A PROPOSED ALGORITHM OF RANDOM VECTOR IN MEASURING SIMILARITY FOR NETWORK TOPOLOGY OF BURSA MALAYSIA

LIM SAN YEE

A thesis submitted in
fulfillment of the requirement for the award of the
Degree of Master of Science

Faculty of Applied Sciences and Technology
Universiti Tun Hussein Onn Malaysia

OCTOBER 2018

ACKNOWLEDGEMENT

I would like to express my sincere appreciations to my supervisor, Dr. Rohayu Mohd Salleh and co-supervisor, Pn. Norhaidah Mohd Asrah for the supports given throughout the duration for this research. Without the supports from my supervisors, I believe that this research was unable to complete. The helps given by my friends, Low Yong Sun, Simon Khoo and Joshua Chieng were also highly appreciated. They lend their hands to me when I faced some problems in the research, both in technical ways and moral supports. I would like to thank my family members who support me physically and mentally throughout the duration of this research. Last but not least, appreciation goes to everyone involved directly or indirectly towards the compilation of this thesis.

ABSTRACT

The stock market is a complex system where the interrelationships between the stocks are complicated because it is in multivariate time series setting which consists of opening, highest, lowest and closing prices. Basically, the Pearson correlation coefficient (PCC) is applied to measure the similarity between two or more univariate time series of stocks. However, the economic information from other variables may inaccurate if the analysis is conducted by applying single variable only. Therefore, multi-dimensional of stocks are considered in this thesis. The similarities between two or more multi-dimensional of stocks are quantified by using Random Vector (RV) coefficient. Based on the preliminary analysis, the computational of RV coefficient is difficult, time-consuming, and tedious when a large number of stocks are involved. Hence, to ease the calculation process and improve the computational efficiency of RV coefficient, an algorithm is proposed. The proposed algorithm is able to measure the similarities among all pairs of stocks in Bursa Malaysia at once. The calculation process of RV coefficient among all pairs of stocks can be shortened and eased as the proposed algorithm consists of time complexity of order of $O(n^2)$. The behaviors and interactions among the stocks in Bursa Malaysia are then determined by using the Forest of all possible minimum spanning trees. In this thesis, MK Land Holdings Berhad was found out to be the predominant stock in Bursa Malaysia as it displays a star-like structure and is located at the central hub of the network.

ABSTRAK

Pasaran saham merupakan sistem kompleks di mana hubungkait di antara saham adalah rumit disebabkan saham berbentuk siri masa multivariat yang terdiri dari harga pembukaan, harga tertinggi, harga terendah, dan harga penutupan. Secara asasnya, pekali kolerasi *Pearson* (PCC) digunakan untuk mengukur kesamaan antara dua atau lebih saham berbentuk siri masa univariat. Walau bagaimanapun, maklumat ekonomi yang diperolehi daripada pembolehubah-pembolehubah yang lain berkemungkinan kurang tepat sekiranya kajian dijalankan dengan menggunakan satu pembolehubah sahaja. Sehubungan itu, dalam tesis ini, saham dalam pelbagai dimensi akan diambil kira. Kesamaan antara dua atau lebih pelbagai dimensi saham akan diukur dengan menggunakan pekali vektor rawak (RV). Berdasarkan analisis awal, pengiraan pekali RV adalah sangat sukar, mengambil masa yang panjang dan rumit apabila melibatkan bilangan saham yang besar. Oleh itu, untuk memperbaiki kesukaran dan kecekapan pengiraan pekali RV, algoritma pekali RV dicadangkan. Algoritma yang dicadangkan dapat mengukur kesamaan antara semua pasangan saham di Bursa Malaysia secara serentak. Proses pengiraan pekali RV antara semua pasangan saham dapat dipendekkan dan dipermudahkan kerana algoritma tersebut mempunyai kerumitan masa susunan $O(n^2)$. Tingkah laku saham dan interaksi di antara saham di Bursa Malaysia akan ditentukan dengan menggunakan '*Forest of all possible Minimum Spanning Trees*'. Dalam tesis ini, MK Land Holdings Berhad dikenalpasti sebagai saham yang dominan di pasaran Bursa Malaysia kerana ia memaparkan struktur bintang dan kedudukannya terletak di hab pusat rangkaian.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

CNN	-	Cable News Network
DCCA	-	Detrended cross-correlation analysis
DTW	-	Dynamic time warping
DXA	-	Detrended cross-correlation analysis
EVC	-	Escoufier vector correlation
Forest	-	Forest of all possible MSTs
GBM	-	Geometric Brownian motion
HT	-	Hierarchical tree
IPC	-	Infrastructure project company
MST	-	Minimum spanning tree
PCA	-	Principal component analysis
PCC	-	Pearson correlation coefficient
REIT	-	Real estate investment trust
RMT	-	Random matrix theory
RV	-	Random vector
SDU	-	Subdominant ultrametric space
OHLC	-	Opening, highest, lowest and closing prices

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

In recent times, the stock has become one of the most important trading objects in the financial market. Every company uses its own organization's shares to trade in the stock market in order to raise capital for their company's expansion purposes. Furthermore, every stock market in the world has its own network structure that is both deeply rooted and complex. The complicated network structure of the stock market is the main focus of investigation in this research. The background of the research is then discussed in Section 1.1. Section 1.2 states the problem statement that has described in this research. In order to solve the problem, there are four research objectives to achieve in this research as stated in Section 1.3. Section 1.4 describes the scope of the research and the significance of study is presented in Section 1.5. Lastly, the organization of the thesis is outlined in Section 1.6.

