Directed network of Shariah-compliant stock in Bursa Malaysia

Muhammad Hasannudin Musa¹ and Fatimah Abdul Razak²
¹,²Jabatan Sains Matematik, Fakulti Sains dan Teknologi, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia

E-mail: ¹muhammadhasannudin91@gmail.com , ²fatima84@ukm.edu.my

Abstract. Stock network is a type of financial network based on stock price data used for analysing stock market dynamics. In this paper, a directed stock network is developed. This model was built using 480 shariah-compliant stocks traded in Bursa Malaysia from the year 2016 until 2018. Transfer Entropy was used as a measuring tool to build the stock network. Different networks are built and evaluated using network analysis methods. To determine the important stocks in the networks, centrality measures are applied such as degree centrality. The findings showed that Borneo Oil Berhad (BRNL) is the most influential and important stock among the 480 shariah-compliant stock in the Bursa Malaysia.

1. Introduction
In network science, network theory is the study of graph as a representation between discrete object. Graph theory is a part of network theory that can be defined as graph which are mathematical structures used to model a relation between objects. A graph in this context is made up of vertices or nodes which are connected by edges. Graph with a large number of nodes and edges can be defined as a complex network [1]. Complex network analysis in recent years has become the most important tool to analyse and investigate the problem in world of research [2]. Network analysis is a popular tool to describe the characteristics and behaviour of complex networks. Many researchers has been using complex network analysis to study complex networks in biology [3], social sciences [4], engineering [5] and finance [2]. This study focusing on the important of stock networks for financial system. There are many analysis and information that can be extracted from the network such as the most influential stock in the market. Usually in network analysis, degree centrality commonly used to identify the important nodes within a network. In this study also we used stock price data from Malaysian stock market that known as Bursa Malaysia.

From the stock listed in main market of Bursa Malaysia, we only consider the shariah-compliant stocks. Shariah-compliant stock were chosen due to the availability of the data and to study the topological structure of shariah-compliant stocks traded in Bursa Malaysia. Moreover, many researchers used conventional stocks in their studies. Shariah-compliant stock are ordinary stock that listed in Bursa which have been classified as shariah permissible for investment based on the company’s compliance with Shariah principles in terms of its primary business and investment activities [6]. Securities Commission Malaysia (SC) is a legal organization with responsibilities of regulating and assessing Malaysia’s capital markets by providing a list of securities that has been examined as shariah compliance [6]. To determine the Shariah-compliant status, the stock have to obey some Shariah criteria for the
stock to be considered as shariah-compliant stock. The company activities must be free from interest, doubtful transaction or uncertainty and forbidden activities in Islam such as alcohol trading and gambling. The shariah principles concentrated on the core business activities of the companies. If the company fulfilling all the criteria then the stock of the company considered as a shariah-compliant stock [6].

Stock network is a graph where each nodes representing the stock and the relationship between stocks represented as an edge. The stock network has been used to observe and analyze the stock market behavior [7]. To build network, there are many methods commonly used such as Minimum Spanning Tree (MST) [2,6,8], the Dynamic Spanning Tree (DST), the Planar Maximally Filtered Graph (PMFG) [9–11] and the Coefficient Threshold Method [8,12]. Among these methods, the MST approach is the simplest network and the most well known to simplify the complexity of the stock network. As an example, MST was used to develop New York stock network as well as the capitalized stock trade in the US equity markets [8,13]. For the Dynamic Spanning Tree (DST) is also common method to study stock networks, but there are drawback that DST shrinks over time [14]. For PMFG method, Tumminello et al. [9] used PMFG to build a network of the portfolio of the 300 most capitalized stocks trade at New York Stock Exchange during the time period 2001 until 2003. The PMFG method was proposed in order to retain more information about network. However, both MST and PMFG suffer loss of information as edges of high correlation are often removed while edges of low correlation are retained because of their topological conditions fitting the topological reduction criteria [15]. In this paper, in order to retain all information of the internal structure of stock network, we used the threshold method to build stock network using Shariah-compliant stocks.

Huang et al. [12] used a threshold method to build a stock correlation network using stock price data which analysed the network properties and topological stability of stock market. Boginski et al. [16] also used correlation threshold method build a stock market network by analyse daily fluctuation of 6546 financial instruments in the US stock markets during 500 consecutive trading days in 2000-2002. Statistical analysis is being conducted to show that network can be described by the power-law model.

