from across the globe using social network analysis approach. The temporal evolution of correlation network and Minimum Spanning Tree (MST) of global stock indices have been analyzed using weekly returns of 93 stock indices for five-year period from the year 2006 through 2010 obtained from Bloomberg. We have chosen this period to study the behaviour of the stock market network before and after the collapse of Lehman Brothers in the USA. Our study attempts to answer the questions about identifying the most influential stock indices in the global stock market, regional influence on the comovement of stock indices, and the impact of the collapse of Lehman Brothers in the USA and the associated global financial crisis that followed on the dynamics of stock market network.

Selection and/or peer-review under responsibility of Prof. Elhadi Shakshuki and Prof. Muhammad Younas.

Keywords: Stock Market Indices; Comovement; Social Network; Centrality; Minimum Spanning Tree; Financial Crisis

1. Introduction

The necessity to understand the complex linkages among the global stock market has increased following the spread of recent financial crisis in the US to the global stock markets and financial institutions. Presence of a large number of heterogeneous interacting elements and non-linearity in their behaviour lead to complex emergent behaviour of stock markets. The behaviour of individual stock indices is normally monitored using a time series data. Temporal analysis of such data becomes very difficult as the number of indices to monitor increases. Recent literature has focused on the network based approach as a viable option to study the complex interrelationship between the stock markets by

* Corresponding author. Tel.: +91 943 301 8562; fax: +91 332 467 8307.
E-mail address: uttam@iimcal.ac.in.

Selection and/or peer-review under responsibility of Prof. Elhadi Shakshuki and Prof. Muhammad Younas. doi:10.1016/j.procs.2011.07.057

presenting the information hidden in the huge volume of time-series stock market data in a manageable fashion. Such approach may help in revealing the internal structure of various stock markets in the world. This may help in international diversification of portfolio and for improving the global financial systems.

Representation of stock market data using network and the properties of the network topology of US stock market have been studied extensively in [1]. It has been shown in [2] that the US stock market netwok obtained from price returns has scale-free degree distribution and this property has been exploited to propose a degree-based index as an indicator of market performance. The concept of networks has been applied to study the behaviour of Korean [3], Chinese [4], Indian [5], and Brazilian [6] stock markets. Most of the literatures on network based models have studied the behaviour of the stock markets using stocks of a particular country that is unable to capture the behaviour of global stock markets. Influential stock indices may influence investor's decision during extreme events and hence indices often tend to move together. The importance of stock indices is normally represented by the market capitalization of the stocks included in the index but the comovement of stock indices are not taken into account in judging their importance. Further, there is no established method to rank the global stock indices and identify the most influential one. Very few works on the evolution of interdependence of global stock market indices have been reported [7, 8]. These papers have studied the regional behaviour of stock markets with relatively small sample of indices. The interdependence in the world equity market with large sample of indices reveals increasing integration of global equity market over the period 1997-2006 [9]. To the best of our knowledge, no work has been reported on identifying influential stock indices using the social network approach and the impact of recent financial crisis during 2008 on the evolutionary behaviour of the global stock market network. This paper attempts to contribute towards bridging this research gap in the literature. Our objective in this paper is two-fold. First, to propose a model to rank the stock indices based on their influence in global stock market and second, to investigate whether there is any significant change in the ranking of these indices after the recent financial crisis.

The rest of the paper is organized as follows. Section 2 gives an overview of some key concepts related to network model used in this study. Section 3 describes the data sets used and section 4 describes the methodology to derive stock market network and identification of influential stock indices from such a network. Section 5 gives the empirical results and discussions. Section 6 draws concluding remarks.

2. Key Metrics

We outline and review a few definitions related to the stock market correlation network and the associated Minimum Spanning Tree (MST) used in our study.

2.1. Correlation network and Minimum Spanning Tree (MST)

A network representation of stock market can be represented by a graph G (V, E) where V is the set of nodes representing the stock indices and E is a set of edges representing interrelations between stock index pairs [1]. The interrelation between a pair of stock indices is captured by the level of similarity between the index returns. The two indices are assumed to behave similarly if the correlation between the index returns is greater or equal to some specified threshold (θ). A Minimum Spanning Tree (MST) is used to reduce the full network into a simpler network that helps in getting better visual insights about the network [10]. The distance d_{ij} between stock indices i and j used for constructing the MST is computed from the correlation ρ_{ij} between stock indices i and j using the transformation as follows, so that a higher value of correlation implies a lower value of distance between the two stock indices.