1.1 Research background

The financial market is a marketplace where all buyers and sellers meet daily to trade securities such as stocks, bonds, derivatives, foreign exchange, and commodities. The transactions of the securities can be done either through physical location or electronic system. As an example, the traders can buy or sell the securities through a physical location, such as the New York Stock Exchange (NYSE), or an electronic system such as the NASDAQ. There are several types of financial market, such as equity market, commodity market, derivatives and stock market, but the stock market is one of the largest markets in the world. It is the place where all the shares of publicly listed companies can be purchased or sold through brokers and traders every day. More importantly, the stock market is considered as an economic indicator of a nation's financial health since it reflects how well all the listed companies are doing.

Each country has its own stock market which is known as the stock exchange. For instance, in Malaysia, the stock exchange is known as Bursa Malaysia. Regarding the Bursa Malaysia (2017), there are thirteen sectors under the Malaysian stock market, such as finance, technology, construction, consumer products, industrial products, mining, plantation, property, hotels, closed-fund, real estate investment trusts (REITs), infrastructure project company (IPC) and trading or services. Every sector has some certain amount of stocks and each stock is related to each other. For example, according to the works that have conducted by Gan & Djauhari (2012a), the construction stocks are straightforwardly connected to the financial stocks. This finding demonstrates that the two divisions of stocks are related to one another. Thus, it can be said that any adjustments in the stock prices will impact other organizations' stock prices as well.

The volatility in the stock market can have a significant economic effect on the economy and individual consumers (Amadeo, 2017). One of the main factors that affect the volatility in the stock market is the confidence among investors towards the corporations. For instance, when investors lose confidence towards the companies, the investors will not invest in those companies. In turn, the companies will lose the raised capital, then the economic growth of a nation will stagnate. Basically, a corporation's condition or circumstance can determine through its stock price, as stock price is the reflection of the corresponding company. At every single minute of the transaction hours, the stock price keeps changing from time to time and does not merely stagnate at a certain price all the time. Besides, the system of the stock market is constantly developing through different heterogeneous associations among the stocks (Kazemilari & Djauhari, 2015). The interrelationships and interactions between the stocks are muddled as the price fluctuates ceaselessly every now and then. Therefore, it can be said that the stock market constitutes a complex system.

In fact, the stock market is considered as a good example of the complex system due to the mass of complicated interrelationships among the stocks. A complex system is a collection of many interdependent components that interact with each other through competitive nonlinear collaboration which leads to emergent and self-organized behavior. Furthermore, regarding the New England Complex Systems Institute (NECSI), the complex system is about how parts of a system leads to the group behavior and how the system interacts with its environment. There are many methods have been using in investigating the complexity of the stock market today such as minimum spanning tree (MST) (Mantegna, 1999), planar maximally filtered graph (Tumminello *et al*., 2005), average linkage-based MSTs (Tumminello *et al*., 2007), and Directed Bubble Hierarchical Tree (Musmeci, Aste & Di Matteo, 2015). Among these approaches, MST is the most popular and the easiest method in analyzing the stock network as seen in the studies by Situngkir & Surya (2005), Coelho *et al.* (2007a), Coelho *et al.* (2007b), Brida & Risso (2008), Gilmore, Lucey & Boscia (2008), Sieczka & Holyst (2009), Tabak, Serra & Cajueiro (2010), Kumar & Deo (2012), Gan & Djauhari (2013), Djauhari & Gan (2015b), Bahaludin, Abdullah & Mat Salleh (2015) and Kazemilari & Djauhari (2015).

The main idea of MST is to break down a complex project's data into its component parts and plotting them to show their interdependencies and interrelationships. In current practice, stocks are investigated by univariate time series of its closing price only, whereby the Pearson correlation coefficient (PCC) is the standard approach to measure the similarity or, equivalently, interrelationships among the stocks where each stock is represented by univariate time series.

Nevertheless, there are some probabilities of losing embodying information from other variables if the analysis applies only to a single variable (Brida & Risso, 2008). Moreover, the phenomenon of social embeddedness can only be detected by a multivariate approach but not by a univariate approach which has mentioned in Halinen & Tornroos (1998). Therefore, it can be said that the information retrieve from the investigation by only utilizing univariate time series is misdirecting as each stock is spoken to by four components of the prices which are opening, highest, lowest and closing prices. Hence, there is a need to investigate the relationships between two or more multidimensional of stocks since every stock is represented by four types of price.

There are several methods to deal with the analysis of the stock market in multivariate setting such as canonical correlation analysis (Hotelling, 1936), factor analysis, vector correlation coefficient (Escoufier, 1973) and principal component analysis (Jolliffe, 1986). Historically, Hotelling (1936) proposed the canonical correlation in the structure of canonical correlation analysis which is used to interpret the linear dependencies between two sets of variables. However, in Josse, Pages & Husson (2008), they remarked that the relationship between the two groups of variables, which is viewed by canonical correlation, is not a global measurement of the intensity of the association since it is expressed component by component.

The vector correlation coefficient which was introduced by Escoufier (1973) provides the notion of global association in multivariate setting (Josse *et al*., 2008). This method assumes that two sets of vectors are superbly correlated if there exists an orthogonal change that can influence one set to agree with the other. The Escoufier vector correlation (EVC), which is also known as random vector (RV) coefficient, can be used as a unifying tool in several multivariate methods (Robert & Escoufier, 1976). It is then applied in measuring the similarity of the stock prices in order to determine the interrelationships among the stocks in the market in multivariate setting.

