ASM Sci. J., **8(2)**, 150–158

Bursa Malaysia Stocks Market Analysis: A Review

M.A. Djauhari[1]* and S.L. Gan[2]

We had reviewed the current practice in stocks market analysis where stock is represented by its closing price, and then found that this approach may be misleading. In actuality, in the daily activity of stocks market, stock is represented by four prices, namely 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. In this paper all four prices will be considered. Then, the notion of multivariate time series similarity among stocks will be developed as a generalisation of univariate time series similarity. The results are used to construct stocks network in multivariate setting. To filter the economic information contained in that network, the standard tools in econophysics is used. Furthermore, the topological properties of stocks are analysed by using the most common centrality measures. As an example, Bursa Malaysia data are investigated and we show that the proposed approach can better figure out the real situation compared to the current one.

Key words: Embeddedness; minimal spanning tree; multivariate time series; *RV*-coefficient; stocks network

In current practice, stock is considered as a univariate time series of closing price. Under the assumption that log price returns are independent and identically distributed (i.i.d), the similarity among stocks is measured by using Pearson correlation coefficient (PCC) that is based on log price returns data. Those consideration and assumption allow us to represent stocks market as a complex system in the form of complex network (Bonanno *et al*. 2003, 2004; Jung *et al*. 2006; Garas & Argyrakis 2007; Brida & Risso 2008; Sieczka & Holyst 2009; Galazka 2011). More specifically, stock is represented as a node and the relationship between two stocks as a link (Tumminello *et al*. 2005; Onnela 2006; Kitsak *et al*. 2010; Newman 2011). According to graph theory, that network is an undirected weighted complete and finite graph (West 2001).

Essentially, in daily market activity, stock is considered as multivariate time series composed by opening, highest, and lowest prices besides closing price. This consideration, which is closer to the reality, is not new in stocks network analysis. For example, Brida & Risso (2008) talk on bivariate time series similarity of closing price and volume transformed into univariate time series through the use of symbolic time series analysis (STSA). They noted that if we used only a univariate time series of closing price to obtain the structure and taxonomy of the stocks market, we would lose the possibility of embodying information from other variables such as trading volume. Another example is the work of Yamashita and Yodahisa (2012), who considered stock as a multivariate time series of its opening, highest, lowest and closing prices. They asserted that the use these four prices provided more information than that of the information provided by only the closing price. Hence, to construct the network among stocks in this setting, they proposed a particular 'correlation' coefficient. However, in the generalisation of PCC, neither of those two functions, which means that their similarity measures cannot be used in a special case where stock is represented by only its closing price. This motivated us to develop a similarity measure among the multivariate time series of opening, highest, lowest and closing prices which generalise PCC.

In this paper we have opted to use Escoufier's operator, which characterises a multivariate data set by an operator, to generalise PCC. If PCC is the cosine of the angle between two variables, then we define the correlation of two variable vectors as the cosine of the angle between two operators. This approach allows us to define the notion of similarity among stocks in multivariate setting and then construct stocks network. The rest is to investigate the topological properties of stocks. For that purpose, as suggested in Mantegna (1999) and Mantegna and Stanley (2000), minimal spanning tree (MST) is used.

In the next section, the notion of multivariate correlation will be introduced by representing random vector as an

[1]Institute for Mathematical Research (INSPEM), Universiti Putra Malaysia, 43400 UPM Serdang, Selangor Darul Ehsan, Malaysia
2 Department of Mathematical Sciences, Faculty of Science, Universiti Teknologi Malaysia, 81310 UTM Skudai, Johor, Malaysia
*Corresponding author (e-mail: maman_abd@upm.edu.my)

Escoufier's operator. Since, algebraically, PCC is the cosine of the angle between two vectors, the notion of similarity among stocks in multivariate setting is developed by using that algebraic concept on two Escoufier's operators. Section 3 is devoted for multivariate stocks network construction and analysis. In Section 4, to illustrate the advantages of the proposed methodology, Bursa Malaysia data will be analysed. The results will be compared to those provided by the standard methodology based on univariate approach and discussed in Section 5. Finally, concluding remarks in the last section is highlighted to close the presentation. The list of company names involved in this study is given in the appendix.