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

The length L of an MST is given by the sum of all the distances l_{ij} between the nodes i and j connected by an edge in the MST.

$$L = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} l_{ij} \tag{2}$$

Where $l_{ij} = d_{ij}$ if the nodes i and j are connected directly in the MST, and $l_{ij} = 0$ otherwise. Therefore, the inverse of the MST length (L^{-1}) is a proxy for the level of integration as implied by the correlation between the stock indices.

2.2. Node Centrality and Network Centralization

Centralization is an important concept in social network analysis and refers to the extent to which the network revolves around a single node or a single edge. Different centrality measures are defined in the literature for the nodes and edges in the network as well as for the network as a whole. These centrality measures capture different aspects of the interaction between the actors represented by the nodes of the network. The centrality of a node is a measure of the potential importance, influence, prominence of the node in a network that is derived from its relative position compared to other nodes in a network. There are four types of centrality measures of the node centrality namely degree centrality, betweenness centrality, closeness centrality, and eigenvector centrality well established in the literature. *Degree centrality* measures the importance of a node due to its direct linkage with other nodes and is measured by the number of ties to other nodes [11]. C*loseness centrality* is a measure of how close a given node is to all other nodes. It is defined as the inverse of average geodesic (shortest) distance of a given node to all other nodes. *Betweenness centrality* measures the relative importance of a node in linking two other nodes through a shortest path in the network. It is measured as the proportion of geodesic paths that pass through a node. *Eigenvector centrality* measures the influence of a node in the network based on who its neighbors are connected to as it takes into account the entire pattern of connections in the network. Eigenvector centrality is given by the eigenvector of the largest eigenvalue of an adjacency matrix [12].

3. Data Description

The daily closing prices for 85 stock indices from 36 countries from across the world for five years from January 2006 to December 2010 obtained from Bloomberg have been used. In addition to these stock indices from various countries, we have included 8 indices namely, SX5E, SX5P, SXXE, SXXP, E100, E300, SPEURO, SPEU from European region in our study to investigate whether the regional indices have any influence on the network structure. We have chosen this period to study the behaviour of the stock market network before and after the collapse of Lehman Brothers in the USA. It has been shown using the random matrix theory (RMT) that for a return correlation matrix computed for n stock indices with T time records in the limit of $T, n \to \infty$ and $Q = (T/n) \geq 1$, the noise in the estimation of correlation matrix using finite period data is greatly reduced. Since we have only 120 weekly return data for post Lehman Brother event, the number of indices is restricted to n ≤ 120. Therefore, we have restricted our samples to only those indices existing for longer period and have data available on Bloomberg (say from 1990) giving us 93 such indices. Finally, we have 260 weekly return data points for the entire period of five years.

4. Methodology

The logarithmic return $R_i(t)$ of a stock index *i* over the period from *(t − 1)* to *t* is given by

$$R_i(t) = \ln\left(\frac{P_i(t)}{P_i(t-1)}\right) \quad (3)$$

Where $P_i(t)$ is the closing price of the stock index *i* at time *t* and the correlation coefficient between stock indices *i* and *j* is given by the standard formula [1].

$$C_{ij} = \frac{\langle R_i R_j \rangle - \langle R_i \rangle \langle R_j \rangle}{\sqrt{\left\langle R_i^2 - \langle R_i \rangle^2 \right\rangle \left\langle R_j^2 - \langle R_j \rangle^2 \right\rangle}} \quad (4)$$

Where $\langle R_i \rangle$ represents the average return of the stock index *i* over a specified time period T consisting of say, N weeks i.e. $\langle R_i \rangle = \frac{1}{N}\sum_{t=1}^{N} R_i(t)$.