Nonetheless, the computational procedure of RV coefficient is complex and time-consuming as RV coefficient is only able to measure the similarity between two multivariate time series of stocks at once. The larger the number of input data, the computational efficiency of RV coefficient is getting worse as well as the time taken to calculate the RV coefficient becomes longer. There is an available built-in function of RV coefficient that can be found in the R software, however, based on our preliminary analysis, the values given from this function are different from the values that are calculated manually. Therefore, it is necessary to construct a new algorithm of RV coefficient in order to obtain the accurate RV coefficient values and reduce the computing time of RV coefficient.

Generally, the stock price keeps floating ceaselessly and is unstable all the time. Thus, it is hard to determine the behaviors of the relationships among the stocks in Bursa Malaysia. Besides, it is also difficult to identify the most influential as well as the most dominant stock in Bursa Malaysia due to the inconstancy of values and complexity of the stock market. There are numerous uncertainties about how this particular stock affects the other stocks' behaviors and how are the interrelationships among the stocks. The majority of the traders are confused about which stock appears to be assuming a dominant role in the stock market within a certain period of time.

Therefore, the complexity of network structure of Malaysia's stock market in the multivariate setting is highly focused in this research. The similarities among the stocks where each stock is represented by multivariate time series are measured by applying the proposed RV coefficient algorithm since the computational of RV coefficient by manual and built-in function in R software are both inefficient. In addition, the interrelationships among the stocks are analyzed by plotting the network topology of the stock market as well as the most dominant and influential stock is investigated in order to help in eliminating the uncertainties. Last but not least, the information provided in this research can be very helpful to the financial market, especially for the financial manager to arrange and construct the stock portfolio and risk management.

1.2 Problem statement

In the literature, there are many studies have conducted to analyze the complex system of the financial market such as the stock market. It has turned out to be a basic component to comprehend the behaviors of the securities and the role played by the stocks in these systems. In current practice, the PCC is used to measure the similarity among stocks due to the stock market is considered as a complex system of univariate time series. However, in daily financial market activities, the stock market is multivariate time series of opening, highest, lowest and closing prices. The information extracted from the analysis by using univariate time series only is incomplete and misleading (Brida & Risso, 2008; Halinen & Tornroos, 1998). Since the stock values keep changing and floating from time to time, it is hard to determine the behaviors of the stocks as well as the interactions among the stocks in the stock market (Mantegna, 1999; Mantegna & Stanley, 2000). To have a complete and comprehensive stock network analysis, four components of stock price are taken into account in this research.

In order to measure the similarities among the multivariate time series of stocks, RV coefficient is used. However, the computational procedure of RV coefficient by manual calculation and built-in function in R software are both complex and time-consuming as both tools are only able to quantify two multivariate time series of stocks at once. Therefore, when the number of stocks gets larger, the efficiency of the RV coefficient gets worsen in terms of the amount of time taken to compute the RV coefficient values. Furthermore, the RV coefficient values compute by the existing built-in function in R software are somehow different to the manual calculation. Thus, based on the findings found, it is necessary to build a new algorithm in Matlab software that is able to provide similar RV coefficient values to the manual calculation values in Microsoft Excel and computationally more efficient than the manual calculation as well as the available built-in function in R software. Some validation of studies by using the three tools, which are built-in function in R software (Tool 1), manual calculation in Microsoft Excel (Tool 2) and proposed algorithm in Matlab software (Tool 3), are conducted in order to show the advantages of the proposed algorithm.

1.3 Research objectives

In this research, there are four objectives as follow:

(i) To build an RV coefficient algorithm.

(ii) To validate the proposed algorithm.

(iii) To construct similarity measure among Malaysian stock market components where each of stock is represented as multivariate time series.

(iv) To investigate the behaviors of the interrelationships among the Malaysian stock market by applying forest of all possible MSTs and centrality measures.

1.4 Research scope

The scope of the research can be divided into three aspects. First, the theoretical aspect. This aspect includes the use of RV coefficient in measuring the similarities among the stocks where each of the stock is represented as four components of price. Then, the stock network in multivariate setting is constructed by using forest of all possible MSTs. The proving of RV coefficient is a multivariate generalization of PCC also involves in this aspect.

Second, the computational aspect. In order to show that, in terms of accuracy, the algorithm developed with Matlab software has higher accuracy than the existing built-in function in R software, empirical and simulation studies are conducted. To prove that the algorithm built has higher computational efficiency than the manual calculation and the available function in R software, time complexity of the algorithm is shown.

Third, the practical aspect. This aspect covers the application of stock network analysis in multivariate setting in real stock market. The stock data used in this research are 738 weekly data of Main Market Bursa Malaysia stocks. The period of investigation of the data collection is taken in weekly basis from August 10, 2015 to

October 1, 2017 from Yahoo Finance (2017). The analysis is conducted by applying Microsoft Excel, Matlab version 2013a, R version 3.3.1 and Pajek software.

