

A SIMILARITY MEASURE OF MULTIVARIATE TIME SERIES IN STOCKS  
NETWORK ANALYSIS

GAN SIEW LEE

A thesis submitted in fulfilment of the  
requirements for the award of the degree of  
Doctor of Philosophy (Mathematics)

Faculty of Science  
Universiti Teknologi Malaysia

MARCH 2016

To my beloved family, for your love and support.  
To my friends, for your wits, intelligence and guidance in life.

## ACKNOWLEDGEMENT

I would like to express my sincere gratitude to everyone who supported and helped me during the preparation of this thesis. In particular, Prof. Dr. Maman A. Djauhari deserves my thanks, for the interesting topic that he introduced to me and his continuous support through these years. He was my main supervisor from the beginning 15 February 2012 until 17 April 2014 when he was contract Professor at Universiti Teknologi Malaysia (UTM). After he joined Universitas Pasundan, Bandung, Indonesia, from May 2014 – September 2014 and then joins Universiti Putra Malaysia (UPM) as Follow Researcher at Institute for Mathematical Research (INSPEM) from 8 September 2014 until present, he is my co-supervisor. I am thankful for his constructive comments, suggestions, guidance, and constant encouragement. I also would like to convey my thanks to Prof. Dr. Zuhaimy bin Haji Ismail. He was my co-supervisor from 15 February 2012 to 17 April 2014 and switched as main supervisor from 21 May 2014 until present. I greatly appreciate his advice, moral support and helpful suggestions. The comments and suggestions from both of my guides are essential to the final presentation of this research. I am thankful to UTM, for providing me all the facilities and documentations. I am also very thankful to the librarians for their professional assistance in providing the valuable literature and also the administrative staffs for their effective services. I am thankful to the Ministry of Higher Education Malaysia and Malaysian Government for sponsorship of MyBrain15 MyPhD. And also thank you to INSPEM, for providing the facilities during meeting Prof Maman at UPM. Last but not least, zillions of thanks to my dear family members and friends for their encouragement and moral support during my studies. Especially, to my mother, Tan Eng, for her endless love and encouragement and also to my late father, Gan Kim Hoo who has been my source of inspiration.

## ABSTRACT

Correlation-based network as a model for financial markets, especially stock market, is a complex system has received much attention. There have been a lot of studies which deals with stocks network analysis, where each stock is represented by a univariate time series of its closing price, and then the similarity between two stocks are quantified by using Pearson correlation coefficient (PCC) on the logarithmic returns. However, in daily stock market activity, stock is represented by a multivariate time series during its opening, highest, lowest, and closing prices. The solely used of the information from closing price may cause the loss of information from other prices. In this thesis, all four prices are considered. The notion of multivariate time series similarity among stocks are developed. The use of Escoufier vector correlation (EVC), a multivariate generalization of PCC, is proposed to measure the similarity between stocks. Then the EVC coefficients are used to construct the stocks network in multivariate setting based on minimal spanning tree (MST). In the case study on BURSA MALAYSIA, the topological properties of stocks in EVC-based MST and in PCC-based MST are different. The total path lengths among stocks in the economic sector according to EVC-based MST is generally smaller than according to PCC-based MST. It means that with the approach of EVC-based MST, the stocks are strongly connected with other stocks in the same sector. Moreover, EVC is proposed to define the similarity between economic sectors, where each sector is represented by a multivariate time series of  $p$  components and each component is a univariate time series of stock's closing price. To the best of our knowledge, there is no previous studies which deals with the similarity between economic sectors using this approach. The methodology for economic sectors network analysis is formulated in this thesis. The current practice of using Kruskal's or Prim's algorithm is to obtain MST, and then sub-dominant ultrametric (SDU) from the MST. It will consume a lot of time when the number of stocks is large. Therefore to solve this problem, an efficient algorithm is developed based on fuzzy relation approach. A comparison study based on the empirical and simulated data shows that the proposed algorithm is faster. The proposed algorithm provides not only MST and SDU, but also the forest of all MSTs.

## ABSTRAK

Jaringan berdasarkan korelasi sebagai sebuah model pasaran kewangan, khususnya pasaran saham, ialah suatu sistem yang kompleks telah mendapat perhatian. Banyak kajian dijalankan berkenaan analisis jaringan saham, dengan setiap saham diwakili oleh siri masa univariat bagi harga penutupnya, dan seterusnya persamaan antara dua saham dikuantifikasikan dengan menggunakan pekali korelasi Pearson (PCC) pada pulangan logaritma tersebut. Walau bagaimanapun, dalam aktiviti pasaran saham harian, saham diwakili oleh siri masa multivariat semasa harga pembukaan, tertinggi, terendah, dan penutupnya. Dengan menggunakan maklumat daripada harga penutup sahaja boleh menyebabkan kehilangan maklumat daripada harga-harga lain. Dalam tesis ini, kesemua empat harga dipertimbangkan. Persamaan siri masa multivariat dalam kalangan unit saham dibangunkan. Penggunaan korelasi vektor Escoufier (EVC), iaitu satu generalisasi multivariat bagi PCC, diusulkan untuk mengukur persamaan antara saham. Seterusnya, korelasi EVC digunakan untuk membina jaringan saham dalam latar multivariat berdasarkan pokok perentangan minimum (MST). Dalam kajian kes terhadap BURSA MALAYSIA, sifat topologi saham dalam MST berdasarkan EVC dan MST berdasarkan PCC adalah berbeza. Jumlah panjang laluan kalangan saham dalam sektor ekonomi menurut MST berdasarkan EVC pada umumnya lebih kecil daripada MST berdasarkan PCC. Ini bermakna dengan pendekatan MST berdasarkan EVC, sahamnya amat kuat terhubung dengan saham-saham lain dalam sektor yang sama. Selain itu, EVC juga diusulkan untuk mendefinisikan persamaan antara sektor ekonomi, dengan setiap sektor diwakili oleh siri masa multivariat bagi komponen  $p$  dan setiap komponen ialah suatu siri masa univariat bagi harga penutup saham tersebut. Sepanjang pengetahuan kami, tiada kajian terdahulu yang memberi fokus kepada persamaan antara sektor ekonomi menggunakan pendekatan ini. Metodologi analisis jaringan bagi sektor ekonomi diformulasikan di dalam tesis ini. Amalan terkini yang menggunakan algoritma Kruskal atau Prim adalah perlu untuk memperoleh MST, dan kemudian ultrametrik subdominan (SDU) daripada MST. Ini akan memakan masa apabila bilangan saham tersebut besar. Oleh itu untuk menyelesaikan masalah ini, satu algoritma yang cekap dibangunkan berdasarkan pendekatan hubungan kabur. Kajian perbandingan berdasarkan data empirikal dan simulasi menunjukkan bahawa algoritma yang diusulkan lebih laju. Algoritma yang diusulkan menyediakan bukan hanya MST dan SDU, tetapi juga *forest* bagi semua MST.