Therefore, C_{ij} is dependent on the specified time period T. An edge connecting stock indices *i* and *j* is added to the graph if the correlation C_{ij} is greater or equal to some specified threshold (θ). The correlation network and MST are constructed using the correlation coefficients. Two indices have been assumed to behave similarly if the correlation between them is greater or equal to 0.6. We have used weekly returns to reduce the impact of variations in trading hours in various stock markets across the world. MATLAB, Microsoft Excel, and Pajek software have been used for the analysis of data and presentation of the results. The data set is partitioned into 36 periods of length 120 weekly returns each with each consecutive period obtained by sliding the sampling window by 4 data points (corresponding to 4 weeks). This partitioning is done to obtain enough number of sample data sets to compute enough data points for the network metrics corresponding to consecutive periods in order to study the dynamics of the network structure. Period 37 corresponds to the entire data set of 260 weeks for five years and this represents the long run behaviour of the stock market network during the entire five-year period of study. The start and end dates for some selected periods in (MM/DD/YYYY) format are shown in Table 1. Using the correlation matrices, the correlation networks and MST are constructed for all the 37 periods. The evolutionary behaviour of the length of the MST has been investigated for the presence of any peculiarity around the transition from tranquil period to crisis period to get more insight into the evolution of the network structure during the crisis. The observations on their behaviour would be explained in section 5.

Table 1: Sample periods with associated start and end date

Period No	Start Date	End Date	Period No	Start Date	End Date
1	1/11/2006	4/23/2008	36	9/17/2008	12/29/2010
6	5/31/2006	9/10/2008	37	1/4/2006	12/29/2010
7	6/28/2006	10/8/2008			

4.1. Finding influential stock indices using centrality measure

We propose a method to identify influential stock indices using the centrality measures widely used in the social network literature. The ranks of all the indices are computed based on all the four centrality measures discussed earlier from the MST and the correlation network (excluding the closeness centrality). Therefore, we have a set of seven ranks for each of the indices. Indices are arranged in descending order of their centrality value based on each of the centrality measures and assigned ranks in ascending order with rank 1 being assigned to the index with the highest centrality. Nodes with same value of centrality are assigned the same rank and the rank of the next node with subsequent lower centrality was adjusted as per the number of nodes higher up in the ranking. For example, if we have a tie for the first position, then both the nodes are assigned a rank 1 and the next node in order is assigned a rank 3 and so on. These ranks are normalized by dividing with the respective maximum rank, yielding normalized ranks ranging from 0 to 1. The final rank of each of the indices is computed based on the average of these seven normalized ranks for all the indices. Hence, the most influential index will have the lowest value of the average rank whereas the least influential node will have the largest value of the average rank. The cross-correlation coefficients of the average normalized rank and all the seven normalized ranks were computed to verify that the average normalized rank had a fairly high correlation with all the seven normalized ranks and all these seven normalized ranks too were positively correlated among themselves.

5. Empirical Findings

We have investigated the evolution of the correlation network as well as the MST derived from the correlation network. A sharp change in the evolutionary pattern following the event of collapse of Lehman Brothers in the USA corresponding to period 7 has been observed. The evolutionary behaviour of the length of MST is shown in Figure 1. This clearly captures the increased level of synchronization in the global stock market at the onset of the recent financial crisis triggered by the collapse of Lehman Brothers as a decrease in the length of the MST is caused by an increased level of correlation between the stock indices. Similar evolutionary behaviour was observed in our recent study on the dynamics of stock returns in various stock markets [13] at the onset of recent crisis. The impact of the failure of Lehman Brothers and various other events on the stock market network is captured by significant decrease in the length of the MST. This shows that the correlation between stock indices around the globe has significantly increased following that event. The last point in the figure corresponding to period 37 represents the long-run behaviour of the length of the MST computed using all the 260 week data for the entire five year duration. Figure clearly reveals the contrasting difference between the short-run behaviour of the stock market network prior to and post Lehman Brothers failure.

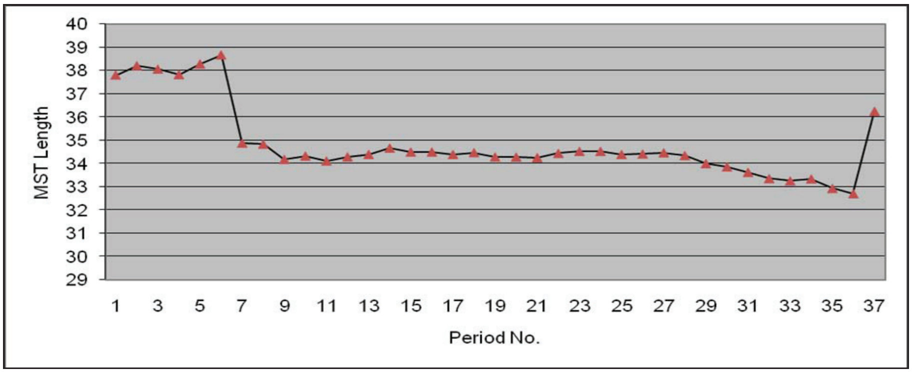

Fig. 1: Evolution of MST length (Period 37 represents the full 5 year period)