Similarity among Stocks in Multivariate Setting

For simplicity, we write $p_i(t,1)$, $p_i(t,2)$, $p_i(t,3)$ and $p_i(t,4)$ the opening, highest, lowest and closing prices of stock i; $i = 1, 2, \ldots, n$. Here n is the number of stocks under study. Let,

$$r_i(t,m) = \ln p_i(t,m) - \ln p_i(t-1,m) \quad (1)$$

the logarithm of the m-th price returns of stock i at time t; $m = 1, 2, 3, 4$. Theoretically, each of the four stock's prices is assumed to be a geometric Brownian motion (GBM) process for all stocks. This means that $r_i(t,m)$ are independent and identically normally distributed (i.i.n.d) for all m (Wilmott 2007). However, in practice, it is sufficient to assume that $r_i(t,m)$ are i.i.d. This is to ensure that PCC and Escoufier's operator are working well.

That assumption allows us to consider $r_i(t,1)$, $r_i(t,1)$, $r_i(t,3)$ and $r_i(t,4)$ in (1) as the components of a random vector. Thus, stock is considered as a multivariate object. In order to be able to speak about the 'correlation' of two random vectors, in the next paragraph random vector will be represented as an Escoufier's operator. Since the set of all such operators is a Hilbert-Schmidt space (Escoufier 1973, 1976), we all know how to define the angle between two operators.

Operator Representation of Random Vector

Escoufier's operator (Escoufier 1973, 1976; Djauhari 2011) is a very powerful tool to simplify multivariate analysis into univariate-like. In practice, it is closely related to the notion of principal components of a random vector. Let us start by considering L_2 (Ω, $\mathcal{A}$, P) the set of all centred random variables of finite variance defined on a probability space (Ω, $\mathcal{A}$, P). We denote $X = (X_1, X_2, \ldots, X_p)^t$ a random vector of p components; X_k is in L_2 (Ω, $\mathcal{A}$, P) for all $k = 1, 2, \ldots, p$. To that random vector, Esoufier (1973, 1976) associated the following operator φ_X from L_2 (Ω, $\mathcal{A}$, P) into itself,

$$\varphi_{\boldsymbol{X}}(Y) = \sum_{k=1}^{p} E(X_k Y) X_k \text{ ; for all Y in } L_2 (\Omega, \mathcal{A}, \text{P}).$$

In matrix notation,

$$\varphi_X(Y) = E(XY)\, X.$$

We recognise $E(XY)$ the covariance matrix of random vector X and random variable Y. The role of operator is to put Y in the coordinate system of $X_1, X_2, \ldots, X_p$. It is then a linear combination of those random variables. The coefficient of Y on X_k, i.e., the scalar product in the space L_2 (Ω, $\mathcal{A}$, P), is $E(X_kY)$ the covariance of Y and X_k.

The following theorem shows that the operator φ_X characterises the random vector X in the sense that there is a one-to-one correspondence between the set $\mathcal{P}$ of all random vectors in L_2 (Ω, $\mathcal{A}$, P) and the set $\mathcal{R}$ of all such operators. The proof can be seen in Escoufier (1973).

Theorem. Let Σ_{XX} be the covariance matrix of X assumed to be positive definite. If $\mathbf{u} \in R^p$ such that $\Sigma_{XX}\mathbf{u} = \lambda\mathbf{u}$, then $Y = \mathbf{u}^t X$ satisfies $\varphi_X(Y) = \lambda Y$. Conversely, if $Y \in L_2$ (Ω, $\mathcal{A}$, P) such that $\varphi_X(Y) = \lambda Y$, then $Y = \mathbf{u}^t X$ where $\mathbf{u}$ satisfies $\Sigma_{XX}\mathbf{u} = \lambda\mathbf{u}$.

From this theorem we learn that (i) φ_X and Σ_{XX} have the same positive eigenvalues, and (ii) if $\lambda \neq 0$, then Y is the principal component of X with variance λ.

Correlation of Two Random Vectors

The operator φ_X is credited to Escoufier and called Escoufier's operator associated with the random vector X. From the theorem we have a very important consequence. Two random vectors are declared similar if and only if their covariance matrices have the same eigensystem. With this, the study of random vectors is transformed into the study of operators. Therefore, since L_2 (Ω, $\mathcal{A}$, P) with the scalar product defined by covariance is a Hilbert space, in $\mathcal{R}$ we can define a scalar product,

$$< \varphi_X, \varphi_Y > = Tr\,(\Sigma_{XY}\Sigma_{YX})$$

for all X and Y in $\mathcal{P}$ where Tr is the trace operator. It is the scalar product in $\mathcal{R}$ induced by the covariance. Consequently,