1.5 Significance of the research

The process of measuring the similarity among the stocks by using RV coefficient is complicated and complex. This is because the RV coefficient is only able to calculate the similarity between a pair of multivariate time series of stocks at once. Thus, by proposing a new algorithm of RV coefficient with Matlab software, the calculation process is eased and the computing time of RV coefficient values is shorten. Besides, with the algorithm proposed, the computational complexity problem can be solved. By analyzing the Bursa Malaysia stock network structure, it provides the visualization of the stock market structure to the stock firms. Therefore, this study not only eases the calculation process of RV coefficient but also provides useful economic information to the risk management and portfolio construction.

1.6 Organization of the thesis

This thesis is organized into five chapters. Chapter 1 is the introduction of the research. This chapter includes research background, problem statement, research objectives, research scopes and significance of the research.

Chapter 2 is the literature review of the Malaysian stock market as well as the similarity measures among the stocks in terms of univariate and multivariate settings. The popular methods in dealing the complex network – correlations-based network analysis and the centrality measures are also included in Chapter 2.

Chapter 3 describes the methodologies used in this research in three sections. The first section is the validity of the proposed algorithm by comparing the computational accuracy and efficiency of the built-in function in R software (Tool 1), manual calculation (Tool 2) and proposed algorithm in Matlab software (Tool 3). The second section is about how the analysis of the research carries out with the details of every method used. Then in the third section, functions of centrality measures such as degree, betweenness, closeness, eigenvector and overall centrality are presented.

Chapter 4 presents the results of the empirical and simulation studies of three tools in terms of computational accuracy and efficiency. Besides, this chapter also presents and discusses the results of the application of the methodology in real world data which is the Main Market of Bursa Malaysia. There are total 738 stocks traded in Main Market of Bursa Malaysia are analyzed.

Chapter 5 is the closing chapter of the presentation of this thesis which consists of a brief summary and conclusions of the results of the thesis. Despite this, contributions of the study and future research directions are also included in this chapter.

1.7 Summary

This chapter discussed about the stock market. Besides, it also discussed about the complex system of stocks and the significance of network analysis in the economic field. Moreover, it can be seen as necessary to analyze the stock market in multivariate setting to ensure the information obtained is complete. Some research gaps in this field had found. Thus, this lead to the formation of this research.

CHAPTER 2

LITERATURE REVIEW

This chapter presents the review of the literature on the Malaysian stock market as well as the similarity measures among the stocks in the stock market based on network analysis approach. This chapter begins with the structure of the Malaysian stock market as one of the largest complex system in the financial market and reviews of historical background in Section 2.1. Then, Section 2.2 and Section 2.3 discuss the similarity measures among the stocks where each stock is represented by univariate time series and multivariate time series respectively. The computational of multivariate analysis techniques is reviewed shortly in Section 2.4. Section 2.5 is then focused on the literature on the correlations-based network analysis since this method is the most popular method in dealing with the complex network. Later on, network topological algorithm is briefly explained in Section 2.6 followed by centrality measures in Section 2.7.

2.1 Malaysian stock market

Other than foreign exchange market, the stock market is one of the most important parts of today's global economy. Nations around the world rely on stock markets for their financial developments. The stock market is a cornerstone where companies sell their shares to raise up the capitals for their businesses. Every country has their own stock markets which known as exchange market. In Malaysia, the Malaysian stock market is known as Bursa Malaysia. However, it was previously known as the Kuala Lumpur Stock Exchange. Bursa Malaysia is one of the biggest markets in Southeast Asia (Yeoh, Hooy & Asrad, 2010) with the total market capitalization of USD 5.44

billion in 2017. Since the early 1990s, Bursa Malaysia has encountered high development regarding market capitalization and trading value. Moreover, the number of companies listed in the exchange market has increased from 271 companies in 1991 to 806 companies in 2017 which has increased 297.42% within 26 years gap. All these improvements indicate that the Malaysian economy has experienced profound structural changes over the most recent couple of decades, developing from a primary producer into an increasingly diversified and broad-based economy (Yeoh *et. al.*, 2010).

Bursa Malaysia has market dynamics that seems to be the more complicated system where the market dynamics describes the changing of pricing signal that outcome from the persistent changes in both supply and demand of any specific stock. The changes of daily stock prices are not only affected by the corporation's own fundamentals (Djauhari & Gan, 2015b) but are also influenced by the other issues such as economic factors and volatility in the market. Furthermore, each stock in Bursa Malaysia has some interactions with other stocks. The interactions among the stocks are ambiguous and confusing as it is hard to determine which particular stock has an interrelationship with the other stocks accurately. Hence, it can be said that Malaysian stock market constitutes a complex system.

There are few studies related to Bursa Malaysia have carried out in order to study the complex network structure of the stock market. In 2012, Sharif & Djauhari (2012a) analyzed stock network analysis of 90 daily closing price of stocks in Bursa Malaysia from January 1, 2007 until December 31, 2009 by applying minimum spanning tree (MST). As a result, CIMB Group Banking, Genting Bhd, RHB Capital Bhd, Wah Seong Corp, MMC Corp Bhd and IJM Corp Berhad were the most influencing stocks among the other stocks where the investors should pay more attention to these six stocks for their investment plans.

Subsequently, Gan & Djauhari (2012a) studied about the analysis of correlation structure among stocks traded in Bursa Malaysia. They analyzed 100 most capitalized stocks in Bursa Malaysia from January 2007 to January 2009 by using MST where each stock is represented by closing price only. In this research, they showed that Bursa Malaysia Berhad stock is the most influencing stocks in the stock network. Another study is the work of Bahaludin *et al*. (2015) who investigated the 100 most capitalized stocks in Bursa Malaysia by using the daily closing price of stocks from

January 1, 2011 until December 31, 2013. Through their analysis, they found out that UEM Sunrise Berhad was the most crucial stock in Bursa Malaysia.