Figure 2 and Figure 3 respectively, show the MST constructed for period 6 corresponding to 120 week data prior to Lehman Brother failure and period 36 corresponding to 120 week data post Lehman Brother failure. The stock indices from different continent have been represented using different geometrical shapes and colours. North American and South American stock indices have been shown using boxes, with black colour for the US indices and white for all others. Asian stock indices have been shown using triangles, with black colour for Japanese indices, gray colour for Turkey and white for all others. European indices have been shown using ellipses, with black colour for the Europe indices, gray for UK and German indices and white for all others. Indices from Oceania are shown with diamond shape of black colour. The only index SEMDEX from Africa is shown with diamond shape of white colour.

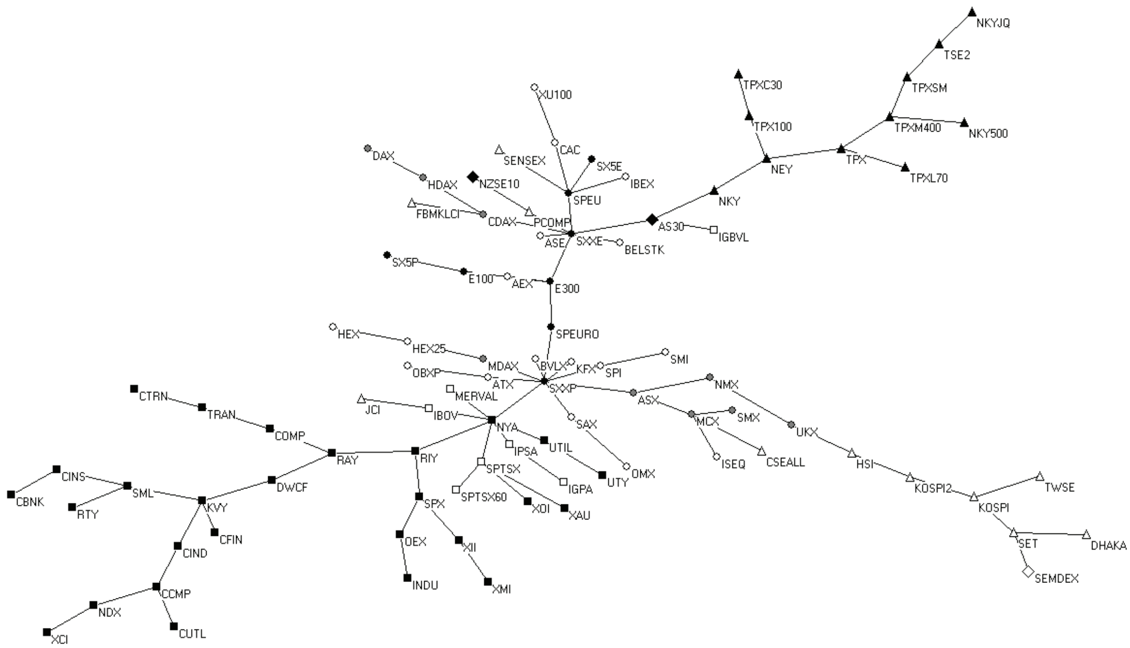

Fig. 2: MST corresponding to Period 6 (Pre Lehman Brothers failure)

The emergent pattern shows clusters of regional stock indices indicating that the network structure of stock indices is influenced by regional and trade interactions in a similar fashion as the network structure of stocks of global stock markets discussed in [14]. Japanese stocks form separate cluster from other Asian indices from other East Asian countries. The cluster of the Japanese stock indices gets linked to the European stock indices through Australian stock indices. Another cluster of East Asian stock indices are linked to the UK index through Hong Kong stock index. Turkey index XU100 from Middle East cluster with indices from Europe showing regional influence. The African stock index SEMDEX forms cluster with the Asian indices. Indices from Europe assume central positions resembling geographical linkages of the countries. A comparison between the two figures reveals some interesting findings about the impact of financial crisis on the topological structure of the network. One can clearly observe that the South American indices except Brazil's IBOV are clustered with Japanese and UK Indices rather than that with USA or Canada. The Indian index Sensex shows an interesting property that it moved out from the cluster of European indices to the clusters on Asian indices in the post Lehman Brothers failure period. This shows that Indian stock market could behave differently during the period of post Lehman Brothers failure indicating its better resilience to the market crash during the recent financial crisis.