$$\|\varphi_{\boldsymbol{X}}\| = \sqrt{Tr\left(\Sigma^2_{\boldsymbol{XX}}\right)} \text{ and } \|\varphi_{\boldsymbol{Y}}\| = \sqrt{Tr\left(\Sigma^2_{\boldsymbol{YY}}\right)},$$

are the length of operators φ_X and φ_Y, respectively, and

$$\rho_{XY} = \frac{\Sigma_{\boldsymbol{XY}}\Sigma_{\boldsymbol{YX}})}{\sqrt{Tr\left(\Sigma_{\boldsymbol{XX}}\right)}\sqrt{Tr\left(\Sigma_{\boldsymbol{YY}}\right)}} \quad (2)$$

is the cosine of the angle between the two φ_X operators and φ_Y. The parameter ρ_{XY} in (2) defines the 'correlation' of random vectors X and Y. In the literature, it is called vector correlation of X and Y.

Sample Version of Vector Correlation

Consider two stocks represented as random vectors X and Y each of $p = 4$ and $q = 4$ components. The m-th component of X and Y are $r_X(t,m)$ and $r_Y(t,m)$; $m = 1, 2, 3, 4$. We denote X and Y data matrices of size $(T \times p)$ and $(T \times q)$, representing a realization of X and Y, respectively. Here, T is the length of time support of the four prices. Without loss of generality, we assume they are centred in column. Then, the sample covariance matrix of X, Y, and between X and Y are,

$$S_{XX} = \frac{1}{T-1}\mathsf{X}^t\mathsf{X},\; S_{YY} = \frac{1}{T-1}\mathsf{Y}^t\mathsf{Y} \text{ and } S_{XY} = \frac{1}{T-1}\mathsf{X}^t\mathsf{Y}.$$

Therefore, see also Robert and Escoufier (1976), the sample form of ρ_{XY} is,

$$RV_{XY} = \frac{Tr\left(S_{\mathbf{XY}}S_{\mathbf{YX}}\right)}{\sqrt{Tr\left(S^2_{\mathbf{XX}}\right)}\sqrt{Tr\left(S^2_{\mathbf{YY}}\right)}} \quad (3)$$

$$= \frac{Tr\left(\mathsf{X}^t\mathsf{Y}\mathsf{Y}^t\mathsf{X}\right)}{\sqrt{Tr\left(\mathsf{X}^t\mathsf{X}\right)^2}\sqrt{Tr\left(\mathsf{Y}^t\mathsf{Y}\right)^2}}.$$

This RV-coefficient is called Escoufier vector coefficient (EVC). It satisfies the following properties:

1. $0 \le RV_{XY} \le 1$. It equals 0 if each column of X is uncorrelated with all columns of Y, and it is equal to 1 if the eigensystem of S_{XX} is equal to that of S_{YY}.
2. If both stocks are represented by their closing prices only, then RV_{XY} is the squared of PCC.
3. If one stock represented by its closing price only, and the other one by its four prices, $RV_{XY} = \frac{1}{\sqrt{p}} R^2_{\mathbf{XY}}$ where R^2_{XY} is the R-squared when we regress the closing price of the first stock, with respect to all prices of the second.

Therefore, EVC generalises the notion of similarity of two random variables into two random vectors. In short, PCC measures the linear relationship of these two stocks each of which is represented by its closing price, whereas EVC measures the linear relationship of those stocks represented by their opening, highest, lowest, and closing prices. Hence, in that sense, EVC generalises PCC.

Four Prices-based Stocks Network

Network Construction

Basically, stocks network is constructed once we have measured the similarity among stocks. Based on the four prices discussed in the previous section, the notion of similarity of stocks i and j will be defined by using RV coefficient. Let S_{ij} be the covariance matrix of the four prices in both stocks i and j. Then, S_{ii} is the covariance matrix of the four prices in stock i. From (3), the RV coefficient of stocks i and j is,

$$RV_{ij} = \frac{Tr\left(\boldsymbol{S}_{ij}\boldsymbol{S}_{ji}\right)}{\sqrt{Tr\left(\boldsymbol{S}^2_{ii}\right)Tr\left(\boldsymbol{S}^2_{jj}\right)}} \quad (4)$$

Since EVC in (4) measures the linear relationship of two stocks each which is represented by its opening, highest, lowest and closing prices, in similar manner as in Mantegna (1999), it can be used as a similarity measure of those stocks in multivariate setting defined by the four prices.

