Information transfer network of global market indices

Yup Kim∗, Jinho Kim, Soon-Hyung Yook∗

Department of Physics and Research Institute for Basic Sciences, Kyung Hee University, Seoul 130-701, Republic of Korea

Abstract

We study the topological properties of the information transfer networks (ITN) of the global financial market indices for six different periods. ITN is a directed weighted network, in which the direction and weight are determined by the transfer entropy between market indices. By applying the threshold method, it is found that ITN undergoes a crossover from the complete graph to a small-world (SW) network. SW regime of ITN for a global crisis is found to be much more enhanced than that for ordinary periods. Furthermore, when ITN is in SW regime, the average clustering coefficient is found to be synchronized with average volatility of markets. We also compare the results with the topological properties of correlation networks.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Due to the availability of large volumes of data amenable to computational analysis, financial systems have been increasingly studied in statistical physics as a part of complex systems in which human activity causes rich interesting phenomena [1,2]. From the empirical studies on financial market data, many interesting universal behaviors which govern the dynamical properties of financial systems have been found [3–5]. Most of those empirical data describe the time evolution of market properties such as price or index. In general the time evolutions of such quantities are irregular and unpredictable. Thus, extracting an essential regularity and universal behavior is not a trivial task. Especially, inferring causal relationship between different time series is very important if the underlying mechanisms are not fully understood. The analysis of complex time series plays an important role in many disciplines such as physics, neuro-science, seismology, and climate science as well as economics [6–9]. In this sense, developing theoretical methods or an efficient way to analyze such time series becomes more important to understand essential properties in various systems.

Studies on financial systems were focused on finding universal properties of financial systems such as (i) the fat-tailed distribution of return, (ii) absence of correlation in the time-series of return, and (iii) long-term correlation in volatility. Such universal properties are relatively well studied, and many stochastic models and agent based models have been

* Corresponding authors.
E-mail addresses: ykim@khu.ac.kr (Y. Kim), syook@khu.ac.kr (S.-H. Yook).

http://dx.doi.org/10.1016/j.physa.2015.02.081
0378-4371/© 2015 Elsevier B.V. All rights reserved.
introduced to uncover the origin of such universal behaviors [1,2]. Especially, understanding the peculiar behavior of volatility is particularly important because the volatility is closely related to the significant events in financial markets such as bubbles and crashes [10–13]. Thus, quantifying such significant economic events is very crucial not only practically but also theoretically. Moreover, the volatility is also known to be closely related to the amount of information that has arrived in the market at a given time and plays an important role in the modeling of the financial systems [1,14,15]. However, the information flow in the financial system during such significant financial events is not yet well understood because of their intrinsic complexities.

Recently, there have been several attempts to investigate the relationship between interacting components in various financial systems using graph theoretical approaches [16–21]. The graph theoretical approaches have provided simple intuitive frameworks to study the underlying properties of financial systems. Since the effect of interaction topology on the observed universal behavior in financial systems is shown to be crucial [22,23], it is important to find the structure of interactions and its topological properties. However, these studies are in general based on the assumption that the interaction between each component of the system is symmetric. For example, since the correlation matrix which is widely used in the analysis of domestic markets with random matrix theory is usually symmetric [24–26], some extensions of such analysis to the studies on global market indices assumed the symmetric interaction [16,17,20]. Thus, the network representation based on such symmetric interaction matrix is generally considered as an undirected network. The assumption of symmetry might give a correct idea to understand the various characteristics observed in financial systems only when the time scale of information flow is smaller than that of the measurement. On the other hand, if the time scale of information flow is large, then the asymmetric relationship between components should be considered due to the causality. The causality is usually measured by the time-delayed mutual information [27] and the delayed cross-correlation [28,29]. But it was shown that the mutual information does not explicitly distinguish the actually exchanged information due to a common history or input signal. In order to improve the measurement of information transfer by excluding those influences, the transfer entropy (TE) was recently introduced [30] and has been widely applied in many disciplines of sciences such as neurosciences [9]. TE is also successfully applied to analyze the information flow in various economical systems [21,31–35].