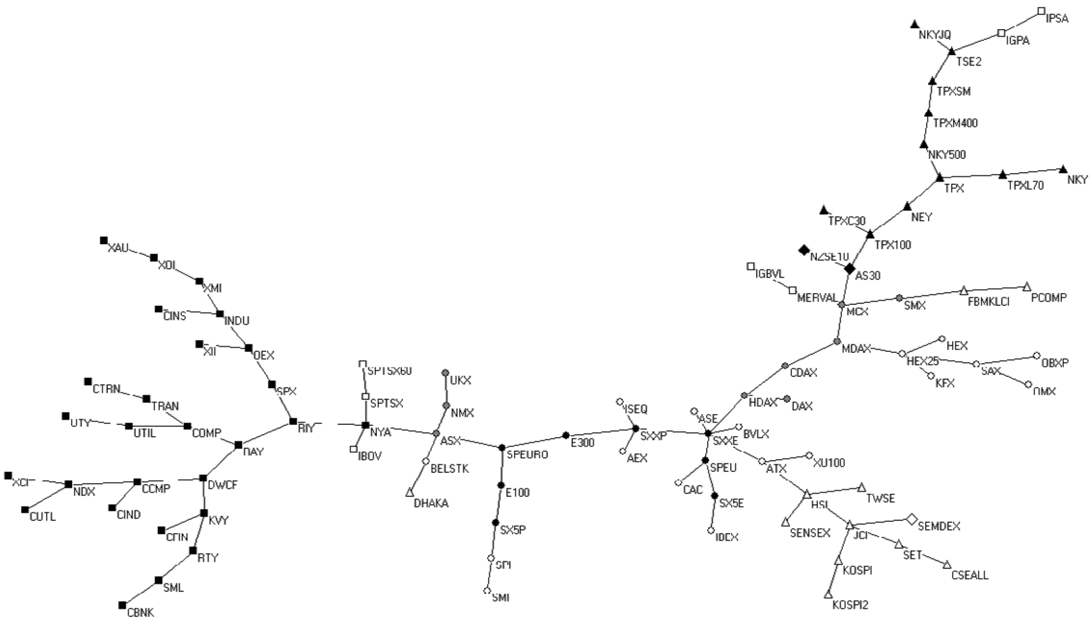

Fig. 3: MST corresponding to Period 36 (Post Lehman Brothers failure)

5.1. Influential stock indices

The ranks of all the indices have been computed based on centrality measures described in Section 4.1 for all the 37 periods. Table 2 shows the 10 most influential stock indices for some selected periods. SXXP and SXXE from Europe emerge as the most influential stock index in the social network of the global stock indices for pre and post Lehman Brothers collapse respectively. We observe that the ranks of the indices do not remain constant in time. The changes in ranks of stock indices ranked higher are relatively low compared to those ranked lower. We have also noticed that the range of change in ranks between period 1 and period 6 is -15 to +22 whereas the range of change in ranks between period 6 and period 7 is -30 to +30 pointing to turbulent market condition in some regions between period 6 and 7.

Table 2: Ranks and standard deviation of ranks for 10 most influential stock indices

Rank	Period 1	Period 6	Period 7	Period 36	Period 37	Country (Period 37)	Continent (Period 37)	Std. Dev of Rank
1	SXXP	SXXP	SXXE	SXXE	SXXE	Europe	Europe	2.3022
2	E300	NYA	SXXP	SXXP	E300	Europe	Europe	0.5477
3	NYA	E300	E300	E300	SXXP	Europe	Europe	0.8367
4	MDAX	SXXE	NYA	SPEURO	SPEU	Europe	Europe	1.9235
5	SPEURO	SPEURO	ASX	NYA	NYA	USA	N America	1.3038
6	SXXE	ASX	MDAX	SPEU	ASX	U K	Europe	1.1402
7	ASX	MDAX	SPEU	MDAX	SPEURO	Europe	Europe	1.6432
8	NMX	SPEU	SPEURO	ASX	ATX	Austria	Europe	7.5166
9	SPEU	NMX	NMX	ATX	CDAX	Germany	Europe	5.3572
10	DWCF	RIY	RAY	CDAX	MDAX	Germany	Europe	2.1679