By using EVC, let $\mathbf{V}$ be a matrix of size $(n \times n)$ with RV_{ij} as the element of its i-th row and j-th column. Then, $\mathbf{V}$ is a symmetric matrix with all diagonal elements equal to 1 and the off-diagonal elements are between 0 and 1. This matrix represents stocks network in multivariate setting of opening, highest, lowest, and closing prices. This vector correlations network generalises the notion of correlations network such as presented in Mantegna (1999), Mantegna and Stanley (2000), Bonanno *et al.* (2003, 2004), and Galazka (2011).

Network Analysis Methodology

Inspired by the work of Mantegna (1999), to analyse the vector correlations network, we define the function,

$$\delta_{ij} = \sqrt{2\left(1 - RV_{ij}\right)}. \quad (5)$$

It defines the distance between two stocks i and j. If $\mathbf{D}$ is a matrix of size $(n \times n)$ with δ_{ij} as the element of its i-th row and j-th column, then $\mathbf{D}$ represents an undirected weighted complete graph of n stocks in multivariate setting. Like in univariate setting, in what follows multivariate stocks network is analysed by using (i) the MST of $\mathbf{D}$ to filter the economic information contained in vector correlations network, and (ii) the four principal centrality measures to interpret the topological properties of each stock in the filtered network.

To obtain an MST of $\mathbf{D}$, we use Kruskal's algorithm introduced by Kruskal (1956) and developed by Djauhari and Gan (2013). Furthermore, to understand the topological properties of each stock, we use degree, betweenness, and closeness centrality measures introduced in Freeman (1979), and eigenvector centrality proposed in Bonacich (1987).

Example: Bursa Malaysia Case

In this section, 30 most capitalised stocks at Bursa Malaysia were studied. Data of stocks' opening, highest, lowest and closing prices were collected during the period of July 26, 2012 until December 31, 2012 (T = 106 trading days). Based on those four prices' data, first, EVC for each pair of stocks was calculated using (3) or more conveniently (4). Then, distance matrix D defined by (5) was calculated, and finally an MST of **D** was constructed. For that purpose, as suggested in the literature on econophysics, Kruskal's algorithm was used (Kruskal, 1956; Graham & Hell 1985; Huang *et al*., 2008; and Djauhari & Gan 2013).

MST of D represents a network consisting of n stocks each of which is in the form of multivariate time series of opening, highest, lowest and closing prices, and (n – 1) most relevant relationships in terms of EVC where each pair of stocks are connected in such a way that the sum of distances between two directly linked stocks is minimum. The result is presented in Figure 1.

The stocks (nodes) of Figure 1 are coloured according to their economic sectors defined in Industry Classification Benchmark (ICB); basic materials (*purple*), consumer goods (*green*), consumer services (*orange*), financials (*yellow*), health care (*cyan*), industrials (*gray*), oil and gas (*red*), telecommunication (*blue*) and utilities (*pink*).

At a glance we observe that CIMB is the most powerful stock in Bursa Malaysia according to the four popular centrality measures; degree, betweenness, closeness and eigenvector centralities. We also see that all telecommunication stocks (MAXIS –TM – DIGI – AXIATA) are linked to each other and directly linked to CIMB. Further details about the topological properties of each sector viewed as a multivariate time series is presented in Section 5.

To illustrate the advantages of multivariate approach and RV-coefficient, the results given in Figure 1 were compared with those given by univariate approach based on closing price only. By using closing price only, MST represents a network consisting of n stocks each of which is in the form of univariate time series of closing price, and (n – 1) most relevant relationships in terms of PCC where each pair of stocks are connected in such a way that the sum of distances between two directly linked stocks is minimum. Here, the distance between the two stocks was defined as $d_{ij} = \sqrt{2\left(1-c_{ij}\right)}$, where c_{ij} is the PCC of stocks i and j (Mantegna, 1999; Mantegna and Stanley, 2000). This distance d_{ij} is a special case of δ_{ij} in (5). An MST issued from closing price approach is presented in Figure 2.

