Graph Perspective of Stock Market Behaviour – Literature Review

Overview

It is often noticed in practice that the prices of different Stocks and other financial investment instruments move together. From economic and investment theoretical point of view this co-movements are very much logical. As many different companies engage in similar type of business. Also many a time one business depends on another business. So those companies are tied together by an invisible thread of relationship. Though it is difficult to find the actual relationship of the companies we can always measure the strength of their relationship by the similarity of movement of their attributes.

Correlation Coefficient is one of the measures of the strength of the relationship between two variables. We can assume if the correlation coefficient of one or more attributes of two Stocks or Financial instruments is more than a threshold value then those two stocks or the instruments are connected or we can say an Edge exists between them.

In this project our objective is to explore these relationships of the stocks from the Graph theory perspective and try to investigate different properties of those graphs.

Literature Review

For this topic I have reviewed about 20 different Academic Papers and consulted many other reference texts and websites as well. Some of the Literatures are based on the different generic approaches of Network Analysis of Stock Market, Some are for specific stock markets. Some papers talk about different important aspects of graph theory and statistical analysis techniques of the stock market in general. For my work I am particularly interested in the clustering of the graph. So I have studied more papers on the graph clustering techniques.

All the papers can be found in the following link for the ready reference.
https://www.dropbox.com/sh/y6fpladv1ovwtne/AAAUrKpp1BzI5gjED_Tvx0pta?dl=0

The academic papers those I studied are as follows :-

[In the description of the papers, I tried to keep the languages as close to the original text as possible]

1. Network_Analysis_of_the_Stock_Market : [Wenyue Sun, Chuan Tian, Guang Yang]

In this paper, they modeled a network for the US stock market based on the correlation of different stock returns. Community detection techniques were then applied to the constructed correlation network.

They found that resulting communities were consistent with the identified market sections that uses Standard Industrial Classification code, which demonstrates that performances of public stocks within the same sector tend to have similar patterns.

For network applications, They looked at the US credit crisis spreading represented by the spread of negative stock performances between Jul/2007 and Feb/2009. Stock performance were classified into 3 categories based on the rate of return (-20%, 0%, and 20%), then a sequence of snapshots of the stock network colored by the three defined categories is used to show the cascading behavior of stock performances.

They observed that the cascading starts from some stocks in different communities, and then spread from the initially infected (negative return) stocks. This observation indicates that the established stock network could potentially be very useful for stock market performance prediction from macroeconomic factors.

They also investigated the application of network in portfolio management. Traditional approaches of portfolio management often rely on certain statistical properties, such as expected return and price variance. These properties, however, generally represent the local behavior of the stocks. But with network analysis important global properties of the stock within the network can be extracted, such as degree centrality, betweenness centrality and closeness centrality.

In this work, they did some preliminary studies of creating investment portfolios using the network properties of different stocks, and their results show that with proper weights given to the top centrality nodes (i.e. stocks), this could outperform the S &P 500 for certain periods.

2. A network perspective of the stock market [Chi K.Tse, Jing Liu, Francis C.M. Lau]

Complex networks are constructed to study correlations between the closing prices for all US stocks that were traded over two periods of time (from July 2005 to August 2007; and from June 2007 to May 2009). The connections between the stocks are determined by cross correlations of the variations of the stock prices, price returns and trading volumes within a chosen period of time. Specifically, a winner-take-all approach is used to determine if two nodes are connected by an edge.

They are the first who attempted to construct a full network of US stock prices that gives full information about their interdependence.

They found that all networks based on connecting stocks of highly correlated stock prices, price returns and trading volumes, display a scale-free (Power law) degree distribution. Also this work clearly suggested that the variations of stock prices are strongly influenced by a relatively small number of stocks.

From the composition of the highly connected stocks, they found that the market is heavily dominated by stocks in the financial sector.

They proposed a new approach for selecting stocks for inclusion in a stock index and compare it with the existing indexes.

3. Complex networks in a stock market [Kyoung Eun Lee, Jae Woo Lee, Byoung Hee Hong]

In this paper they consider cross-correlations among stock prices in the Korean stock-market. They uses the daily Korean stock market prices of KOSPI 200 for 4 years from January 3, 2000 to December 29, 2004.