## TABLE OF CONTENTS

CHAPTER	TITLE	PAGE
	<b>DECLARATION</b>	ii
	<b>DEDICATION</b>	iii
	<b>ACKNOWLEDGEMENTS</b>	iv
	<b>ABSTRACT</b>	v
	<b>ABSTRAK</b>	vi
	<b>TABLE OF CONTENTS</b>	vii
	<b>LIST OF TABLES</b>	x
	<b>LIST OF FIGURES</b>	xi
	<b>LIST OF APPENDICES</b>	xiii
<b>1</b>	<b>INTRODUCTION</b>	<b>1</b>
	1.1 Background of the Problem	1
	1.2 Statement of the Problem	3
	1.3 Objectives of the Study	5
	1.4 Scope of the Study	5
	1.5 Organization of the Thesis	7
	1.6 Research Framework	8
<b>2</b>	<b>LITERATURE REVIEW</b>	<b>10</b>
	2.1 Similarity among Univariate Time Series	10
	2.2 Similarity among Multivariate Time Series	13
	2.3 Correlations in Stock Market	14
	2.4 Network Topological Analysis	20
	2.4.1 Centrality Measures	21

2.5	Summary	23
<b>3</b>	<b>RESEARCH METHODOLOGY</b>	<b>28</b>
3.1	Stocks Network Construction: A Standard Practice	28
3.2	Multidimensional Prices-based Stocks Network Construction	34
3.2.1	Characterization of Random Vector	35
3.2.2	Correlation between Two Operators	37
3.2.3	Similarity among Stocks	40
3.3	Economic Sectors Network Construction	42
3.4	An Efficient Algorithm	43
3.4.1	Construction of Forest of All MSTs	44
3.4.2	Finding an MST in $D$	46
3.5	Summary	50
<b>4</b>	<b>CASE STUDIES OF EVC-BASED MST: BURSA MALAYSIA</b>	<b>51</b>
4.1	Stocks Network Analysis of BURSA MALAYSIA	51
4.1.1	Topological Structure of Stocks	52
4.1.2	Centrality Measures of PCC-Based MST	56
4.1.3	Centrality Measures of EVC-Based MST	61
4.1.4	Jaccard Index	66
4.1.5	Summary	67
4.2	Economic Sectors Network Analysis of BURSA MALAYSIA	69
4.2.1	Centrality Measures	70
4.2.2	Topological Structure of Inter and Intra Sectors	71
4.2.3	Summary	73
<b>5</b>	<b>CASE STUDIES OF EVC-BASED MST: NYSE</b>	<b>75</b>
5.1	Stocks Network Analysis of NYSE	75
5.1.1	Topological Structure of Stocks	76
5.1.2	Centrality Measures of EVC-Based MST	78

	5.1.3 Summary	83
5.2	SICs Network Analysis of NYSE	84
	5.2.1 Centrality Measures	88
	5.2.2 Topological Structure of Inter and Intra SICs	91
	5.2.3 Summary	93
<b>6</b>	<b>AN EFFICIENT ALGORITHM: EMPIRICAL AND SIMULATION CASES</b>	<b>94</b>
6.1	An Illustrative Example	94
6.2	Performance of Algorithm	96
	6.2.1 Performance of Algorithms in Empirical Cases	96
	6.2.2 Performance of Algorithms in Simulation Cases	99
6.3	Summary	101
<b>7</b>	<b>CONCLUSION</b>	<b>102</b>
7.1	Conclusions	102
7.2	Contributions of the Study	104
7.3	Future Research	105
	<b>REFERENCES</b>	<b>106</b>
	Appendices A – G	118 - 201



## LIST OF TABLES

<b>TABLE NO.</b>	<b>TITLE</b>	<b>PAGE</b>
2.1	The summary of previous studies in financial markets	24
4.1	Total path lengths among stocks in each sector	55
4.2	Topological Properties of MSTs	56
4.3	Centrality scores for 50 stocks in PCC-based MST	56
4.4	Centrality scores for 50 stocks in EVC-based MST	61
4.5	P-values for GAMUDA and other stocks	68
4.6	Color of nodes and number of stocks in each economic sector of BURSA MALAYSIA	69
4.7	Centrality scores for 8 economic sectors of BURSA MALAYSIA	70
5.1	Top-10 stocks on four centrality measures	78
5.2	List of 25 most influential stocks at NYSE	83
5.3	Industry sector, color, number of SICs and number of stocks of NYSE	84
5.4	Industry sector, SIC code, industry name and number of stocks of NYSE	85
5.5	Centrality scores for 63 SICs of NYSE	88
6.1	Running time (in second) of each algorithm in four stocks markets	97
6.2	Running time (in second) of each algorithm in four simulated markets	99
7.1	Hypothetical distance matrix	105

## LIST OF FIGURES

<b>FIGURE NO.</b>	<b>TITLE</b>	<b>PAGE</b>
1.1	Summary of research framework	9
3.1	Algorithm of forest of all MSTs	46
3.2	Flow chart of the proposed algorithm	49
4.1	PCC-based MST of 50 stocks traded at BURSA MALAYSIA	53
4.2	EVC-based MST of 50 stocks traded at BURSA MALAYSIA	54
4.3	Degree centrality of PCC-based MST	58
4.4	Betweenness centrality of PCC-based MST	58
4.5	Closeness centrality of PCC-based MST	59
4.6	Eigenvector centrality of PCC-based MST	59
4.7	Overall centrality of PCC-based MST	60
4.8	Degree centrality of EVC-based MST	63
4.9	Betweenness centrality of EVC-based MST	63
4.10	Closeness centrality of EVC-based MST	64
4.11	Eigenvector centrality of EVC-based MST	64
4.12	Overall centrality of EVC-based MST	65
4.13	EVC-based MST of 8 economic sectors of BURSA MALAYSIA	70
4.14	Topological structure of inter and intra sectors of BURSA MALAYSIA	72
5.1	EVC-based MST of 1515 stocks traded at NYSE	77
5.2	Top-10 stocks on degree centrality	79
5.3	Top-10 stocks on betweenness centrality	80
5.4	Top-10 stocks on closeness centrality	81
5.5	Top-10 stocks on eigenvector centrality	82

5.6	EVC-based MST of 63 SICs of NYSE	87
5.7	Topological structure of inter and intra SICs of NYSE	91
6.1(a)	Forest of all MSTs $F$	95
6.1(b)	Sub-graph $H(1)$	95
6.1(c)	Sub-graph $H^*$	95
6.1(d)	MST $M$ in $H^*$	95
6.2	MST in $\mathcal{D}$	96
6.3	Comparison between KA and PAK in stock markets	98
6.4	Comparison between PA and PAP in stock markets	99
6.5	Comparison between KA and PAK in simulated markets	100
6.6	Comparison between PA and PAP in simulated markets	101
7.1	Two MSTs obtained from Kruskal's algorithm	105

**LIST OF APPENDICES**

<b>APPENDIX</b>	<b>TITLE</b>	<b>PAGE</b>
A	Publications	118
B	The relation between Escoufier operator and basis theorem in linear algebra	121
C	Asymptotic distribution of EVC	123
D	50 capitalized stocks traded at BURSA MALAYSIA	129
E	Definition of sector classified at BURSA MALAYSIA	131
F	1515 stocks traded at NYSE	132
G	Centrality scores for 1515 NYSE stocks in EVC-based MST	168

## **CHAPTER 1**

### **INTRODUCTION**

The aim of this chapter is to introduce the significance of this research. In Section 1.1, the background of the problem is discussed followed by the problem statement in Section 1.2. After that, the objectives of the study are presented in Section 1.3. In Section 1.4, the scope of the study is presented. The organization of the thesis is delivered in Section 1.5. Finally, the research framework in Section 1.6 closed this chapter.

#### **1.1 Background of the Problem**

The existence of interrelationships among stocks is a well-known fact. This phenomenon as can be seen from the stock price of a company is not only influenced by company fundamental itself but also the news released from the other listed companies. Meanwhile, the economic factors presented in financial markets drive the prices of several stocks. Stocks are highly transmitting behavior or information with the stocks in the same economic sector. Understanding the interrelationships among stocks and assessing the relative importance of stocks are essential in asset allocation, portfolio selection and risk management. The interrelationships can be expressed by measuring the similarities among the stocks.

The similarity between stocks is customarily quantified by using Pearson correlation coefficient (PCC) on the time series of logarithmic closing price returns. For studying the correlations of stocks, the analysis based on network model has been proposed by Mantegna (1999). The correlation structure among stocks summarized in a symmetric matrix is conceptualized as a network. A network is defined by a set of nodes connected by a set of edges; nodes represent the stocks and edges represent the interrelationships among stocks. The network is usually large with  $n$  stocks are connected by  $n(n-1)/2$  edges. Since the stocks network is complex and difficult to analyze, Mantegna (1999) introduced a filtering process based on minimal spanning tree (MST) to extract the most important information contained therein and provides a simple representation of the network.