Based on these three studies that have done, it can be said that different investigation period of data, as well as a different number of stocks data use, will provide assorted results in the analysis. Due to the constantly evolving of the stock prices, it is not necessary that the most significant stock obtained from the analysis must be under either the Trading or Services sector or Industrial Products sector, although these two sectors are the main elements in Malaysia's economy since 2006 (Azer *et al*., 2016).

2.2 Similarity measure among univariate time series

Univariate time series is defined as a time series that consists of single variable that is collected sequentially over regular time interval. In stock market, a time series tracks a stock price movement over a specified period of time with price recorded at regular time intervals. The stock price is a reflection of the corresponding company's condition or circumstance. Any factors such as the volatility in the market or reputation of the corresponding company get affected will directly influence the stock price. Hence, studying the similarity among the stock price changes is able to help the investors and portfolio managers to construct the investment portfolio.

The system of the stock market is ceaselessly developing through different heterogeneous associations among the stocks (Kazemilari & Djauhari, 2015) and the stock prices are following the geometric Brownian motion (GBM) process. GBM can be expressed as the stock price follows a random walk pattern and is technically known as Markov process. With reference to the stock market fluctuations, Louis Bachelier was the first who introduced the mathematical theory of Brownian motion in 1900 (Bachelier, 1900; Taqqu, 2001). He used the increments of Brownian motion to model the absolute price change $|P(t)-P(t-\Delta t)|$ where $P(t)$ is the stock price at time t (Silva, 2005; Gan, 2016). Still, there is a limitation on his model as the stock price cannot become negative as suggested in the arithmetic Brownian process (Samuelson, 1965). Osborne (1959) was then introduced the logarithmic return of stock price to model the price change, $R(t)=\ln P(t)-\ln P(t-\Delta t)$. In 1965, Paul Samuelson suggested the GBM as a model for stocks as under the GBM, the logarithmic return of stock price is independent and identically normally distributed (Samuelson, 1965; Wilmott, 2007; Gan, 2016).

The movement or behavior of a particular stock can be tracked by the time series, such as univariate time series of stock price return. In the literature, many researchers applied univariate time series of stocks closing price to measure the similarities among the stocks which can be seen in the Korean stock market (Jung *et al.*, 2006), London stock market (Coelho *et al.*, 2007a), Malaysian stock market (Lam *et al.*, 2012; Bahaludin *et al.*, 2015), China stock market (Huang, Gao & Yang, 2009), Brazil stock maket (Tabak *et al*., 2010), global stock markets (Kumar & Deo, 2012), Turkish stock market (Kantar, Mustafa & Deviren, 2012), Indonesian stock market (Gan & Djauhari, 2012b), and NYSE stock market (Djauhari & Gan, 2015b). In these studies, the data collected are considered as synchronous and redundant data.

However, the real-world data is usually characterized by a high degree of non-stationary (Podobnik & Stanley, 2007) and carries plentiful of non-redundant economic information (Mantegna & Stanley, 2000; Bonanno, Vandewalle & Mantegna, 2000). The financial data is the non-synchronous nature of the data as the stock markets start operating at different time in different regions (Coelho *et al*., 2007b) as well as the highest and lowest stock prices do not occur at the same time for all stocks in a trading day (Kazemilari & Djauhari, 2015). In order to minimize the non-synchronous problem, Bonanno *et al.* (2000) and Coelho *et al.* (2007b) suggested to convert the daily stock price data to weekly stock price data.

There are numbers of methods to deal with the similarities among the stocks where each stock is represented as univariate time series of its closing price, such as Pearson correlation coefficient (PCC), detrended cross-correlation analysis (DXA), detrended cross-correlation analysis (DCCA) cross-correlation coefficient, σ_{DCCA} and dynamic time warping (DTW). Among these methods, PCC is the standard approach in measuring the similarities among the stocks in univariate setting due to PCC is being a simple and easy method to apply. However, the PCC is not a robust measurement and it can be misleading if outliers are presented (Zebende, 2011). Therefore, Podobnik & Stanley (2007) introduced detrended cross-correlation analysis (DXA) in order to quantify the cross-correlations between two univariate time series of stocks in the presence of non-stationarity. After that, a robust similarity measure is then proposed by Zebende (2011) in order to quantify the level of cross-correlation between

REFERENCES

Abdi, H. (2007). RV coefficient and congruence coefficient. *Encyclopedia of Measurement and Statistics*, 849-853.

Amadeo, K. (2017). The Balance. Retrived on August 15, 2017 from https://www.thebalance.com/how-do-stocks-and-stock-investing-affect-the-u-s-economy-3306179.

Asrah, M. A., Djauhari, M. A. & Suleiman, E. S. (2014). Work attitude among Malaysian academicians in the public universities: A social network analysis. *Modern Applied Science*, 8(5), 9-18.

Azer, I., Che Hamzah, H., Mohamad, S. A., & Abdullah, H. (2016). Contribution of Economic Sectors to Malaysian GDP. *Regional Conference on Science, Technology and Social Sciences (RCSTSS 2014)*, 17, 183-189.

Bachelier, L. (1900). *Theorie de la Speculation*. University of Paris: Ph.D. Thesis.