They calculate the cross correlations between the stock prices using their Log Return. The metric distance between a pair of stocks is defined by SQRT(2*(1-CorrelationCoefficient))

Based on this distance they constructed a minimal spanning tree (MST) from the distance matrix in MST graph. They obtained the average shortest path length 5.05. Also they observed that the degree distribution of the MST follow a power-law [with $\gamma = 2.5$]

4. **Identifying influential stock indices from global stock markets: A social network analysis approach [Ram Babu Roy, Uttam Kumar Sarkar]**

In this article they have proposed a method to rank the stock indices from across the globe using social network analysis approach.

They have used the weekly returns of 93 stock indices for five-year period from the year 2006 through 2010 obtained from Bloomberg.

The temporal evolution of correlation network and Minimum Spanning Tree (MST) of global stock indices have been analyzed before and after the collapse of Lehman Brothers in the USA.

They studied about identifying the most influential stock indices in the global stock market, regional influence on the co-movement of stock indices, and the impact of the collapse of Lehman Brothers in the USA and the associated global financial crisis using the dynamics of stock market network.

5. **Mining market data-A network approach [Vladimir Boginski, Sergiy Butenk, Panos M.Pardalos]**

They constructed the Market Graph by calculating cross-correlations between pairs of stocks based on the opening prices data over a certain period of time.

They studied the evolution of the structural properties of the market graph over time.

They found out that the power-law structure of the market graph is quite stable over the considered time intervals; therefore the stock market can be considered as a "self-organized" system.

Another important result they found is the fact that the edge density of the market graph, as well as the maximum clique size, steadily increased during the last several years, which supports the well-known idea about the globalization of economy which has been widely discussed.

They have also indicated the natural way of dividing the set of financial instruments into groups of similar objects (clustering) by computing a clique partition of the market graph. Their methodology can be extended by considering quasi-cliques in the partition.

6. **Community detection in graphs [Santo Fortunato]**

This paper deals with one of the most important of the Graph Theory.

One of the most relevant features of graphs representing real systems is community structure, or clustering, i.e. the organization of vertices in clusters, with many edges joining vertices of the same cluster and comparatively few edges joining vertices of different clusters.

Such clusters, or communities, can be considered as fairly independent compartments of a graph, playing a similar role where systems are often represented as graphs.

This problem is very hard and not yet satisfactorily solved, despite the huge effort of a large interdisciplinary community of scientists working on it over the past few years.

This paper tries to consolidate the works done so far.

7. **Graph clustering [Satu Elisa Schaeffer]**

This is a survey covering the different aspects of graph clustering. It reviews different definitions of Graph Clustering and the Measure of Cluster Quality. This also presents global algorithms for producing a clustering for the entire vertex set of an input graph, after which we discuss the task of identifying a cluster for a specific seed vertex by local computation. Some ideas on the application areas of graph clustering algorithms are given. This also addresses the problems of evaluating clusters and benchmarking different cluster algorithms.

8. **Spectral methods for graph clustering − A survey [Maria C.V. Nascimento, Andre C.P.L.F. de Carvalho]**

This is also a Survey. But this focuses on a particular class of graph clustering algorithms, known as spectral clustering algorithms. These algorithms are mostly based on the Eigen-Decomposition of Laplacian matrices of either weighted or unweighted graphs. This survey presents different graph clustering formulations, most of which based on graph cut and partitioning problems, and describes the main spectral clustering algorithms found in literature that solve these problems.

9. **Unveiling the connectivity structure of financial networks via high-frequency analysis [Donatello Materassi, Giacomo Innocenti]**

The paper deals with the problem of reconstructing the internal link structure of a network of agents subject to mutual dependencies. They are opining that standard multivariate approaches based on a correlation analysis are not well suited to detect mutual influences and dependencies, especially in the presence of delayed or propagative relations and when the sampling rate is sufficiently high to capture them.

In particular, they developed and applied a metric based on the coherence function to take into account these dynamical phenomena. The effectiveness of the proposed approach is illustrated through numerical examples and through the analysis of a real complex networked system, i.e. a set of 100 high volume stocks of the New York Stock Exchange, observed during March 2008 and sampled at high frequency.