The MST method provides a simple way to study the correlations of stocks. The topological structure of stocks generated by MST is useful for studying the topological properties such as the relative importance of each stock in the investigated portfolio. It is useful for recognizing the specific characteristic of stocks with their relative positions in the network. The previous studies have shown that the individual stocks tend to gather with the stocks in the same economic sector. In the current practice of stocks network analysis, sub-dominant ultrametric (SDU) is obtained from MST and then presented in the form of an indexed hierarchical tree (HT). MST is used to study the topological properties of stocks and the corresponding SDU to understand the taxonomic structure of stocks. The important roles of MST and SDU can be found in many studies of stock market, currency market, commodity market, and automotive industry. The studies showed that the MST provides meaningful economic explanation. Most of previous studies of stocks network analysis are focused on the United States (US) market, such as New York Stock Exchange (NYSE).

According to the business activities of companies in the market, stocks are classified into specific economic sectors. Therefore, each economic sector is composed by a number of homogenous stocks. Generally, the numbers of stocks in two different economic sectors are not necessary the same. Understanding the topological properties of economic sectors is as important as understanding the

topological properties of stocks. In order to construct the economic sectors network, the similarities between economic sectors are quantified. A few studies work on the economic sectors network analysis. In the existing studies, the similarity between economic sectors is defined as PCC between the logarithmic returns of economic sector indices and then with the analysis based on MST.

In current practice, there are four steps for conducting stocks network analysis. The first step is to define the similarities between stocks considering their corresponding time series of prices. In the existing studies, each stock is typically represented by a univariate time series of its closing price. Then, the similarity between stocks is quantified by PCC between time series of logarithmic closing price returns. The second step is to transform the correlation coefficient matrix into distance matrix. The third step is to analyze the corresponding network based on MST method and then followed by constructing the SDU from MST in the last step.

In terms of computation, in the literature, there are two algorithms commonly applied to construct the MST, namely Kruskal's algorithm and Prim's algorithm. These algorithms are directly applied on the distance matrix to obtain the MST and then the SDU from the MST. The Kruskal's algorithm is mathematically appealing and simple. Based on the frequency of algorithms applied in the literature, it is clear that researchers prefer to use Kruskal's algorithm than Prim's algorithm. This maybe because of the former is easier to formulate than the latter. The search for efficient ways to construct the MST and SDU is one of the most important problems in the modern financial industry. This study proposes an efficient algorithm.

## **1.2 Statement of the Problem**

In the existing studies, PCC is customarily used to calculate the similarity between stocks. The current stocks network analysis is based on the PCCs between the logarithmic closing price returns. The use of only information from closing price in the analysis might cause losing of information from other variables in the system.

This is because, in daily stock market activity, the price information for each stock is recorded in opening, highest, lowest, and closing prices. Thus, stock is a multivariate time series of those four prices and not a univariate time series of closing price only. The use of those four prices provides more information than the information from closing price only. It could help to understand the market better. In order to have a comprehensive stocks network analysis, we have to take into account the information of all four prices. To the best of our knowledge, so far, there is no study attempt to construct such stocks network.

Although there are a lot of studies in stocks network analysis, a few studies in the economic sectors network analysis have been conducted. As the importance of analyzing the topological properties of stocks, the topological properties of economic sectors are crucial for investment activities. If each stock is represented by a univariate time series of its closing price, then each economic sector can be considered as a multivariate time series of several closing prices each of which represents a stock in that sector. In practice, the numbers of stocks in two different economic sectors are not necessary the same. Based on this idea, the similarity between economic sectors is the similarity between two multivariate time series which might be of different dimensions. To the best of our knowledge, there is no study dealing with the notion of economic sectors network in multivariate setting.

In order to analyze the topological properties of stocks, the MST issued from Kruskal's algorithm or Prim's algorithm is used. The algorithm is directly applied on the distance matrix to obtain the MST. Therefore, the size of the distance matrix or, equivalently, the size of the correlation matrix influences the performance of the algorithm in constructing the MST. The performance of the algorithm becomes slower and slower when the number of stocks gets larger and larger. This is one of the fundamental problems in stocks network analysis as the number of stocks that involved in the analysis is always large. In this thesis, a faster algorithm is presented.



### 1.3 Objectives of the Study

The objectives of this research are:

- i. To define the similarity between multivariate time series of stocks' opening, highest, lowest, and closing prices by using Escoufier's operator.
- ii. To develop a filtering process for stocks network analysis in multivariate setting.
- iii. To formulate a methodology for economic sectors network analysis.
- iv. To develop an efficient algorithm having faster computational running time than Kruskal's algorithm and Prim's algorithm.

### 1.4 Scope of the Study

The scope of this research can be divided into three aspects.

1. Theoretical aspect. This aspect covers:
  - i. The use of Escoufier's operator to define Escoufier vector correlation (EVC) that can be used to quantify the similarity between two multivariate time series of different dimensions. We propose the use of EVC to measure the similarity between stocks, each of which is represented by a multivariate time series of its opening, highest, lowest, and closing prices. Then, construct stocks network in multivariate setting based on MST. We also introduce the use of EVC in economic sectors network analysis when each stock is represented by a univariate time series of its closing price.
  - ii. Prove that EVC is a multivariate generalization of PCC.
  - iii. Develop an efficient algorithm based on the combination of fuzzy relation theory and the properties of forest of all MSTs.

2. Computational aspect.

Empirical and simulated networks are conducted to show that, in terms of the running time, the proposed algorithm has higher computational efficiency than Kruskal's algorithm and Prim's algorithm.

3. Practical aspect.

It covers the applications in real stock markets to show the advantages of the methods developed in points 1 and 2 for (i) stocks network analysis, and (ii) economic sectors network analysis.

To illustrate the advantages of the proposed stocks network analysis and economic sectors network analysis, we perform empirical studies for Malaysian and US stock markets. The data from BURSA MALAYSIA and NYSE are used in this study. For BURSA MALAYSIA case study, we analyze the data of 50 most capitalized stocks during the time period from January 2, 2007 until September 26, 2011. The data of opening, highest, lowest, and closing prices were downloaded from <http://www.klse.info/downloads>. For NYSE case study, we analyze the data of 1515 stocks continuously traded at NYSE from December 16, 2004 until November 21, 2014. The data of four prices were downloaded from <http://www.nasdaq.com/>.

To illustrate the advantages of the proposed algorithm in finding an MST and corresponding SDU, a comparison study based on empirical data from stock markets and simulated data are presented in this study. The considered stock markets are Singapore Stock Exchange (SGX), BURSA MALAYSIA, NYSE, and Shanghai and Shenzhen Stock Exchange (SHSZ). The number of stocks in SGX, BURSA MALAYSIA, NYSE and SHSZ are 28, 90, 100, and 300, respectively. The data were collected during the period of January 2, 2008 until December 31, 2010, of January 2, 2007 until December 31, 2009, of November 18, 2010 until March 28, 2012, and of July 28, 2011 until June 1, 2012, respectively. The data were downloaded from Yahoo Finance.

## 1.5 Organization of the Thesis

This thesis is organized into seven chapters. Chapter 1 is the Introduction. The research background, problem statement, objectives, and scopes of this research are given in this chapter.

In Chapter 2 the literature review about the similarity measures on univariate and multivariate time series are presented. The importance of correlation-based network analysis in stock market and the centrality measures used in network analysis are presented.

The methodologies of this research are given in Chapter 3 which is divided into four sections. In the first section the standard practice of stocks network construction proposed by Mantegna (1999) is recalled. The second section is about defining the similarity between stocks in multivariate setting by using Escoufier's operator and then with the stocks network construction. Based on the EVC defined from Escoufier operator, the construction of economic sectors network in multivariate setting is developed in the third section. In the last section, an efficient algorithm is developed based on the properties of SDU from fuzzy relation viewpoint, uniqueness theorem and properties of forest of all MSTs in Djauhari (2012).