Bahaludin, H., Abdullah, M. H., & Mat Salleh, S. (2015). Minimal spanning tree for 100 companies in Bursa Malaysia. *AIP Conference Proceedings*, 1643, 609-615.

Bhattacharya, A. (2015). CNN Money. Retrieved on March 31, 2018 from http://money.cnn.com/2015/06/03/news/economy/oecd-world-us-economy-slower-growth/index.html.

Bonnano, G., Vandewalle, N., & Mantegna, R. N. (2000). Taxonomy of stock market indices. *Physical Review E*, 62, R7615-R7618.

Bonacich, P. (1972). Factoring and weighting approaches to status scores and clique identification. *Journal of Mathematical Sociology*, 2, 113-120.

Borgatti, S. P. (2005). Centrality and network flow. *Social Networks*, 27, 55-71.

Brida, J. G., & Risso, W. A. (2008). Multidimensional minimal spanning tree: The Dow Jones case. *Physica A: Statistical Mechanics and its Application*, 387(21), 5205-5210.

Bursa Malaysia. (2017). Retrieved on July 15, 2017 from http://www.bursamalaysia.com/market/listed-companies/list-of-companies/plc-profile.html?stock_code=9148

Cheng, T. L. (2017). The Star Online. Retrieved on March 22, 2018 from https://www.thestar.com.my/business/business-news/2017/12/30/taking-stock-of-property-sector/

Coelho, R., Hutzler, S., Repetowicz, P., & Richmond, P. (2007a). Sector analysis for a FTSE portfolio of stocks. *Physica A*, 373, 615-626.

Coelho, R., Gilmore, C. G., Lucey, B., Richmond, P., & Hutzler, S. (2007b). The evolution of interdependence in world equity markets – Evidence from minimum spanning trees. *Physica A*, 376, 455-466.

Djauhari, M. A. (2012). A robust filter in stock network analysis. *Physica A: Statistical Mechanics and its Applications*, 391(20), 5049-5057.

Djauhari, M. A., & Gan, S. L. (2013a). Network topological property of English dialects similarity: A robust filter approach. *American Journal of Applied Sciences*, 10(7), 646-653.

Djauhari, M. A., & Gan, S. L. (2013b). Minimal spanning tree problem in stock networks analysis: An efficient algorithm. *Physica A: Statistical Mechanics and Its Applications*, 392(9), 2226–2234.

Djauhari, M. A., & Gan, S. L. (2015a). Bursa Malaysia stocks market analysis: a review. *ASM Science Journal*, 8(2), 150-158.

Djauhari, M. A., & Gan, S. L. (2015b). Optimality problem of network topology in stocks market analysis. *Physica A: Statistical Mechanics and its Application*, 419, 108-114.

Djauhari, M. A., & Gan, S. L. (2016). Network topology of economic sectors. *Journal of Statistical Mechanics: Theory and Experiment*, 2016, 1-20.

Escoufier, Y. (1973). Le traitement des variables vectorielles. *Biometrics*, 29(4), 751-760.

Escoufier, Y. (2006). Operator related to a data matrix: a survey. *Compstat 2006 - Proceedings in Computational Statistics*. Rome, Italy: Physica-Verlag HD. pp. 285-297.

Fallani, F. D. V., Costa, L. da F., Rodriguez, F. A., Astolfi, L., Vecchiato, G., Toppi, J., Borghini, G., Cincotti, F., Mattia, D., Salinari, S., Isabella, R., & Babiloni, F. (2010). A graph-theoretical approach in brain functional networks. Possible implications in EEG studies. *Nonlinear Biomed Physics*, 4(1), S8. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2880805/

Financial Time (2018). Indices. Retrieved on March 21, 2018 from https://markets.ft.com/data/indices/tearsheet/summary?s=KPRP:KLS

Forouzan, B. A. (2012). *Data Communications and Networking* – Fourth Edition. New York, NY: McGraw Hill.

Freeman, L. C. (1979). A set of measures of centrality based on betweenness. *Sociometry*, 35-41.

Gan, S. L. (2016). *A Similarity Measure of Multivariate Time Series in Stocks Network Analysis*. Universiti Teknologi Malaysia: Ph.D. Thesis.

Gan, S. L., & Djauhari, M. A. (2012a). Stock networks analysis in Kuala Lumpur Stock Exchange. *Malaysian Journal of Fundamental & Applied Sciences*, 8(2), 60-66.

Gan, S. L., & Djauhari, M. A. (2012b). Network topology of Indonesian stock market. *Proceedings of IEEE International Conference on Cloud Computing and Social Networking*. Bandung, Indonesia: IEEE. pp. 1-4.

Gan, S. L., & Djauhari, M. A. (2013). Multidimensional stock network analysis: an Escoufier's RV coefficient approach. *Proceedings ICMSS 2013 International Conference on Mathematical Sciences and Statistics*. Kuala Lumpur, Malaysia: AIP Conference Proceedings. pp. 550-555.

Gan, S. L., & Djauhari, M. A. (2015). New York Stock Exchange performance: evidence from the forest of multidimensional minimum spanning trees. *Journal of Statistical Mechanics: Theory and Experiment*, 2015, 1-20.

Gilmore, C. G., Lucey, B. M., & Boscia, M. (2008). An ever-closer union? Examining the evolution of linkages of European equity markets via minimum spanning trees. *Physica A: Statistical Mechanics and its Application*, 387(25), 6319-6329.