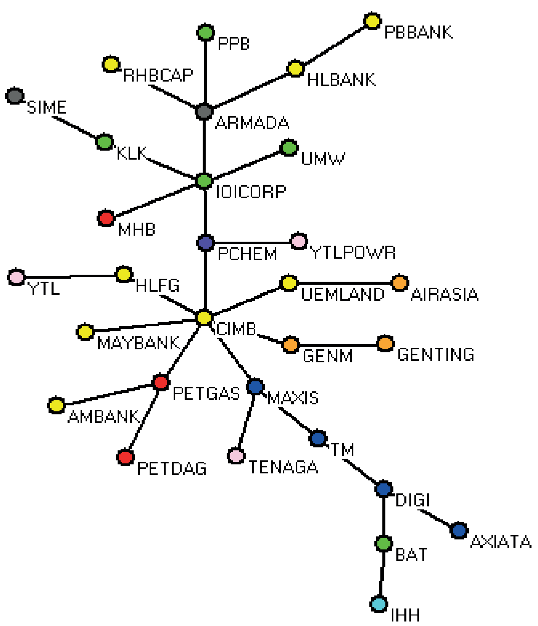

Figure 1. Multivariate MST of 30 stocks at Bursa Malaysia.

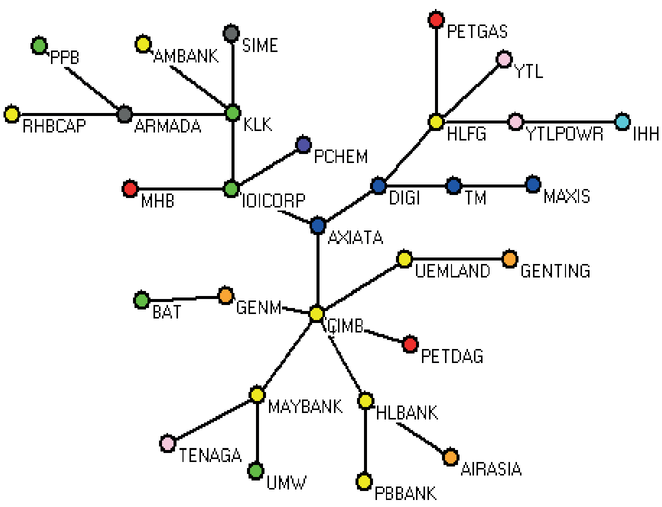

Figure 2. Univariate MST of Bursa Malaysia 30 stocks.

This MST is totally different from that in Figure 1. Just to mention their difference, although CIMB is still the most powerful stock according to the four popular centrality measures and also all telecommunication stocks (MAXIS –TM – DIGI – AXIATA) are linked to each other and directly to CIMB, their degrees are different.

DISCUSSION

According to the 30 most capitalised stocks, during the period of July 26, 2012 until December 31, 2012, univariate approach of Bursa Malaysia stocks market analysis as represented in Figure 2 gives different topological properties compared to those given by multivariate approach in Figure 1. As we will show, the real situation of stocks market in Bursa Malaysia, among those 30 stocks during the period of study, can better be figured out by the latter than the former. See, for example, the connections of PETDAG and PETGAS, and also of GENM and GENTING in Figure 1. These connections, which show the phenomenon of social embeddedness (Halinen & Tornroos 1998), can only be detected by multivariate approach and not by univariate one. We also learnt from company's website that PETDAG and PETGAS are led by two common directors. On the other hand, GENTING is the management company and investment holding of Genting Group where GENM is its subsidiary. The board of directors in both companies is chaired by the same person and the same deputy chairman.

Further properties, for example, are the number of leaves and the diameter of MST. Based on the graph theory, a leaf is a node of degree one, while diameter is the longest path between two leaves. Therefore, according to the four centrality measures, a leaf represents the worst stock while diameter shows how far at most the information from one worst stock could reach the other one. According to multivariate approach, the number of leaves is 16 and diameter is 10 (from PBBANK to IHH). On the other hand, based on univariate approach, the number of leaves is 17 and diameter is 8 (from PPB or RHBCAP to IHH). This does make sense as the information in multivariate setting is more complex than in the univariate one.

From Figures 1 and 2 we learn that AIRASIA, AMBANK, GENTING, IHH, MHB, PBBANK, PETDAG, PPB, RHBCAP, SIME, TENAGA, UMW, and YLT are considered as the worst in Bursa Malaysia by both approaches. We are then more confident to say so as an evident from MST. On the other hand, Table 1 shows the stocks considered as worst by multivariate approach but not by univariate one, and vice versa.