In Chapter 4, as a case study, the topological properties of 50 stocks traded at BURSA MALAYSIA are analyzed based on PCC-based MST and EVC-based MST. The stocks are classified according to their economic sector classifications defined in BURSA MALAYSIA. Then, the economic sectors network based on EVC is presented. The topological structures of inter and intra economic sectors are presented and discussed.

In Chapter 5, the case study of NYSE involving 1515 stocks that continuously traded in market are analyzed based on EVC-based MST. The topological properties of those stocks are analyzed. According to Standard Industrial Classification (SIC) system, the stocks are classified into 63 SICs. These 63 SICs are

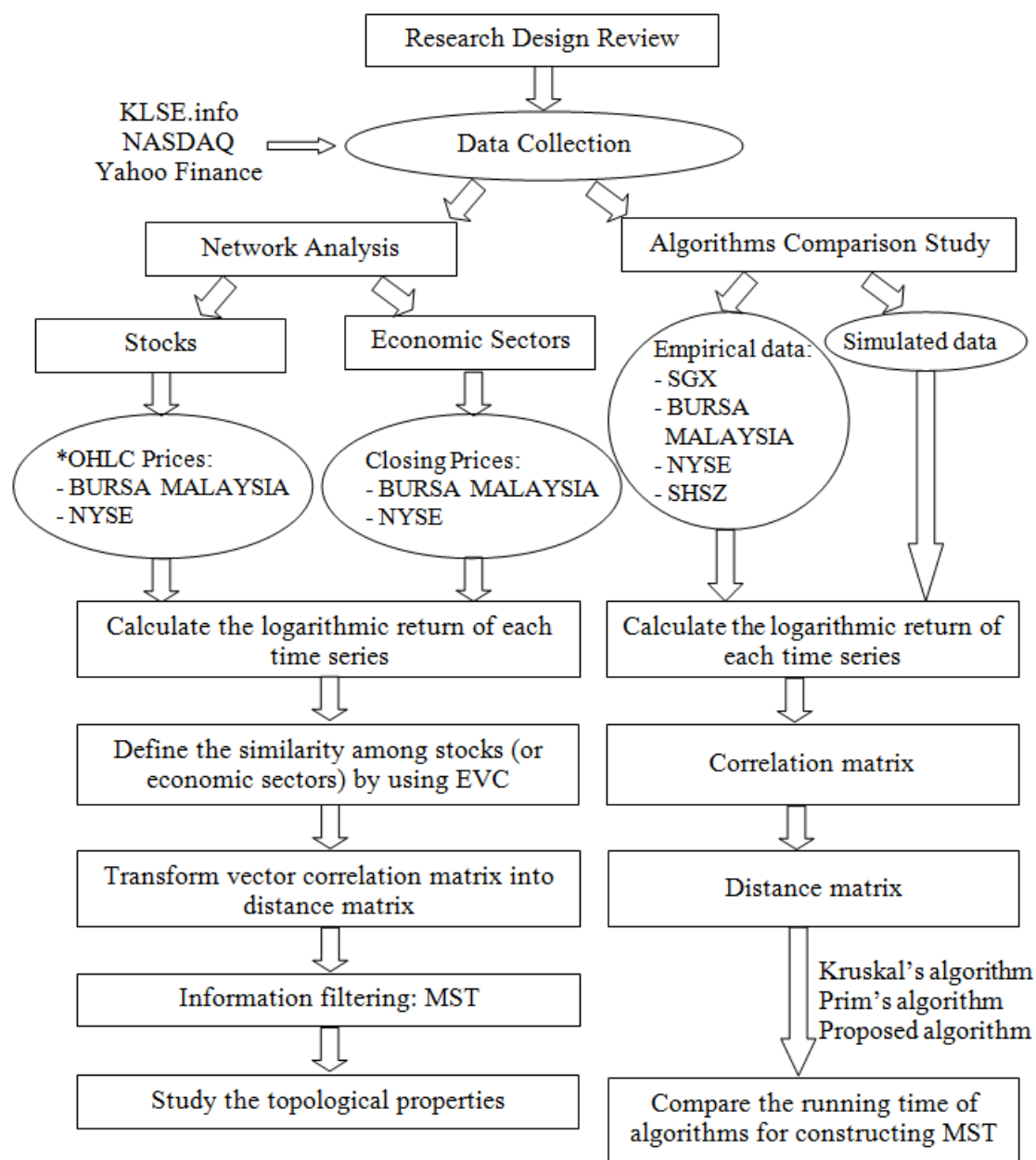
again classified into nine main industry sectors. The EVC-based SICs network is presented. The topological structures of inter and intra SICs are presented and discussed.

In Chapter 6, an illustrative example on how the proposed algorithm works is given. In order to show the advantage of the proposed algorithm, its running time is compared to Kruskal's algorithm and also to Prim's algorithm. For that purpose, four empirical data and four simulated data are used. For each data set, the running time of each of those algorithms are reported.

Finally, conclusion and recommendations for future research in the Chapter 7 which consists of some conclusions, contributions of the study and future research directions will close the presentation of this thesis.

## **1.6 Research Framework**

To facilitate the readers, the flow chart presented in Figure 1.1 gives the summary of the research framework. The left-hand side of the figure depicts the steps of conducting the stocks network analysis and economic sectors network analysis of the stock markets, i.e., BURSA MALAYSIA and NYSE in Chapters 4 and 5, respectively. The right-hand side shows the steps of conducting the comparison study between algorithms in Chapter 6.



\*OHLC is the abbreviation formed from opening, highest, lowest and closing.

**Figure 1.1** Summary of research framework

## REFERENCES

- Agarwal, R., Faloutsos, C. and Swami, A. (1993). Efficient Similarity Search in Sequence Databases. *Proceedings of the International Conference on Foundations of Data Organization and Algorithms*. 13-15 October. Chicago, USA, 69–84
- Allaire, J. and Lepage, Y. (1992). A Procedure for Assessing Vector Correlations. *Annals of the Institute of Statistical Mathematics*. 44(4), 755-768.
- Anderson, T. W. (2003). *An Introduction to Multivariate Statistical Analysis*. (3<sup>rd</sup> ed.) New York: John Wiley and Sons, Inc.
- Anton, H. (1991). *Elementary Linear Algebra*. (6<sup>th</sup> ed.) U.S.: Von Hoffmann Press.
- Bachelier, L. (1900). *Theorie de la Spéculation*. Thesis for the Doctorate in Mathematical Sciences. Gauthier-Villars, Paris. English translation in *The Random Character of Stock Market Prices*. (1964). Cootner, P. editor, MIT Press, Cambridge, pp.17-78.
- Bazlamacci, C. F. and Hindi, K. S. (2001). Minimum-weight Spanning Tree Algorithms: A Survey and Empirical Study. *Computers & Operations Research*. 28, 767-785.
- Becker, K. G., Finnerty, J. E. and Gupta, M. (1990). The Intertemporal Relation between the U.S. and Japanese Stock Markets. *Journal of Finance*. 45(4), 1297-1306.
- Bedi, J., Richards, A. J. and Tennat, P. (2003). The Characteristics and Trading Behavior of Dual-listed Companies. Reserve Bank of Australia Research Discussion Paper No. 2003-06. Australia: Reserve Bank of Australia.
- Boginski, V., Butenko, S. and Pardalos, P. M. (2005). Mining Market Data: A Network Approach. *Computational Statistics & Data Analysis*. 48(2), 431-443.
- Bonacich, P. (1972). Factoring and Weighting Approaches to Status Scores and Clique Identification. *Journal of Mathematical Sociology*. 2, 113-120.