Hair, J. F., Black, W. C., Babin, B. J. & Anderson, R. A. (2010). *Multivariate Data Analysis* – 7th Edition. London: Pearson Prentice Hall.

Halinen, A., & Tornroos, J. (1998). The role of embeddedness in the evolution of business networks. *Scandinavian Journal of Management*, 14(3), 187-205.

Hotelling, H. (1936). Relations between two sets of variates. *Biometrika*, 28(3/4), 321-377.

Huang, F., Gao, P., & Wang, Y. (2009). Comparison of Prim and Kruskal on Shanghai and Shenzhen 300 index hierarchical structure tree. *International Conference on Web Information Systems and Mining*. Shanghai, China: IEEE. pp. 237-241.

Husson, F., Josse, J., Le, S., Mazet, J. (2017). FactoMineR Version 1.39: Multivariate Exploratory Data Analysis and Data Mining with R. Retrieved on December 1, 2017 from https://cran.r-project.org/web/packages/FactoMineR/index.html.

Husson, F., Josse, J., Le, S., Mazet, J. (2018). FactoMineR Version 1.40: Multivariate Exploratory Data Analysis and Data Mining with R. Retrieved on March 28, 2018 from https://cran.r-project.org/web/packages/FactoMineR/index.html.

Investopedia. (2017). Retrieved on January 23, 2018 from https://www.investopedia. com/terms/t/timeseries.asp

Jolliffe, I. T. (1986). Principal component analysis and factor analysis. In *Principal component analysis*. NY: Springer.

Josse, J., Pages, J., & Husson, F. (2008). Testing the significance of the RV coefficient. *Computational Statistics and Data Analysis*, 53, 82-91.

Jung, W. S., Chae, S., Yang, J. S., & Moon, H. T. (2006). Characteristics of the Korean stock market correlations. *Physica A: Statistical Mechanics and its Application*, 361(1), 263-271.

Lam, S. X., Zainuddin, H., Shitan, M., & Krishnarajah, I. (2012). The market effect on Malaysian stock correlation network. *Applied Mathematical Sciences*, 6(104), 5161-5178.

Kantar, E., Keskin, M. & Deviren, B. (2012). Analysis of the effects of the global financial crisis on the Turkish economy, using hierarchical methods. *Physica A: Statistical Mechanics and its Applications*, 391, 2342-2352.

Kazemilari, M., & Djauhari, M. A. (2013). Analysis of a correlation network in world currency exchange market. *International Journal of Applied Mathematics and Statistics*, 44(14), 202-209.

Kazemilari, M., & Djauhari, M. A. (2015). Correlation network analysis for multi-dimensional data in stocks market. *Physica A: Statistical Mechanics and its Applications*, 429, 108-114.

Kollo, T. & von Rosen, D. (2005). *Advanced Multivariate Statistics with Matrices*. Netherlands: Springer.

Krichel, T., & Bakkalbasi, N. (2006). A social network analysis of research collaboration in the economics community. *Journal of Information Management and Scientometrics*, 3(1), 12.

Kumar, S., & Deo, N. (2012). Correlation and Network Analysis of Global Financial Indices. *Physical Review E*, 86, 026101-1-8.

Mantegna, R. N. (1999). Hierarchical structure in financial markets. *The European Physcial Journal B*, 11, 193-197.

Mantegna, R. N., & Stanley, H. E. (2000). An introduction to econophysics: correlations and complexity in finance. *Physics Today*, 53(12), 70.

Mizuno, T., Takayasu, H., & Takayasu, M. (2006). Correlation networks among currencies. *Physica A: Statistical Mechanics and its Application*, 364, 336-342.

Musmeci, N., Aste, T., & Di Matteo, T. (2015). Relation between financial market structure and the economy: Comparison between clustering methods. *PLOS ONE*, 10(4): e0126998.

National Property Information Centre (2018). Retrieved on March 21, 2018 from http://napic.jpph.gov.my/portal/web/guest/mainpage?p_p_id=ViewStatistics_WAR_ViewStatisticsportlet&p_p_lifecycle=2&p_p_state=normal&p_p_mode=view&p_p_resource_id=fileDownload&p_p_cacheability=cacheLevelPage&p_p_col_id=column-2&p_p_col_count=2&fileURI=5316.

Naylor, M. J., Rose, L. C., & Moyle, B. J. (2007). Topology of foreign exchange markets using hierarchical structure methods. *Physica A: Statistical Mechanics and its Application*, 382(1), 199-208.

Osborne, M. F. M. (1959). Brownian motion in the stock market. *Operations Research*, 7, 145-173.

Park, K., & Yilmaz, A. (2010). A social network analysis approach to analyze road networks. *ASPRS Annual Conference*, San Diego, CA.

Podobnik, B., & Stanley, E. (2007). Detrended cross-correlation analysis: A new method for analyzing two non-stationary time series. *Physical Review Letters*, 100(8), 084102.

Pro Programming. (2017). Retrieved on March 24, 2018 from https://proprogramming.org/find-time-complexity-of-algorithms/.

Resovsky, M., Horvath, D., Gazda, V., & Sinicakova, M. (2013). Minimum spanning tree application in the currency market. *Biatec*, 21(7), 21-23.

Robert, P., & Escoufier, Y. (1976). A unifying tool for linear multivariate statistical methods: the RV-coefficient. *Journal of the Royal Statistical Society Series C*, 25(3), 257-265.