The Pearson Correlation Coefficient is frequently used to capture the relationship between stocks [17]. However, Pearson correlation coefficient can only measure the linear relationships [2]. Khoojine et al. [18] used a nonlinear measure which is mutual information (MI) to measure the relationship between Chinese Stock Market during the turbulence in 2015-2016. Chinese Stock market network which is built using MI also being analysed. MI is a statistical tool to measure dependence between two random variables or systems [16,17]. Briefly MI is a measurement tool to measure the amount of shared information between two variables [19]. It also useful tool to quantify the signal interaction [19]. It can be used to identify the relationship between data sets that are not detected by the linear measure of correlation [20]. MI only detect the interdependence between two system which is stock in this case but MI cannot indicate the way of information flowing. Therefore, in this work, we propose to develop a directed stock network by using Transfer Entropy (TE) using the price data from Bursa Malaysia. TE is a special case or conditional MI. Two measure used to quantify the signal interaction or shared information between two variables are MI and TE [19]. MI measures the amount of shared information between two variables while, TE measures the amount of information transfer from one variable two another [19]. Korbel et al. [21] used TE to analyse information flows between communities of financial market. So, from this idea, we want to build a directed stock network of shariah-compliant stock in Bursa Malaysia using TE measure that demonstrate relationship between stocks. The rest of this paper is organized as follows. In section 2, the computation of TE for developing stock network is defined. In section 3, we develop stock networks using TE and finally discuss and analyse of this network using network analysis. For this purpose, the open-source social network analysis software “Gephi” by Bastian et al. [22] is used to depict the stock network.
2. Developing directed stock network

2.1. Transfer Entropy

In this section, we will describe the concept of Transfer Entropy (TE) which is a useful tool in information theory. Before we explain the concept of TE in details, we have to know the basic measure for TE and MI which is Shannon Entropy. Mutual information associates with Shannon’s entropy, which is a measure of the unpredictability of a random variable [18]. The concept of Shannon entropy was introduced by Claude Shannon in his seminal paper in 1948 [23]. The Shannon entropy measure the expected uncertainty in random variable and signal [19]. According to Shannon [23], we suppose a system $X$, and it contains a series of possible events whose probabilities of occurrence is $P(X = x)$ and the Shannon entropy formula given by the expression:

$$H(X = x) = - \sum_{x \in X} P(X = x) \log P(X = x)$$ (1)

Then, Shannon entropy formula given by the expression:

$$H(X = x, Y = y) = - \sum_{x \in X} \sum_{y \in Y} P(X = x, Y = y) \log P(X = x, Y = y)$$ (2)

where $P(X = x)$ is a probability of $X$ and $P(X = x, Y = y)$ is the joint probability distribution of $(X, Y)$.

If the two system are independent, then the Shannon entropy can be expressed as:

$$H(X = x, Y = y) = - \sum_{x \in X} \sum_{y \in Y} P(X = x)P(Y = y) \log P(X = x)P(Y = y)$$ (3)

According to the Shannon entropy, mutual information for two discrete random variables, $X$ and $Y$ can be written as:

$$I(X = x, Y = y) = \sum_{x \in X} \sum_{y \in Y} P(X = x, Y = y) \log \frac{P(X = x, Y = y)}{P(X = x)P(Y = y)}$$ (4)

As we mentioned before, a variant of conditional Mutual Information namely Transfer Entropy was first defined by Schreiber in [24,25]. With some modifications, we consider the joint probability and conditional the Transfer Entropy can be expressed as:

$$T_{Y \rightarrow X} = \sum_{x_{n+1}} \sum_{y_{n}} \sum_{x_{n}} P(X_{n+1} = x_{n+1}, X_{n} = x_{n}, Y_{n} = y_{n}) \log \frac{P(X_{n+1} = x_{n+1} | X_{n} = x_{n}, Y_{n} = y_{n})}{P(X_{n} = x_{n} | X_{n} = x_{n})}$$ (5)

where $x_{n}$ is the value of variable $X_{n}$ and $y_{n}$ is the value of variable $Y_{n}$ and $x_{n+1}$ is the value of variable $X$ at time $n+1$.

From all the definitions as mentioned, we need to do trivial part of this measure which is calculation of probability distribution to estimate the Transfer Entropy. TE is a directional measure, in that $X$ may provide information to $Y$ but $Y$ may not inform $X$ [19]. First, we used the numerical method to compute the Transfer Entropy of stock price data sets. In this study we used the daily closing price data the transformed into stock return data. Considering a network on $n$ stocks with $P_{i,t}$ denoted as the closing stock price of a company $i$ ($i=1,...,N$) at time $t$ and $R_{i,t}$ as the log-return of stock $i$ at day $t$, the logarithmic of the stock return is calculated as [26]:

$$R_{i,t} = \ln \frac{P_{i,t}}{P_{i,t-1}}, (t = 2,...,d; i = 1,2,...,n)$$ (6)

Then we have $R_{i,t}$ stock return data set. The next step is letting $C$ be a correlation matrix of $N\times N$ matrix, where $N$ is the number of stocks. Before computing the Transfer entropy, we have to normalize and discretize the data of stock price. The normalization of data due to the price different for all 480
stock companies is too large. The discretization on data using a binning method. We find maximum value of the stock price and minimum value to build the 10 divided intervals. Then we sort the data according to the range value. Then we compute the Transfer Entropy and store the value of TE in matrix $C$. The matrix of TE is asymmetry so we do some data management to determine the causal relation and to detect the direction of information in order to build directed stock network. Based on the calculation of TE, which value bigger will dominate the direction between pair of stock relation.