10. **Network analysis of a financial market based on genuine correlation and threshold method [A. Namaki, A.H. Shirazi, R. Raei, G.R. Jafari]**

This paper uses Random Matrix Theory (RMT) notion for specifying the largest eigenvector of correlation matrix as the market mode of stock network, which can be thought of an adaptive complex network consisting of many interacting units.

For a better risk management, they cleaned the correlation matrix by removing the market mode from data and then construct this matrix based on the residuals. They show that this technique has an important effect on correlation coefficient distribution by applying it for Dow Jones Industrial Average (DJIA).

To study the topological structure of a network they applied the “removing market mode” technique and the threshold method. They used Tehran Stock Exchange (TSE) as an example. They showed that this network follows a power-law model in certain intervals. They also showed the behavior of clustering coefficients and component numbers of this network for different thresholds. These outputs are useful for both theoretical and practical purposes such as asset allocation and risk management.

11. **A social network model of investment behaviour in the stock market [L. Bakker, W. Hareb, H. Khosravi, B. Ramadanovic]**

This paper deals with the "trust factor" by studying different trader behaviors.
To consider the psychological factors that impact market valuation, a model is formulated for investment behaviour of traders whose decisions are influenced by their trusted peers' behavior. The model is implemented and several different ``trust networks'' are tested. Simulation results demonstrate that real life trust networks can significantly delay the stabilization of a market.

12. **The Relation between Degree and Strength in the Complex Network Derived from an Individual Stock [Zelin Zhang, Maokang Luo,Ke Deng, and Li Lai]**

The relationship of Connections (Degree) and their influences (Strength) are studied for individual stocks in a market network.

A method based on coarse-graining to construct a directed weighted complex network which models the transformation of the trading data of an individual stock is introduced. The degree (strength) distribution of derived network follows a power-law. A moderated regression equation with interaction effects of average return and out-degree (in-degree) on out-strength (in-strength) is established. Moreover, they found that the differences of nodes affect the network's structure and average return level impacts nodes' eigenvector centrality and page-rank, significantly.

13. **Structure Characteristics of the International Stock Market Complex Network in the Perspective of Whole and Part [Guangxi Cao, Yingying Shi and Qingchen Li]**

International stock market forms an abstract complex network through the fluctuation correlation of stock price index. Most of the past studies of complex network almost focus on single country's stock market. This paper we investigate the whole and partial characteristics of international stock market network.

For the analysis on the whole network, they first determined the reasonable threshold as the basic of the following study. Robustness is applied to analyze the stability of the network and the result shows that the network has robustness against random attack but intentional attack breaks the connection integrity of the network rapidly.

In the partial network, the sliding window method is used to analyze the dynamic evolution of the relationship between the Chinese (Shanghai) stock market and the international stock market. The connections between the Chinese stock market and foreign stock markets have become increasingly closer and these show a significant enhancement especially after China joined the WTO.

In general, they conclude that transnational investors pay more attention to some significant event of the stock market with large degree for better risk-mitigation.

14. **Correlation of financial markets in times of crisis [Leonidas Sandoval Junior, Italo De Paula Franca]**

Using the eigenvalues and eigenvectors of correlations matrices of some of the main financial market indices in the world, this paper show that high volatility of markets is directly linked with strong correlations between them. This means that markets tend to behave as one during great crashes.

In order to do so, they investigate financial market crises that occurred in the years 1987 (Black Monday), 1998 (Russian crisis), 2001 (Burst of the dotcom bubble and September 11), and 2008 (Subprime Mortgage Crisis), which mark some of the largest downturns of financial markets in the last three decades.

15. **Statistical analysis of Financial networks [VladimirBoginski, Sergiy Butenko , Panos M. Pardalos]**

In many practically important cases, a massive dataset can be represented as a very large graph with certain attributes associated with its vertices and edges. Studying the structure of this graph is essential for understanding the structural properties of the application it represents. It turns out that the degree distributions of most of these graphs can be described by the power-law model.