Table 1 shows that, according to multivariate approach which is more comprehensive than univariate, AXIATA, MAYBANK, and YLTPOWER should be considered as the worst stocks even though they have degree 2 or 3 in univariate approach. A more general comparison of the two approaches is presented in Table 2 in terms of the four principal centrality measures; degree centrality (D), betweenness centrality (B), closeness centrality (C), and eigenvector centrality (E). Furthemore, the details of the first three centrality measures can be found in Freeman (1979), and the last one in Bonacich (1987). We can also find all those measures in the works of many such as Borgatti (1995), Borgatti (2005), Borgatti and Everett (2006), and Newman (2008).

From Table 2 we learn, for example, that zero betweenness refers to the stock of degree one which represents leaf. If we exclude those leaves in further analysis, then we arrive at the 10 most powerful stocks as presented in Table 3.

We see a different conclusion given by the two approaches. Since multivariate approach is more comprehensive, we recommend the policy makers and stock players to use the results issued by that approach.

CONCLUDING REMARKS

The EVC generalises PCC into multivariate setting. In this setting, stock is represented as a multivariate time series of its opening, highest, lowest, and closing prices, whereas in univariate setting, stock is represented by its closing price only. EVC allows us to define the similarity among stocks in multivariate case in similar manner as in univariate case

TABLE 1. THE WORST STOCKS ACCORDING TO UNIVARIATE AND MULTIVARIATE APPROACHES.

Stock's name	Degree	
	Univariate approach	Multivariate approach
AXIATA	3	1
BAT	1	2
MAYBANK	3	1
MAXIS	1	3
PCHEM	1	3
PETGAS	1	3
YLTPOWER	2	1

TABLE 2. CENTRALITY MEASURES OF EACH STOCK.

Stock's name	Multivariate setting				Univariate setting			
	D	B	C	E	D	B	C	E
AIRASIA	0.035	0	0.227	0.079	0.035	0	0.220	0.098
AMBANK	0.035	0	0.230	0.091	0.035	0	0.220	0.057
ARMADA	0.138	0.259	0.264	0.131	0.103	0.136	0.227	0.077
AXIATA	0.035	0	0.181	0.017	0.103	0.680	0.392	0.368
BAT	0.069	0.069	0.184	0.019	0.035	0	0.216	0.083
CIMB	0.241	0.771	0.392	0.603	0.207	0.603	0.363	0.567
DIGI	0.103	0.197	0.220	0.050	0.103	0.404	0.330	0.206
GENM	0.069	0.069	0.290	0.232	0.069	0.069	0.274	0.233
GENTING	0.035	0	0.227	0.079	0.035	0	0.216	0.083
HLBANK	0.069	0.069	0.213	0.050	0.103	0.136	0.279	0.273
HLFG	0.069	0.069	0.290	0.232	0.138	0.259	0.269	0.124
IHH	0.035	0	0.156	0.007	0.035	0	0.179	0.018
IOICORP	0.172	0.505	0.322	0.244	0.138	0.446	0.337	0.254
KLK	0.069	0.069	0.250	0.094	0.138	0.313	0.279	0.159
MAXIS	0.103	0.352	0.322	0.275	0.035	0	0.204	0.030
MAYBANK	0.035	0	0.284	0.205	0.103	0.136	0.279	0.273
MHB	0.035	0	0.246	0.083	0.035	0	0.254	0.091
PBBANK	0.035	0	0.177	0.017	0.035	0	0.220	0.098
PCHEM	0.103	0.512	0.363	0.326	0.035	0	0.254	0.091
PETDAG	0.035	0	0.230	0.091	0.035	0	0.269	0.203
PETGAS	0.103	0.136	0.296	0.267	0.035	0	0.213	0.044
PPB	0.035	0	0.210	0.044	0.035	0	0.186	0.027
RHBCAP	0.035	0	0.210	0.044	0.035	0	0.186	0.027
SIME	0.035	0	0.201	0.032	0.035	0	0.220	0.057
TENAGA	0.035	0	0.246	0.094	0.035	0	0.220	0.098
TM	0.069	0.246	0.264	0.111	0.069	0.069	0.254	0.085
UEMLAND	0.069	0.069	0.290	0.232	0.069	0.069	0.274	0.233
UMW	0.035	0	0.246	0.083	0.035	0	0.220	0.098
YTL	0.035	0	0.227	0.079	0.035	0	0.213	0.044
YTLPOWR	0.035	0	0.269	0.111	0.069	0.069	0.216	0.051

TABLE 3. TEN MOST POWERFUL STOCKS. (Ordered from the most to less powerful).