- Bonacich, P. (1987). Power and Centrality: A Family of Measures. *American Journal of Sociology*. 92, 1170-1182.
- Bonacich, P. (1991). Simultaneous Group and Individual Centralities. *Social Networks*. 13(2), 155-168.
- Bonacich, P. (2007). Some Unique Properties of Eigenvector Centrality. *Social Networks*. 29(2), 555-564.
- Bonanno, G., Caldarelli, G., Lillo, F. and Mantegna, R. N. (2003). Networks Topology of Correlation Based Minimal Spanning Trees in Real and Model Markets. *Physical Review E*. 68, 046130(1)-046130(4).
- Bonanno, G., Caldarelli, G., Lillo, F., Miccichè, S., Vandewalle, N. and Mantegna, R. N. (2004). Networks of Equities in Financial Markets. *The European Physical Journal B*. 38, 363-371.
- Bonanno, G., Vandewalle, N. and Mantegna, R. N. (2000). Taxonomy of Stock Market Indices. *Physical Review E*. 62, R7615-R7618.
- Borgatti, S. P. and Everett, M. G. (2006). A Graph-Theoretic Perspective on Centrality. *Social Networks*. 28, 466-484.
- Borgatti, S. P. (1995). Centrality and AIDS. *Connections*. 18(1), 112-115.
- Borgatti, S. P. (2005). Centrality and Network Flows. *Social Networks*. 27, 55-71.
- Brida, J. G. and Risso, W. A. (2008). Multidimensional Minimal Spanning Tree: The Dow Jones Case. *Physica A*. 387, 5205-5210.
- Brida, J. G. and Risso, W. A. (2010). Hierarchical Structure of the German Stock Market. *Expert Systems with Applications*. 37, 3846-3852.
- Caldarelli, G., Marchetti R. and Pietronero, L. (2000). The Fractal Properties of Internet. *Europhysics Letters*. 52(4), 386-391.
- Campbell, J. Y., Lo, A. W. and MacKinlay, A. C. (1997). *The Econometrics of Financial Markets*. Princeton, N.J: Princeton University Press.
- Chakraborti, A. (2006). *An Outlook on Correlations in Stock Prices. Economics of Stock and other Markets New Economic Window*. Italy: Springer Verlag.
- Chan, L. and Lien, D. (2003). Using High, Low, Open, and Closing Prices to Estimate the Effects of Cash Settlement on Futures Prices. *International Review of Financial Analysis*. 12, 35-47.
- Chen, N.-F., Roll, R. and Ross, S. A. (1986). Economic Forces and the Stock Market. *The Journal of Business*. 59(3), 383-403.

- Cheong, S. A., Forna, R. P., Lee, G. H. T., Kok, J. L., Yim, W. S., Xu, D. Y. and Zhang, Y. (2012). The Japanese Economy in Crises: A Time Series Segmentation Study. *Economics: The Open-Access, Open Assessment E-Journal*. 6, 1-81. ZBW-Leibniz Information Centre for Economics.
- Christley, R. M., Pinchbeck, G. L., Bowers, R. G., Clancy, D., French, N. P., Bennett, R. and Turner, J. (2006). Contact Network Structure and Risk of Infection. *Proceedings of the 11<sup>th</sup> International Symposium on Veterinary Epidemiology and Economics*. 6-11 August. Cairns, Australia, 275-277.
- Clèroux, R. and Ducharme, G. R. (1989). Vector Correlation for Elliptical Distribution. *Communications in Statistics – Theory and Methods*. 18(4), 1441-1454.
- Coelho, R., Hutzler, S., Repetowicz, P. and Richmond, P., (2007). Sector Analysis for a FTSE Portfolio of Stocks. *Physica A*. 373, 615-626.
- Coelho, R., Gilmore, C. G, Lucey, B., Richmond, P. and Hutzler, S. (2007). The Evolution of Interdependence in World Equity Markets – Evidence From Minimum Spanning Trees. *Physica A*. 376, 455-466.
- Coronnello, C., Tumminello, M., Lillo, F., Miccichè, S. and Mantegna, R. N. (2008). Economic Sector Identification in a Set of Stocks Traded at the New York Stock Exchange: A Comparative Analysis. *Advanced Topics on Cellular Self-Organizing Nets and Chaotic Nonlinear Dynamics to Model and Control Complex Systems*. Singapore: World Scientific, pp. 159-180.
- Costenbader, E. and Valente, T. W. (2003). The Stability of Centrality Measures When Networks Are Sampled. *Social Networks*. 25, 283-307.
- Cupal, M., Deev, O. and Linnertová, D. (2012). Network Structures of the European Stock Markets. *Proceedings of 30th International Conference Mathematical Methods in Economics*. 11-13 September. Karvina, Czech Republic, 79-84.
- Djauhari, M. A. (1998). Operators Averaging and Its Application in Qualitative Variables Selection. *Jurnal Matematika & Sains*. 3(1), 41-49.
- Djauhari, M. A. (2011). Geometric Interpretation of Vector Variance. *MATEMATIKA*. 27(1), 51-57.
- Djauhari, M. A. (2012). A Robust Filter in Stock Networks Analysis. *Physica A*. 391, 5049-5057.
- Dykstra, R. L. (1970). Establishing the Positive Definiteness of the Sample Covariance Matrix. *The Annals of Mathematical Statistics*. 41(6), 2153-2154.



- Elton, E. J. and Gruber, M. J. (1971). Improved Forecasting Through the Design of Homogeneous Groups. *The Journal of Business*. 44(4), 432-450.
- Eom, C., Oh, G. and Kim, S. (2007). Topological Properties of a Minimal Spanning Tree in the Korean and the American Stock Markets. *Journal of the Korean Physical Society*. 51(4), 1432-1436.
- Eom, C., Oh, G. and Kim, S. (2008). Statistical Investigation on Connected Structures of Stock Networks in a Financial Time Series. *Journal of the Korean Physical Society*. 53(6), 3837–3841.
- Eom, C., Oh, G., Jung, W.-S., Jeong, H. and Kim, S. (2009). Topological Properties of Stock Networks Based on Minimal Spanning Tree and Random Matrix Theory in Financial Time Series. *Physica A*. 388, 900-906.
- Eryiğit, M. and Eryiğit, R. (2009). Network Structure of Cross-Correlations among the World Market Indices. *Physica A*. 388, 3551-3562.
- Escoufier, Y. (1973). Le Traitement des Variables Vectorielles. *Biometrics*, 29, 751-760.
- Escoufier, Y. (2006). Operator Related to A Data Matrix: A Survey. *Comstat 2006 – Proceedings in Computational Statistics*. 28 August – 1 September. Rome, Italy, 285-297.
- Figueiredo, A., Figueiredo, F., Monteiro, N. P. and Straume, O. R. (2012). Restructuring in Privatised Firms: A Statis Approach. *Structural Change and Economic Dynamics*. 23(1), 108- 116.
- Freeman, L. C. (1979). Centrality in Social Networks: I. Conceptual Clarification. *Social Networks*. 1, 215-239.
- Friedkin, N. A. (1991). Theoretical Foundations for Centrality Measures. *American Journal of Sociology*. 96(6), 1478-1504.
- Gałaszka, M. (2011). Characteristics of the Polish Stock Market Correlations. *International Review of Financial Analysis*. 20, 1-5.
- Gan, S. L. and Djauhari, M. A. (2012). Stock Network Analysis in Kuala Lumpur Stock Exchange. *Malaysian Journal of Fundamental & Applied Science*. 8(2), 60-66.
- Garas, A. and Argyrakis, P. (2007). Correlation Study of the Athens Stock Exchange. *Physica A*. 380, 399-410.
- Garman, M. B. and Klass, M. J. (1980). On The Estimation of Security Price Volatilities from Historical Data. *Journal of Business*. 53, 67–78.