Therefore, in this paper we investigate the information transfer networks (ITN) for different periods to understand how the information transfer changes when there occur significant economic events. In addition we also study the relationship between the topological property of ITN and volatility. For this purpose, we select 10 global stock indices over the world. The time series of each index is divided into six different periods. Four of them correspond to the periods of significant financial events such as subprime mortgage crisis, and two of them are ordinary periods for comparison. ITN for each period is constructed from the measured TE. By applying the threshold method [17], we show that ITN undergoes a crossover to a small–world (SW) network from the complete graph. The obtained ITN in SW regime is not static, i.e. the topology of ITN in SW regime depends on time, which shows a behavior of dynamically evolving temporal network [36–38]. Especially, we find that the SW regimes for the periods of global crisis and bear market become larger than those for ordinary periods. Furthermore, when ITN is in SW regime the behavior of clustering coefficient of ITN coincides with that of the volatility. We also compare the results with that for the correlation networks (CN) for each period.

2. Data set

In order to study how the information transfer between different countries is changed when there are significant financial events, we mainly use 10 major global market indices traded from 01/01/1991 to 31/12/2012. These indices are USA, S&P 100; Japan, Nikkei 225; Hong Kong, HSI; Singapore, STI; Australia, AORD; China, SSE; Swiss, SMII; Germany, DAX30; France, CAC40; and UK, FTSE. Though more global indices would be preferable, using more indices results in a substantial shortening of trading periods due to available data. Because of the differences in national holidays and weekends among the countries, the data are adjusted according to the following rules as suggested by Ref. [17]: when more than 30% of markets did not open on a certain day, the data of the day is removed. On the other hand, if it was fewer than 30%, we keep existing indices and insert the last closing value of index for each unopened market. From the obtained time series of index $I$, $Y(t)$, the return of market index $I$ is defined as [1].

$$R_I(t) = \ln(Y_I(t)) - \ln(Y_I(t - \Delta t)).$$

(1)

Here we use $\Delta t = 1$ day.

As an example, $Y(t)$ and volatility $|R_I(t)|$ of S&P100 index are displayed in Fig. 1. The shaded intervals represent the six different periods which are labeled with a number 1–6. Four of them are the periods for significant financial events (periods 3–6 in Fig. 1). For comparison’s purpose, two ordinary periods during which any remarkable financial event does not occur (periods 1 and 2 in Fig. 1) are also investigated [39]. The details of the financial events related to each period are listed in Table 1. As shown in Fig. 1(a), during the ordinary periods (periods 1 and 2), $Y(t)$ increases very slowly and the corresponding $|R_I(t)|$ is relatively small. In contrast to the ordinary periods, in the periods 3–6, $Y(t)$ changes more rapidly, thus $|R_I(t)|$ becomes large, which are the characteristics of significant financial events.

3. Definition of states

To calculate TE between indices, we first define the state, $i_t$, of index $I$ at the day $t$. One of the simplest choices of $i_t$ for a market index $I$ was suggested in Ref. [35] depending on the sign of $R_I(t)$. However, for this binary state, $i_t$ describes only
of changes of each index into the increase or decrease of index $I$. Here we usetheshorthand notation $i|\ell$ as aregarded as non-regular. On the other hand, if $R(t)$ is largert than $|R|^*$, then the changes of indicesare regarded as non-regular. On the other hand, if $-|R|^*/3 \leq R(t) < |R|^*/3$, then the index is regarded as effectivelyunchanged. Otherwise, the change of index is regarded as regular. Assigning rules of state are given in Table 2. The state $i = -1$ ($i = +1$) corresponds to regular decrease (regular increase) of the index and $i = 0$ represents that the index is effectivelyunchanged. On the other hand, $i = -2$ ($i = +2$) stands for the non-regular decrease (non-regular increase) of the index.

4. Transfer entropy

Let $i_t$ ($j_t$) be the state of market index $I$ ($J$) at the day $t$. TE which represents the information flow from $J$ to $I$ is defined as \[ E_{\text{TE}} = \sum_{i_t} p(i_{t+1}, i_t^{(k)} \cdot j_t^{(I)}) \left( \frac{p(i_{t+1} | i_t^{(k)}) \cdot j_t^{(I)}}{p(i_{t+1}^{(k)})} \right) \log \left( \frac{p(i_{t+1} | i_t^{(k)}) \cdot j_t^{(I)}}{p(i_{t+1}^{(k)})} \right). \] Here we use the shorthand notation $i_t = (i_t, \ldots, i_{t-k})$. The sum in Eq. (2) means the sum over all available realization of state $(i_t, i_t^{(k)} \cdot j_t^{(I)})$ in a time series. The joint probability density function $p(i_{t+1}, i_t^{(k)} \cdot j_t^{(I)})$ is the probability that the combination of $i_{t+1}$, $i_t^{(k)}$, and $j_t^{(I)}$ has a particular value. The conditional probability density function $p(i_{t+1} | i_t^{(k)})$ is the probability that $i_{t+1}$ has a particular value when the values of the previous samples $i_t^{(k)}$ and $j_t^{(I)}$ are known. From now on, $k$ and $\ell$ in Eq. (2) are set as $k = \ell = 1$.