No.	Multivariate setting	Univariate setting
1	CIMB	CIMB
2	PCHEM	HLFG
3	IOICORP	IOICORP
4	MAXIS	KLK
5	ARMADA	AXIATA
6	PETGAS	DIGI
7	GENM	HLBANK
8	TM	MAYBANK
9	HLFG	GENM
10	UEMLAND	UEMLAND

suggested by Mantegna (1999) and Mantegna and Stanley (2000). Then, we can construct a multivariate stocks network and filter multivariate economic information by using MST. Since multivariate approach is closer to the real activity than univariate one, the results are more reliable. The topological properties of stocks derived from Figure 1 are then more reliable than those issued from Figure 2. It is also important to note that the phenomenon of social embeddedness can only be detected by multivariate approach and not by univariate one. Thus, the real situation of Bursa Malaysia's 30 most capitalised stocks, during the period of study, can better be figured out by the former than the latter.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the Editor and anonymous referees for their comments and suggestions that led to the final version of this presentation. Special thanks goes to the Institute for Mathematical Research (INSPEM), Universiti Putra Malaysia, for providing research facilities as well as to Universiti Teknologi Malaysia, where the second author is doing her PhD programme.

Date of submission: April 2013
Date of acceptance: January 2014

REFERENCES

Borgatti, SP 1995, 'Centrality and AIDS', *Connections*, vol. 18, no. 1, pp. 112–115, viewed 17 August 2013, http://www.analytictech.com/networks/centaids.htm

Borgatti, SP 2005, 'Centrality and network flows', *Social Networks*, vol. 27, no. 1, pp. 55–71.

Borgatti, SP & Everett, MG 2006, 'A graph-theoretic perspective on centrality', Social Networks, vol. 28, no. 4 pp. 466–484.

Bonacich, P 1987, 'Power and Centrality: A Family of Measures', *American Journal of Sociology*, vol. 92, no. 5, pp. 1170–1182.

Bonanno, G et al. 2003, 'Topology of correlation-based minimal spanning trees in real and model markets', *Physical Review E*, vol. 68, no. 4, pp. 046130–046133.

Bonanno, G et al. 2004, 'Networks of equities in financial markets', *European Physics Journal B*, vol. 38, no. 2, pp. 363–371.

Brida, JG & Risso, WA 2008, 'Multidimensional minimal spanning tree: the Dow Jones case', *Physica A*, vol. 387, no. 21, pp. 5205–5210.

Coelho, R *et al*. 2007, 'The evolution of interdependence in world equity markets – Evidence from minimum spanning tree', *Physica A*, vol. 376, no. 15, pp. 455–466.

Djauhari, MA 2011, 'Geometric interpretation of vector variance', *Matematika*, vol. 27, no. 1, pp. 51–57.

Djauhari, MA 2012, 'A robust filter in stock networks analysis', *Physica A*, vol. 391, no. 20, pp. 5049–5057.

Djauhari, MA & Gan, SL 2013, 'Minimal spanning tree problem in stock network analysis: an efficient algorithm', *Physica A*, vol. 392, no. 9, pp. 2226–2234.

Escoufier, Y 1973, 'Le traitement des variables vectorielles', *Biometrics*, vol. 29, no. 4, pp. 751–760.

Escoufier, Y 1976, 'Opérateur associé à un tableau de données', Annales de l'INSEE, 22–23, pp. 751–760, Paris.

Freeman, LC 1979, 'Centrality in social networks: I. conceptual clarification', *Social Networks*, vol. 1, pp. 215–239.

Galazka, M 2011, 'Characteristics of Polish stock market correlations', *International Review of Financial Analysis*, vol. 20, pp. 1–5.

Garas, A & Argyrakis, P 2007, 'Correlation study of the Athens stock exchange', *Physica A*, vol. 380, no. 1-2, pp. 399-410.

Graham, RL & Hell, P 1985, 'On the history of the minimum spanning tree problem', *Annales of Historical Computation*, vol. 7, no. 1, pp. 43–57.

Halinen, A & Tornroos, J 1998, 'The role of embeddedness in the evolution of business network', *Scandinavian Journal of Management*, vol. 14, no. 3, pp. 187–205.

Huang, F, Gao, P & Wang, Y 2008, 'Comparison of Prim and Kruskal on Shanghai and Shenzhen 300 Index Hierarchical Structure Tree', in *International Conference on Web Information Systems and Mining*, pp. 237–241, Shanghai.

Jang, W, Lee, J & Chang, W 2011, 'Currency crises and evolution of foreign exchange market: Evidence from minimum spanning tree', *Physica A*, vol. 390, no. 4, pp. 707–718.