- Goetzmann, W. N., Li, L. and Rouwenhorst, K. G. (2005). Long-Term Global Market Correlations. *Journal of Business*. 78(1), 1-38.
- Gower, J. C. (1966). Some Distance Properties of Latent Root and Vector Methods Used in Multivariate Analysis. *Biometrika*. 53(3/4), 325-338.
- Graham, R. L. and Hell, P. (1985). On the History of the Minimum Spanning Tree Problem. *Annals of the History of Computing*. 7(1), 43-57.
- Greenberg, H. J. (1998). Greedy Algorithms for Minimum Spanning Tree. Denver: University of Colorado. Available at <http://www.cudenver.edu/hgreenbe/>.
- Gu, R., Shao, Y. and Wang, Q. (2013). Is the Efficiency of Stock Market Correlated with Multifractality? An Evidence from the Shanghai Stock Market. *Physica A*. 392, 361-370.
- Haksever, C. and Render, B. (2013). Chapter 1: The Important Role Services Play in An Economy. In *Service Management: An Integrated Approach to Supply Chain Management and Operations*. New Jersey: FT Press.
- Halinen, A. and Tornroos, J. (1998). The Role of Embeddedness in the Evolution of Business Network. *Scandinavian Journal of Management*. 14(3), 187-205.
- Hayashi, T. and Yoshida, N. (2005). On Covariance Estimation of Non-synchronously Observed Diffusion Processes. *Bernoulli*. 11(2), 359-379.
- Holton, G. A. (2003). *Value-at-Risk: Theory and Practice*. San Diego: Academia Press. Retrieved from <http://value-at-risk.net>.
- Horst, E. T., Rodriguez, A., Gzyl, H. and Molina, G. (2012). Stochastic Volatility Models Including Open, Close, High, and Low Prices. *Quantitative Finance*. 12, 199-212.
- Huang, F., Gao, P. and 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*. 7-8 November. Shanghai, China, 237-241.
- Huang, W. Q., Zhuang, X. T. and Yao, S. (2009). A Network Analysis of the Chinese Stock Market. *Physica A*. 388, 2956-2964.
- Jang, W., Lee, J., and Chang, W. (2011). Currency Crises and the Evolution of Foreign Exchange Market: Evidence from Minimum Spanning Tree. *Physica A*. 390, 707-718.
- Josse, J. and Holmes, S. (2014). Test of Independence and Beyond. Available at arXiv:1307.7383v3.

- Josse, J., Pagès, J. and 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. and Moon, H.-T. (2006). Characteristics of the Korean Stock Market Correlations. *Physica A*. 361, 263-271.
- Kantar, E., Keskin, M. and Deviren, B. (2012). Analysis of the Effects of the Global Financial Crisis on the Turkish Economy, Using Hierarchical Methods. *Physica A*. 391, 2342-2352.
- Kenett, D. Y., Tumminello, M., Madi, A., Gur-Gershgoren, G., Mantegna, R. N. and Ben-Jacob, E. (2010). Dominating Clasp of the Financial Sector Revealed by Partial Correlation Analysis of the Stock Market. *PLoS ONE*. 5(12), e15032(1)- e15032(14).
- Keogh, E. and Ratanamahatana, C. A. (2005). Exact Indexing of Dynamic Time Warping. *Knowledge and Information Systems*. 7(3), 358-386.
- Keskin, M., Deviren, B. and Kocakaplan, Y. (2011). Topology of the Correlation Networks among Major Currencies Using Hierarchical Structure Methods. *Physica A*. 390, 719-730.
- Kim, D.-H. and Jeong, H. (2005). Systematic Analysis of Group Identification in Stock Markets. *Physical Review E*. 72, 046133-1-046133-8.
- Kitsak, M., Gallos, L. K., Havlin, S., Liljeros, F., Muchnik, L., Stanley, H. E. and Makse, H. A. (2010). Identification of Influential Spreaders in Complex Networks. *Nature Physics*. 6, 888-893.
- Kocakaplan, Y., Doğan, Ş., Deviren, B. and Keskin, M. (2013). Correlations, Hierarchies and Networks of World's Automotive Companies. *Physica A*. 392, 2736-2774.
- Kollo, T. and von Rosen, D. (2005). *Advanced Multivariate Statistics with Matrices*. The Netherlands: Springer.
- Kruskal, J. B. (1956). On the Shortest Spanning Subtree of a Graph and the Traveling Salesman Problem. *Proceedings of the American Mathematical Society*. 7, 48-50.
- Laloux, L., Cizeau, P., Bouchaud, J.-P. and Potters, M. (1999). Noise Dressing of Financial Correlation Matrices. *Physical Review Letters*. 83(7), 1467.
- Landherr, A., Friedl, B. and Heidemann, J. (2010). A Critical Review of Centrality Measures in Social Networks. Discussion Paper WI-282, University of Augsburg.

- Lin, W.-L., Engle, R. F. and Ito, T. (1994). Do Bulls and Bears Move Across Borders? International Transmission of Stock Returns and Volatility. *The Review of Financial Studies*. 7(3), 507-538.
- Lo, A. W. and MacKinlay, A. C. (1990). An Econometric Analysis of Nonsynchronous Trading. *Journal of Econometrics*. 45, 181-212.
- Malkevitch, J. Trees: A Mathematical Tool for All Seasons. American Mathematical Society, Feature Column, May 2012.
- Mantegna, R. N. and Stanley, H. E. (2000). *An Introduction to Econophysics: Correlation and Complexity in Finance*. Cambridge University Press, United Kingdom.
- Mantegna, R. N. (1999). Hierarchical Structure in Financial Markets. *The European Physical Journal B*. 11, 193-197.
- Ma, R. (2009). Designing Interactive Visualization Methods for Comparing Multivariate Stock Market Data over Time. Master Thesis. Faculty of Informatics, TU Vienna.
- McMillan, D. G. (2002). Non-linear Predictability of UK Stock Market Returns. *Oxford Bulletin of Economics and Statistics*. 65(5), 557-573.
- Miccichè, S., Bonanno, G., Lillo, F. and Mantegna, R. N. (2003). Degree Stability of Minimum Spanning Tree of Price Return and Volatility. *Physica A*. 342, 66-73.
- Miccichè, S., Lillo, F. and Mantegna, R. N. (2005). Correlation Based Hierarchical Clustering in Financial Time Series. *Proceedings of the 31<sup>st</sup> Workshop of the International School of Solid State Physics*. 20-26 July. Sicily, Italy.
- Mizuno, T., Takayasu, H. and Takayasu, M. (2006). Correlation Networks among Currencies. *Physica A*. 364, 336-342.
- Muirhead, R. J. (1982). *Aspect of Multivariate Statistical Theory*. New York: John Wiley & Sons, Inc.
- Münnix, M. C., Shimada, T., Schäfer, R., Leyvraz, F., Seligman, T. H., Guhr, T. and Stanley, H. E. (2012). Identifying States of a Financial Market. *Nature Scientific Reports*. 2, 644(1)-644(6).
- Murtagh, F. (2005). Identifying the Ultrametricity of Time Series. *The European Physical Journal B*. 43, 573-579.
- Naylor, M. J., Rose, L. C. and Moyle, B. J. (2007). Topology of Foreign Exchange Markets Using Hierarchical Structure Methods. *Physica A*. 382, 199-208.