2.2. Constructing the network
The main idea of constructing the stock network is as follow. Let the set of stocks represent the set of nodes of the network. The value of Transfer Entropy as weight for directed network stock. Let graph $G = (V, E)$ represent the stock network, where $V$ and $E$ are the set of nodes and edges respectively. $E$ is defined as:

$$E = \begin{cases} 
e e_{ij}, & i \ne j \text{ and } c_{ij} \geq \theta \\ e_{ij} = 0, & i = j \end{cases} \quad (7)$$

and $c_{ij}$ is the element of connection or relation matrix for TE measure and $\theta$ is certain threshold TE value. We know that the value TE for $T_{Y \rightarrow X}$ and $T_{X \rightarrow Y}$ will be stored in correlation matrix. After some filtering and data management the bigger value will stay to indicate the cause of direction. Finally, from the value in correlation matrix, open-source network analysis and visualization software “Gephi” [22] is used to visualize the stock network.

3. Results and discussion
There are more than 1000 companies listed stocks in Bursa Malaysia but in this study we only used 480 shariah-compliant stock listed in Bursa Malaysia. The data is from 1st January 2016 until 4th January 2019 equal to 776 trading days. The period is chosen due to the availability of the data. In Figure 1, we show the TE distribution for log-return data set that have has been used in this study. As can we see, the result in Figure 1 resembles the shape of log-normal distribution. Based on this shape, it is easy to see that the increment of threshold value $\theta$ increase will make the number of edges in the stock network decreases.

![Figure 1. Frequency distribution of the transfer entropy value](image-url)

Then we now proceed to analyse the stock network for different values of $\theta$, as mentioned in the previous section. We will show certain threshold $\theta$ that has been analysed. Social network analysis is the process or method to investigate the social structures through networks [7]. There are many
centralities measure are quite popular as a measuring tool to identify most important nodes within a network such as degree centrality, closeness centrality, betweenness centrality and eigenvector centrality [7]. In this study we are focusing in using degree centrality to determine the important nodes according to our definitions. Degree of every nodes is one of the most crucial fundamental in network analysis. It represents number of nodes connected to a specific node. For directed network, degree centrality usually define two separate measures of degree centrality, namely in-degree and out-degree. Therefore, in our definition, out-degree centrality measure is used to detect the nodes that have a large out-degree value. In that sense, we presume that nodes which have a large out-degree value are the important for cause effect. We tried to look the effect of filtering the threshold on the important nodes. In each of the following figures, according to the centrality metric used, most important nodes (stocks) are shown with larger size.

Figure 2 presenting stock network with threshold value $\theta > 0.06$, where out-degree centrality is used. The stock network in Figure 1 show that Borneo Oil Berhad (BRNL) are the most important nodes lie in the center of the network. BRNL stock are the stock that influencing other stock in sense of price fluctuation.

Figure 2. Out-degree centrality, threshold $\theta > 0.06$.

In the following figures, the density of stock network is reduced for more meaningful interpretation. Stock network in Figure 3 was created using the out-degree centrality. Based on the observation, BRNL is the most important stock in the network. BRNL are the company in the sector of industrial product and services.
In Figure 4, we keep increasing $\theta$ by 0.01 and show that most important nodes according to out-degree centrality. For this stock network at threshold $\theta > 0.08$ the BRNL still dominate to become the most important node. In this stock network, the density of this network also decreasing as we increase the $\theta$. In Figure 4 also shown that reduction of edges were not affecting the status of BRNL stock as nodes with highest degree.

For $\theta > 0.09$, only 17 stocks were found to be connected in this stock network. In this figure, we can clearly the BRNL stock became the most important stock and Omesti Berhad (OMES) from technology sector also a stock with highest in-degree centrality.
4. Conclusion
In this paper, our motivation was to propose the application of nonlinear measure which is TE to measure the relation between stock instead of linear measure which is Pearson Correlation Coefficient. The limitation of Pearson Correlation Coefficient to indicate the direction of information flowing motivated us to build a directed stock network using TE. Correlation matrix was created to store the value of TE for each entry to form a table. The tables became an adjacency matrix which can be converted to form a stock network using open-source software which is Gephi. The tool from network analysis which is degree centrality were used to determine the most important nodes in the stock network. From the analysis conducted and according to our definition, BRNL are the most important and influential stock in Malaysian market. In future studies, we suggest analysing stock network using more centralities measure such as closeness centrality, betweenness centrality, eigenvector centrality and clustering coefficient to extract more information in stock network in Malaysian market.