Jung, W-S *et al*. 2006, 'Characteristics of Korean stock market correlation', *Physica A*, vol. 361, no. 1, pp. 263–271.

Kruskal, JB 1956, 'On the shortest spanning subtree of a graph and travelling salesman problem', *American Mathematical Society*, vol. 7, no. 1, pp. 48–50.

Mantegna, RN 1999, 'Hierarchical structure in financial markets', *European Physics Journal B*, vol. 11, no. 1, pp. 193-197.

Mantegna, RN & Stanley, HE 2000, *An introduction to econophysics: Correlations and complexity in finance*. Cambridge, UK: Cambridge University Press

Micciche, S *et al*. 2003, 'Degree stability of minimum spanning tree of price return and volatility', *Physica A*, vol. 342, no. 1–2, pp. 66–73.

Newman, MEJ 2008, 'Mathematics of Networks. The New Palgrave Encyclopedia of Economics', 2nd edn, eds Blume, L.E. & Durlauf, S.N., Palgrave Macmillan, Basingstoke.

Newman, MEJ 2011, 'Complex system: a survey', *American Journal of Physics*, vol. 79, no. 8, pp. 800–810.

Onnela, JP 2006, 'Complex system in the study of financial and social systems', Ph.D. Thesis, Helsinki University of

Technology, Finland, viewed July 22, 2014, https://aaltodoc.aalto.fi/handle/123456789/5/discover

Robert, P & Escoufier, Y 1976, 'A unifying tool for linear multivariate statistical methods: The RV coefficients. *Applied Statistics*, Vol. 25, no. 3, pp. 257–265.

Sieczka, P & Holyst, J 2009, 'Correlations in commodity markets', *Physica A*, vol. 388, no. 8, pp. 1621–1630.

Tumminello, M *et al*. 2005, 'A tool for filtering information in complex systems', *Proceedings of the National Academy of Science*, vol. 102, no. 30, pp. 10421–10426.

Wang, G-J *et al.* 2012, 'Similarity measure and topology evolution of foreign exchange markets using dynamic time warping method: Evidence from minimal spanning tree', *Physica A*, vol. 391, no. 16, pp. 4136–4146.

West, DB 2001, *Introduction to graph theory*. 2nd edn, Prentice Hall, New Jersey.

Wilmott, P 2007, *Paul Wilmott Introduces Quantitative Finance*, John Wiley & Sons, England.

Yamashita, Y & Yodahisa, H 2012, 'Similarity Measure and Clustering Algorithm for Candlestick Valued Data', *Joint Meeting of Japanese and Italian Classification Societies* (JCS-CLADAG).

Zhang, Y, *et al*. 2011, 'Will the US economy recover in 2010? A minimal spanning tree study', *Physica A*, vol. 390, no. 11, pp. 2020–2050.

Zhuang, R, Hu, B & Ye, Z 2008, 'Minimal Spanning Tree for Shanghai-Shenzhen 300 Stock Index', in *Proceedings IEEE World Congress on Computational Intelligence*, pp.1417–1424, Hong Kong.

Appendix: List of Stocks and their Company's Name.

No.	Stock	Company's Name
1	AIRASIA	AirAsia
2	AMBANK	AMMB Holdings
3	ARMADA	Bumi Armada
4	AXIATA	Axiata Group
5	BAT	British American Tobacco
6	CIMB	CIMB Group Holdings
7	DIGI	Digi. Com
8	GENM	Genting Malaysia/ Resorts World
9	GENTING	Genting
10	HLBANK	Hong Leong Bank
11	HLFG	Hong Leong Financial Group
12	IHH	IHH Healthcare
13	IOICORP	IOI Corporation
14	KLK	Kuala Lumpur Kepong
15	MAXIS	Maxis
16	MAYBANK	Malayan Banking
17	MHB	Malaysia Marine and Heavy Engineering
18	PBBANK	Public Bank
19	PCHEM	PETRONAS Chemicals Group
20	PETDAG	PETRONAS Dagangan
21	PETGAS	PETRONAS Gas
22	PPB	PPB Group
23	RHBCAP	RHB Capital
24	SIME	Sime Darby
25	TENAGA	Tenaga Nasional
26	TM	Telekom Malaysia
27	UEMLAND	UEM Land
28	UMW	UMW Holdings
29	YTL	YTL Corporation
30	YTLPOWR	YTL Power International