- Nesetril, J. (1997). A Few Remarks on the History of MST-Problem. *Archivum Mathematicum*. 33, 15-22.
- Newman, M. E. J. (2008). *Mathematics of Networks*. *The New Palgrave Encyclopedia of Economics*. (2<sup>nd</sup> ed.) Basingstoke: Palgrave Macmillan.
- Newman, M. E. J. (2011). Complex Systems: A Survey. *American Journal of Physics*. 79, 800-810.
- Ni, C., Sugimoto, C. R. and Jiang, J. (2011). Degree, Closeness, and Betweenness: Application of Group Centrality Measurements to Explore Macro-Disciplinary Evolution Diachronically. *Proceedings of the 13<sup>th</sup> Meetings of International Society for Scientometrics and Informetrics (ISSI)*. 4-7 July. Durban, South Africa, 1-13.
- Oh, G. (2014). Grouping Characteristics of Industry Sectors in Financial Markets. *Physica A*. 395, 261-268.
- Onnela, J.-P., Chakraborti, A., Kaski, K. and Kertész, J. (2002). Dynamic Asset Trees and Portfolio Analysis. *The European Physical Journal B*. 30, 285-288.
- Onnela, J.-P., Chakraborti, A. and Kaski, K. (2003a). Dynamics of Market Correlations: Taxonomy and Portfolio Analysis. *Physical Review E*. 68, 056110-1-056110-12.
- Onnela, J.-P., Chakraborti, A., Kaski, K. and Kertész, J. (2003b). Dynamic Asset Trees and Black Monday. *Physica A*. 324, 247-252.
- Onnela, J.-P., Chakraborti, A., Kaski, K., Kertész, J. and Kanto, A. (2003c). Asset Trees and Asset Graphs in Financial Markets. *Physica Scripta*. T106, 48-54.
- Osborne, M. F. M. (1959). Brownian Motion in the Stock Market. *Operations Research*. 7, 145-173.
- Papoulis, A. and Pillai, S. U. (2002). *Probability, Random Variables and Stochastic Processes*. McGraw-Hill, pg. 211. ISBN 0-07-366011-6.
- Park, H., Han, S., Rojas, E., Son, J. and Jung, W. (2011). Social Network Analysis of Collaborative Ventures for Overseas Construction Projects. *Journal of Construction Engineering Management*. 137(5), 344-355.
- Parkinson, M. (1980). The Extreme Value Method for Estimating the Variance of the Rate of Return. *Journal of Business*. 53, 61-65.
- Plerou, V., Gopikrishnan, P., Rosenow, B., Nunes Amaral, L. A., Guhr, T. and Stanley, H. E. (2002). Random Matrix Approach to Cross Correlations in Financial Data. *Physical Review E*. 65, 066126(1)- 066126(18).

- Plerou, V., Gopikrishnan, P., Rosenow, B., Nunes Amaral, L. A. and Stanley, H. E. (1999). Universal and Nonuniversal Properties of Cross Correlations in Financial Time Series. *Physical Review Letters*. 83(7), 1471-1474.
- Podobnik, B., Jiang, Z.-Q., Zhou, W.-X. and Stanley, H. E. (2011). Statistical Tests for Power-Law Cross-Correlated Processes. *Physical Review E*. 84 066118(1)- 066118(8).
- Podobnik, B. and Stanley, H. E. (2008). Detrended Cross-Correlation Analysis: A New Method for Analyzing Two Nonstationary Time Series. *Physical Review Letters*. 100, 084102.
- Preis, T., Kenett, D. Y., Stanley, H. E., Helbing, D. and Ben-Jacob, E. (2012). Quantifying the Behavior of Stock Correlations under Market Stress. *Nature Scientific Reports*. 2, 752(1)-752(5).
- Prim, R. C. (1957). Shortest Connection Networks and Some Generalizations. *Bell System Technical Journal*. 36(6), 1389-1401.
- Rammal, R., Toulouse, G. and Virasoro, M. A. (1986). Ultrametricity for Physicists. *Reviews of Modern Physics*. 58(3), 765-788.
- Robert, P. and Escoufier, Y. (1976). A Unifying Tool for Linear Multivariate Statistical Methods: The RV-Coefficient. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*. 25(3), 257-265.
- Robert, P., Cleroux, P. and Ranger, N. (1985). Some Results on Vector Correlation. *Computational Statistics & Data Analysis*. 3, 25-32.
- Rogers, L. C. G. and Satchell, S. E. (1991). Estimating Variance from High, Low and Closing Prices. *Annals of Applied Probability*. 1, 504–512.
- Rogers, L. C. G. and Zhou, F. (2008). Estimating Correlation from High, Low, Opening and Closing Prices. *The Annals of Applied Probability*. 18(2), 813-823.
- Ross, S. A. (1976). The Arbitrage Theory of Capital Asset Pricing. *Journal of Economic Theory*. 13, 341-360.
- Roy, R. B. and Sarkar, U. K. (2011). Identifying Influential Stock Indices from Global Stock Markets: A Social Network Analysis Approach. *Procedia Computer Science*. 5, 442-449.
- Roy, R. B. and Sarkar, U. K. (2013). Identifying Dominant Economic Sectors and Stock Markets: A Social Network Mining Approach. *Trends and*

*Applications in Knowledge Discovery and Data Mining Lecture Notes in Computer Science*. 7867, 59-70.

- Samsi, S. M., Yusof, Z. and Cheong, K.-C. (2012). Financial Sector and Economic Growth in Malaysia: Sectoral Shock Analysis. *IPEDR*. 55(16), 82-85.
- Samuelson, P. A. (1965). Rational Theory of Warrant Pricing. *Industrial Management Review*. 6(2), 13-31.
- Schott, J. R. (1997). *Matrix Analysis for Statistics*. New York: John Wiley & Sons, Inc.
- Sensoy, A. and Tabak, B. M. (2014). Dynamic Spanning Trees in Stock Market Networks: The Case of Asia Pacific. Working Paper 351, Banco Central Do Brasil, Brazil.
- Sieczka, P. and Hołyst, J. A. (2009). Correlations in Commodity Markets. *Physica A*. 388, 1621-1630.
- Silva, A. C. (2005). *Applications of Physics to Finance and Economics: Returns, Trading Activity and Income*. Doctor of Philosophy, University of Maryland, College Park.
- Sinha, S., Chatterjee, A., Chakraborti, A. and Chakrabarti, B. K. (2010). *Econophysics: An Introduction*. Berlin: Wiley-VCH.
- Skaradzinski, D. A. (2003). The Nonlinear Behavior of Stock Prices: The Impact of Firm Size, Seasonality, and Trading Frequency. Doctor of Philosophy, Virginia Polytechnic Institute and State University.
- Sreedharan, N. (2004). A Vector Error Correction Model (VECM) of Stock Market Returns. Discussion Paper 2004-06. University of Tasmania, Australia.
- Stanley, H. E. (2003). Statistical Physics and Economic Fluctuations: Do Outliers Exist? *Physica A*. 318, 279-292.
- Szabo, G., Alava, M. and Kertesz, J. (2003). Geometry of Minimum Spanning Tree on Scale-Free Networks. *Physica A*. 330, 31-36.
- Tabak, B. M., Serra, T. R. and Cajueiro, D. O. (2010). Topological Properties of Stock Market Networks: The Case of Brazil. *Physica A*. 389, 3240-3249.
- Taqqu, M. S. (2001). Bachelier and His Times: A Conversation with Bernard Bru. *Finance and Stochastic*. 5(1), 3-32. Reprinted in *Mathematical Finance-Bachelier Congress 2000*, German H. *et al.* editors, 2002, Springer-Verlag Berlin Heidelberg, 1-45.