This paper conduct the statistical analysis of the stock market graph and showed that it also follows the power-law model. Moreover, it detects cliques and independent sets in this graph. These special formations have a clear practical interpretation, and their analysis allows one to apply a new data mining technique of classifying financial instruments based on stock prices data, which provides a deeper insight into the internal structure of the stock market.

16. **A network analysis of the Chinese stock market [Wei-Qiang Huang, Xin-Tian Zhuang, Shuang Yao]**

In this paper, they used a threshold method to construct China's stock correlation network and then study the network's structural properties and topological stability. They conduct a statistical analysis of the network and showed that it follows a power-law model. They also detect components, cliques and independent sets in this network.

Their analysis allows one to apply a new data mining technique of classifying financial instruments based on stock price data, which provides a deeper insight into the internal structure of the market.

They also test the topological stability of this network and find that it displays a topological robustness against random vertex failures, but it is also fragile to intentional attacks. Such a network stability property would be also useful for portfolio investment and risk management.

17. **Network Topologies of Shanghai Stock Index [Jianhua Zhang, Huaxi Zhou, Lu Jiang, Yougui Wang]**

In this paper, they analyzed time series of Shanghai stock index with complex network theory. The degree distribution of the network extracted from the original series is found to be well fitted with a power law, while the network from return series is governed by an exponential degree distribution.

Compared with the time series of standard Brownian motion, they found that the dynamics of the original series can be identified, but the return series has the similar topology with a random one. Moreover, in the scale-free networks from original series, the small-world property is detected and the time interval distribution between connected pairs decays as an exponential function, which implies that nodes correlated with a given one appear in a Poisson process.

18. **Extracting hidden fluctuation patterns of Hang Seng stock index from network topologies [Ping Li, Bing-Hong Wang]**

This paper presents a model of complex network generated from Hang Seng index (HSI) of Hong Kong stock market, which encodes stock market relevant both interconnections and interactions between fluctuation patterns of HSI in the network topologies.

In the network, the nodes (edges) represent all kinds of patterns of HSI fluctuation (their interconnections).

Based on network topological statistic, this paper presents efficient algorithms, measuring betweenness centrality (BC) and inverse participation ratio (IPR) of network adjacency matrix, for detecting topological important nodes. They have obtained at least three uniform nodes of topological importance, and find the three nodes, i.e. 18.7% nodes undertake 71.9% betweenness centrality and closely correlate other nodes. From these topological important nodes, they could extract hidden significant fluctuation patterns of HSI. They also find these patterns are independent the time intervals scales.

The results contain important physical implication, i.e. the significant patterns play much more important roles in both information control and transport of stock market, and should be useful to more understand fluctuations regularity of stock market index.

They have concluded that Hong Kong stock market, rather than a random system, is statistically stable, by comparison to random networks.

19. A review of two decades of correlations, hierarchies, networks and clustering in financial markets [Gautier Marti, Frank Nielsen, Mikołaj Binkowskia, Philippe Donnat]

This is an extremely helpful document for the research.

This document tries to gather in one place the relevant material that can help the researcher in the field to have a bigger picture, the quantitative researcher to play with this alternative modelling of the financial time series, and the decision maker to leverage the insights obtained from these methods.

This document forms a basis for implementation of an open toolbox of standard tools to study correlations, hierarchies, networks and clustering in financial markets. They put their research in an website called “www.datagrapple.com/tech”.

20. A Network Analysis of the Greek Stock Market [Kydros Dimitriosa , Oumbailis Vasileios]

This gives the network perspective of the stock market which is relatively small.

In this paper they analyse stock relationships in the Greek Stock Market. They propose a model that can depict such relationships and create networks of stocks. They investigate all stocks in the Greek Stock Market for years 2007 and 2012 (one year before and during the current economic crisis).

Different networks are created according to the degree of correlation of stocks. These networks are visualized and evaluated, using methods from Social Network Analysis. A number of metrics, mainly centrality measurements, are calculated and interpreted. They have discussed the hypothesis that Greek stocks follow the “herd” rule and investigate the role of important actors (stocks) in these networks.

Their results show that the Greek Market is a “shallow” market, easily affected by a few big investors or the economic climate