Samuelson, P. A. (1965). Rational theory of warrant pricing. *Industrial Management Review*, 6(2), 13-31.

Shabanali, R. (2017). *Webmindset*. Definition: What is a complex system? Retrieved on July 18, 2017 from http://webmindset.net/terminology-of-the-complexity-approach/

Shaffer, C. A. (2012). *Data Structures and Algorithm Analysis in Java*. USA: Dover Publication.

Sharif, S., & Djauhari, M. A. (2012a). A proposed centrality measure: The case of stocks traded at Bursa Malaysia. *Modern Applied Science*, 6(10), 62-69.

Sharif, S., & Djauhari, M. A. (2012b). Worker's behaviour in manufacturing industry: An evidence from a minimum spanning tree. *Malaysian Journal of Fundamental & Applied Sciences*, 8(1), 1-6.

Sharif, S., Yusoff, N. S., & Djauhari, M. A. (2012). Network topology of foreign exchange rate. *Modern Applied Science*, 6(11), 35–43.

Sharif, S., & Djauhari, M. A. Minimal Spanning Tree and Centrality Measure. In: Sharif, S., Yahaya, S. S. S., & Saaban, A. (Ed). *Scientific Investigation on Univariate Quantitative Methods*. Sintok: UUM Press. Pp. 58-67; 2016.

Shikar, M. A. K. (2012). Multivariate statistical analysis. *British Journal of Science*, 6(1), 55-66.

Silva, A. C. (2005). *Applications of Physics to Finance and Economics: Returns, Trading Activity and Income*. University of Maryland: Ph.D. Thesis.

Situngkir, H., & Surya, Y. (2005). *On stock market dynamics through ultrametricity of minimum spanning tree*. Indonesia: Bandung de Institute. pp. 1-18.

Sieczka, P., & Holyst, J. A. (2009). Corelations in commodity markets. *Physica A: Statistical Mechanics and its Applications*, 388, 1621-1630.

Tabachnick, B. G. & Fidell, L. S. (1989). *Using Multivariate Statistics*. New York: Dover Publication.

Tabak, B. M., Serra, T. R., & Cajueiro, D. O. (2010). Topological properties of stock market networks: The case of Brazil. *Physica A: Statistical Mechanics and its Application*, 389(16), 3240-3249.

Taqqu, M. S. (2001). Bachelier and his times: A conversation with Bernard Bru. *Finance and Stochastic*, 5(1), 3-32.

The Wall Street Journal. (2018). Retrieved on 28th February, 2018 from http://quotes.wsj.com/MY/XKLS/MKLAND/company-people.

Toh, K. I. (2017). *The success story of IGB REIT*. Retrieved on 28th February, 2018 from https://www.thestar.com.my/business/business-news/2017/09/04/the-success-story-of-igb-reit/.

Tumminello, M., Aste, T., Di Matteo, T., & Mantegna, R. N. (2005). A tool for filtering information in complex systems. *Applied Physical Sciences*, 102, 10421-10426.

Tumminello, M., Coronnello, C., Lillo, F., Micciche, S., Mantegna, R. N. (2007). Spanning trees and bootstrap reliability estimation in correlation based networks. *International Journal Bifurcation Chaos*, 17, 2319.

Vandewalle, N., Brisbois, F., & Tordoir, X. (2001). Non-random Topology of Stock Markets. *Quantitative Finance*, 1, 372-374.

Wang, G. J., Xie, C., Chen, Y. J., & Chen, S. (2013). Statistical properties of the foreign exchange network at different time scales: Evidence from Detrended Cross-Correlation Coefficient and minimum spanning tree. *Entropy*, 15(5), 1643–1662.

Wang, J., Zhu, Y., Li, S., Wan, D., & Zhang, P. (2014). Multivariate time series similarity searching. *The Scientific World Journal*, 2014, 1-8.

Wilmott, P. (2007). *Paul Wilmott Introduces Quantitative Finance*. New York: John Wiley & Sons, Inc.

Yahoo Finance (2017). Retrieved on July 16, 2017 from https://finance.yahoo.com/

Yamashita, Y., & Yadohisa, H. (2012). Similarity measure and clustering algorithm for candlestick valued data. *Joint Meeting of Japanese and Italian Classification Societies (JCS-CLADAG)*. Capri Island, Italy. pp. 17-20.

Yang, K. & Shahabi, C. (2004). A PCA-based similarity measure for multivariate time series. *Proceedings of the 2nd ACM International Workshop on Multimedia Databases*. Washington, USA: Association for Computing Machinery. pp. 65-74.

Yeoh, B. K., Hooy, C. W., & Arsad, Z. (2010). Time-varying world integration of the Malaysian stock market: a Kalman filter approach. *Asian Academy of Management Journal of Accounting and Finance*, 6(2), 1-17.

Yeoh, K. H., & Hooy, C. W. (2011). A DEA analysis on the efficiency of Malaysian Commercial Banks over 2003-2008. *The Journal of the Institute of Bankers Malaysia*, 136, 3-11.

Zebende, G. F. (2011). DCCA cross-correlation coefficient: Quantifying level of cross-correlation. *Physica A*, 390, 614-618.

Zhang, J., & Luo, Y. (2017). Degree centrality, betweenness centrality, and closeness centrality in social network. *Advances in Intelligent Systems Research*, 132, 300-303