Acknowledgements
The work was financially supported by the research grants FRGS/1/2020/STG06/UKM/02/8.

References
[1] Albert R and Barabási A L 2002 Statistical mechanics of complex networks Rev. Mod. Phys. 74 47–97
[2] Guo X, Zhang H and Tian T 2018 Development of stock correlation networks using mutual information and financial big data PLoS One 13 1–16
[3] Ma C Y and Liao C S 2020 A review of protein–protein interaction network alignment: From pathway comparison to global alignment Comput. Struct. Biotechnol. J. 18 2647–56
[4] Kim J and Hastak M 2018 Social network analysis: Characteristics of online social networks after a disaster Int. J. Inf. Manage. 38 86–96
[5] Qian Y S, Wang M, Kang H X, Zeng J W and Liu Y F 2012 Study on the road network connectivity reliability of valley city based on complex network Math. Probl. Eng. 2012
[6] Mahamood F N A, Bahaludin H and Abdullah M H 2019 A Network Analysis of Shariah-Compliant Stocks across Global Financial Crisis: A Case of Malaysia Mod. Appl. Sci. 13 80
[7] Dimitrios K and Vasileios O 2015 A Network Analysis of the Greek Stock Market Procedia Econ. Financ. 13 80
[8] Bonanno G, Caldarelli G, Lillo F and Mantegna R N 2003 Topology of correlation-based minimal spanning trees in real and model markets Phys. Rev. E - Stat. Physics, Plasmas, Fluids, Relat. Interdiscip. Top. 68 4–7
[9] Tumminello M, Matteo T Di, Aste T and Mantegna R N 2007 Correlation based networks of equity returns sampled at different Eur. Phys. J. B 217 209–17
[10] Tumminello M, Aste T, Matteo T Di and Mantegna R N 2005 A tool for filtering information in complex systems 102 10421–6
[11] Yan X, Xie C, Wang G and Management I 2015 Stock market network’s topological stability: Evidence from planar maximally filtered graph and minimal spanning tree 29 1–19
[12] Huang W Q, Zhuang X T and Yao S 2009 A network analysis of the Chinese stock market Phys. A Stat. Mech. its Appl. 388 2956–64
[13] Micciché S, Bonanno G, Lillo F and Mantegna R N 2003 Degree stability of a minimum spanning tree of price return and volatility Phys. A Stat. Mech. its Appl. 324 66–73
[14] Sensoy A and Tabak B M 2014 Dynamic spanning trees in stock market networks: Physica A 414 387–402
[15] Tse C K, Liu J and Lau F C M 2010 A network perspective of the stock market J. Empir. Financ. 17 659–67
[16] Boginski V, Butenko S and Pardalos P M 2005 Statistical analysis of financial networks 48 431–43
[17] Omela J P, Kaski K and Kertész J 2004 Clustering and information in correlation based financial networks Eur. Phys. J. B 38 353–62
[18] Khoojine A S and Han D 2019 Network analysis of the Chinese stock market during the turbulence of 2015–2016 using log-returns, volumes and mutual information Phys. A Stat. Mech. its Appl. 523 1091–109
[19] Sendrowski A, Sadid K, Meselhe E, Wagner W, Mohrig D and Passalacqua P 2018 Transfer entropy as a tool for hydrodynamic model validation Entropy 20 1–24
[20] Lopes F, Pijn J P and Bocciinga P 1989 Interdependence of EEG Signals: Linear vs. Nonlinear Associations and the Significance of Time Delays and Phase Shifts 2 9–18
[21] Korbel J, Jiang X and Zheng B 2019 Transfer entropy between communities in complex financial networks Entropy 21 1–13
[22] Bastian M, Heymann S and Jacomy M 2009 Gephi: An open source software for exploring and manipulating networks. BT - International AAAI Conference on Weblogs and Social Int. AAAI Conf. Weblogs Soc. Media 361–2
[23] Shannon C E 1948 A Mathematical Theory of Communication Bell Syst. Tech. J. 27 623–56
[24] Schreiber T 2000 Measuring information transfer Phys. Rev. Lett. 85 461–4
[25] Abdul Razak F and Jensen H J 2014 Quantifying “causality” in complex systems: Understanding transfer entropy PLoS One 9 1–14
[26] Mantegna R N 1999 Hierarchical structure in financial markets Eur. Phys. J. B 11 193–7