It's interesting to note that the change in the ranks of influential stocks due to financial crisis is quite low as we can see from the table that the standard deviation of ranks of top seven leading indices corresponding to period 37 are within 3. We have also observed that the change in the ranks of indices between period 6 and period 36 are more that 10 for around one-third of the stock indices. Period 6 corresponds to 120 week data prior to the Lehman Brothers failure and period 36 corresponds to 120 week after the Lehman Brothers failure. This shows that the failure of Lehman Brothers and several other events during September 2008 had huge impact on the topology of stock market indices network.

6. Conclusions

We have presented a novel method to identify influential stock indices using centrality measures of social network analysis literature. Our findings indicate that SXXP and SXXE from Europe emerge as the most influential stock index in the social network of the global stock indices for pre and post Lehman Brothers collapse, respectively. The study also reveals an increased level of synchronization in the stock index returns around the collapse of Lehman Brothers in the USA. The study provides an empirical evidence of regional influences on the network dynamics. The regional clustering patterns point to the presence of different types of inter-market interaction and communication mechanism influencing the stock market behaviour. Such comovements of stock indices may be due to greater trade or social linkages between stock markets of different regions of the globe.

References

[1] Boginski, V., Butenko, S., and Pardalos, P.M. Mining market data: A network approach. Computers & Operations Research, 33 (2006), 3171–3184.

[2] Tse, C.K., Liu, J., and Lau, F.C.M. A network perspective of the stock market, Journal of Empirical Finance, doi:10.1016/j.jempfin.2010.04.008 (2010).

[3] Kim, K., Kim, S.Y., and Ha, D.-H. Characteristics of networks in financial markets., Computer Physics Communications, 177 (2007), 184–185.

[4] Huang, W.-Q., Zhuang, X.-T., and Yao, S. A network analysis of the Chinese stock market, Physica A, 388 (2009), 2956-64.

[5] Pan, R.K., and Sinha, S. Collective behaviour of stock price movements in an emerging market. Phy. Rev. E, 76, 4 (2007).

[6] Tabak, B.M., Takami, M.Y., Cajueiro, D.O., and Petiting, A. Quantifying price fluctuations in the Brazilian stock market. Physica A, 388 (2009), 59-62.

[7] Ravichandran, K., and Maloain, A.M. The Global Financial Crisis and Stock Market Linkages: Further Evidence on GCC Market, Journal of Money, Investment and Banking, 16, ISSN 1450-288X (2010).

[8] Morana, C., and Beltratti, A. Comovements in International Stock Markets. Journal of International Financial Markets Institutions and Money, 18 (2008), 31-45.

[9] Coelho, R., Gilmore, C.G., Lucey, B., Richmond, P., and Hutzler, S. The evolution of interdependence in world equity markets—Evidence from minimum spanning trees. Physica A: Stat. Mech. and its Applications, 376, 15 (2007), 455-466.

[10] Garas, A., and Argyrakis, P. Correlation study of the Athens Stock Exchange. Physics A, 380 (2007), 399-410.

[11] Freeman, L.C. Centrality in social networks: Conceptual clarification. Social Networks, 1, 3 (1979), 215-239.

[12] Bonacich, P. Power and Centrality: A Family of Measures. The American Journal of Sociology, 92, 5 (1987).

[13] Roy, R.B., and Sarkar, U.K. Capturing Early Warning Signal for Financial Crisis from the Dynamics of Stock Market Networks: Evidence from North American and Asian Stock Markets. Society for Computational Economics 16th International Conference on Computing in Economics and Finance, London, UK, (2010).

[14] Roy, R.B., and Sarkar, U.K. A social network approach to examine the role of influential stocks in shaping interdependence structure in global stock markets. *Forthcoming: International Conference on Advances in Social Network Analysis and Mining (ASONAM)*, Kaohsiung, Taiwan, 2011.