- Tola, V., Lillo, F., Gallegati, M. and Mantegna, R. N. (2008). Cluster Analysis for Portfolio Optimization. *Journal of Economic Dynamics & Control*. 32, 235-258.
- Tumminello, M., Aste, T., Di Matteo, T. and Mantegna, R. N. (2005). A Tool for Filtering Information in Complex Systems. *PNAS*. 102, 10421-10426.
- Tumminello, M., Coronello, C., Lillo, F., Miccichè, S. and Mantegna, R. N. (2007). Spanning Trees and Bootstrap Reliability Estimation in Correlation-Based Networks. *International Journal of Bifurcation Chaos*. 17, 2319-2329.
- Tumminello, M., Lillo, F. and Mantegna, R. N. (2010). Correlation, Hierarchies, and Networks in Financial Markets. *Journal of Economic Behavior & Organization*. 75, 40-58.
- Umadevi, V. (2013). Automatic Co-Authorship Network Extraction and Discovery of Central Authors. *International Journal of Computer Applications*. 74(4), 1-6.
- Valente, T. W., Coronges, K., Lakon, C. and Costenbader, E. (2008). Hoe Correlated Are Network Centrality Measures? *Connect (Tor)*. 28(1), 16-26.
- Vandewalle, N., Brisbois F. and Tordoir, X. (2001). Non-Random Topology of Stock Markets. *Quantitative Finance*. 1, 372-374.
- Wambeke, B. W., Liu, M. and Hsiang, S. M. (2012). Using Pajek and Centrality Analysis to Identify a Social Network of Construction Trades. *Journal of Construction Engineering Management*. 138(10), 1192-1201.
- Wang, G.-J., Xie, C., Chen, S., Yang, J.-J. and Yang, M.-Y. (2013). Random Matrix Theory Analysis of Cross-Correlations in the US Stock market: Evidence from Pearson's Correlation Coefficient and Detrended Cross-Correlation Coefficient. *Physica A*. 392, 3715-3730.
- Wang, G.-J., Xie, C., Han, F. and Sun, B. (2012). Similarity Measure and Topology Evolution of Foreign Exchange Markets Using Dynamic Time Warping Method: Evidence from Minimal Spanning Tree. *Physica A*. 391, 4136-4146.
- Wang, J., Zhu, Y., Li, S., Wan, D. and Zhang, P. (2014). Multivariate Time Series Similarity Searching. *The Scientific World Journal*. 2014, 1-8.
- West, D. B. (2001). *Introduction to Graph Theory*. (2<sup>nd</sup> ed.) New Jersey: Prentice Hall.
- Wigner, E. P. (1951). On A Class of Analytic Functions of Quantum Theory of Collisions. *The Annals of Mathematics*. 53, 36-67.



- Wilmott, P. (2007). *P. Wilmott Introduces Quantitative Finance*. New York: John Wiley & Sons, Inc.
- Xu, Y., Ma, J. Sun, Y., Hao, J., Sun, Y. and Zhao, Y. (2009). Using Social Network Analysis as a Strategy for Ecommerce Recommendation. *Pacific Asia Conference on Information Systems (PACIS)*. 10-12 July. Hyderabad, India.
- Yamashita, Y. and Yodahisa, H. (2012). Similarity Measure and Clustering Algorithm for Candlestick Valued Data. *Joint Meeting of Japanese and Italian Classification Societies (JCS-CLADAG)*. 3-4 September. Capri Island, Italy, 1-4.
- Yang, D. and Zhang, Q. (2000). Drift-Independent Volatility Estimation Based on High, Low, Open, and Close Prices. *Journal of Business*. 73(3), 477-491.
- Yang, K. and Shahabi, C. (2004). A PCA-based Similarity Measure for Multivariate Time Series. *Proceedings of the 2<sup>nd</sup> ACM International Workshop on Multimedia Databases*. 8-13 November. Washington, USA, 65-74.
- Zebende, G. F. (2011). DCCA Cross-Correlation Coefficient: Quantifying Level of Cross-Correlation. *Physica A*. 390, 614-618.
- Zhang, Y., Lee, G. H. T., Wong, J. C., Kok, J. L., Prusty, M. and Cheong, S. A. (2011). Will the US Economy Recover in 2010? A Minimal Spanning Tree Study. *Physica A*. 390, 2020-2050.
- Zhuang, R., Hu, B. and Ye, Z. (2008). Minimal Spanning Tree for Shanghai-Shenzhen 300 Stock Index. *IEEE World Congress on Computational Intelligence*. 1-6 June. Hong Kong, 1417-1424.

## APPENDIX A

### PUBLICATIONS

#### I. Journal Publications

1. Gan, S. L. and Djauhari, M. A. (2012). Stock Network Analysis in Kuala Lumpur Stock Exchange. *Malaysian Journal of Fundamental & Applied Sciences*. 8(2), 60-66. **(Indexed in Google Scholar)**
2. Gan, S. L. and Djauhari, M. A. (2012). An Overall Centrality Measure: The Case of U.S Stock Market. *International Journal of Basic & Applied Sciences IJBAS-IJENS*. 12(6), 99-103 **(Indexed in Google Scholar with I.F: 0.9878)**
3. Djauhari, M. A. and Gan, S. L. (2013). Minimal Spanning Tree Problem in Stock Networks Analysis: An Efficient Algorithm. *Physica A: Statistical Mechanics and its Applications*. 392(2), 2226-2234. **(Indexed in SCOPUS with I.F: 1.722)**
4. Djauhari, M. A. and Gan, S. L. (2013). Network Topology Property of English Dialects: A Robust Filter Approach. *American Journal of Applied Sciences*. 10(7), 646-653. **(Indexed in SCOPUS)**
5. Djauhari, M. A. and Gan, S. L. (2014). Dynamics of Correlation Structure in Stock Market. *Entropy*. 16(1), 455-470. **(Indexed in SCOPUS with I.F: 1.564)**
6. Djauhari, M. A. and Gan, S. L. (2015). Optimality Problem of Network Topology in Stocks Market Analysis. *Physica A: Statistical Mechanics and its Applications*. 419, 108-114. **(Indexed in SCOPUS with I.F: 1.722)**
7. Djauhari, M. A. and Gan, S. L. (2015). Bursa Malaysia Stocks Market Analysis: A Review. *ASM Science Journal*. 8(2), 150. **(Indexed in SCOPUS)**
8. Gan, S. L. and 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, P12005. **(Indexed in SCOPUS with I.F: 2.404)**

9. Djauhari, M. A. and Gan, S. L. (2014). Network Topology of Economic Sectors. *Physica A: Statistical Mechanics and its Applications*. **(Revised version is under review) (Indexed in SCOPUS with I.F: 1.722)**

## II. International and National Conference Proceedings

1. Gan, S. L. and Djauhari, M. A. (2012). Network Topology of Indonesian Stock Market. *Proceedings of IEEE International Conference on Cloud Computing and Social Networking*. 26-27 April 2012. Bandung, Indonesia, 1-4. **(Indexed in SCOPUS)**
2. Gan, S. L. and Djauhari, M. A. (2012). Filtering Network Topology in Stock Market. *Proceedings of the 1<sup>st</sup> ISM International Statistical Conference*. 4-6 September 2012. Johor Bahru, Malaysia, 155-162.
3. Gan, S. L. and Djauhari, M. A. (2012). Sub-Dominant Ultrametric Hierarchical Structures of NYSE 100 Stocks. *Proceedings ICSSBE 2012 International Conference on Statistics in Science, Business and Engineering*. 10-12 September 2012. Langkawi, Malaysia, 429-433. **(Indexed in SCOPUS)**
4. Gan, S. L. and Djauhari, M. A. (2013). Multidimensional Stock Network Analysis: An Escoufier's RV Coefficient Approach. *Proceedings ICMSS 2013 International Conference on Mathematical Sciences and Statistics*. 5-7 February 2013. Kuala Lumpur, Malaysia, 550-555. **(Indexed in SCOPUS)**
5. Gan, S. L., Djauhari, M. A. and Ismail, Z. (2014). Economic Sectors Network at Kuala Lumpur Stock Exchange. *ISI-RSC 2014 International Statistical Institute Regional Statistics Conference 2014*. 16-19 November 2014. Kuala Lumpur, Malaysia.
6. Gan, S. L., Djauhari, M. A. and Ismail, Z. (2014). Monitoring Correlation Structures Stability in Foreign Exchange Market. *IEEE 2014 International Conference on Industrial Engineering and Engineering Management*. 9-12 December 2014. Selangor, Malaysia, 848-852. **(Indexed in SCOPUS)**

### **III. Book Chapters**

1. Djauhari, M. A. and Gan, S. L. (2014). *Book Chapters: Statistics in Scientific Investigation III Financial Network Analysis*. Chapter 2: A Robust Filtered Network Analysis of Malaysian Stocks Market, pp. 20-41. **(In press)**