Since stock exchange markets have different closing times for a given day $t$ depending on the geographical location of the market, $j_t$ in Eq. (2) should be carefully defined. Thus, we replace $j_t$ in Eq. (2) by $j_t'$ and $E_{\text{TE}} = \sum_{i_t} p(i_{t+1}, i_t, j_t') \left( \frac{p(i_{t+1} | i_t^{(k)}) \cdot j_t^{(I)}}{p(i_{t+1}^{(k)})} \right) \log \left( \frac{p(i_{t+1} | i_t^{(k)}) \cdot j_t^{(I)}}{p(i_{t+1}^{(k)})} \right)$. Here $t'$ represents the first closing day of market $J$ after market $I$ is closed at $t$. If the two markets for $I$ and $J$ are in the same

Table 1

<table>
<thead>
<tr>
<th>Period</th>
<th>Event Description</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>No significant event (ordinary)</td>
<td>1990.01.01–1992.12.31</td>
</tr>
<tr>
<td>2</td>
<td>No significant event (ordinary)</td>
<td>1994.01.01–1995.12.31</td>
</tr>
<tr>
<td>3</td>
<td>Dot Com bubble (bull market) \cite{40}</td>
<td>1998.07.23–2000.08.31</td>
</tr>
<tr>
<td>4</td>
<td>Dot Com bubble (bear market) \cite{41}</td>
<td>2000.09.01–2002.10.09</td>
</tr>
<tr>
<td>5</td>
<td>Subprime mortgage crisis \cite{17}</td>
<td>2007.03.09–2009.03.09</td>
</tr>
<tr>
<td>6</td>
<td>Eurozone crisis \cite{42}</td>
<td>2010.03.09–2012.03.09</td>
</tr>
</tbody>
</table>

Table 2

| Assignment of state $i(t)$ for an index $I$ at day $t$. |
|------------------|------------------|
| $i(t)$           | Range of $R(t)$  | |
| $-2$             | $R(t) < -|R|^*$  | Non-regular decrease |
| $-1$             | $-|R|^* \leq R(t) < -|R|^* / 3$ | Regular decrease |
| $0$              | $-|R|^* / 3 \leq R(t) < |R|^*$ | Unchanged |
| $+1$             | $|R|^* / 3 \leq R(t) < |R|^*$ | Regular increase |
| $+2$             | $R(t) \geq |R|^*$ | Non-regular increase |

Fig. 1. (a) Time evolution of S&P100 index. (b) Plot of $|R(t)|$. The gray-shaded intervals represent the six different periods.
continent, both should have nearly the same closing time ($t = t'$) and $T_{j \to l} = \sum p(l_{t+1}, l'_{t}, J_{t}) \log_2 \frac{p(l_{t+1}|l'_{t}, J_{t})}{p(l_{t+1}|l_{t}, J_{t})}$. On the other hand, if the two markets are in different continents, then $T_{j \to l}$ is calculated in two different ways. If $t' = t + 1$, $T_{j \to l} = \sum p(l_{t+1}, l'_{t}, J_{t+1}) \log_2 \frac{p(l_{t+1}|l'_{t+1}, J_{t+1})}{p(l_{t+1}|l_{t}, J_{t})}$. If $t' = t$, $T_{j \to l} = \sum p(l_{t+1}, l'_{t}, J_{t}) \log_2 \frac{p(l_{t+1}|l'_{t+1}, J_{t+1})}{p(l_{t+1}|l_{t}, J_{t})}$. Therefore, for any case, the causality is not violated.

5. Network analysis

Since TE is asymmetric, ITN constructed by TE is a directed weighted complete graph, in which a node $l$ represents a global market index $l$. The obtained value of $T_{j \to l}$ is assigned to the directed edge from node $J$ to $l$ as its weight, $w_{j \to l}$. For a systematic analysis of the topological properties of ITN, we use the threshold method suggested in Ref. [17]. Starting from a directed weighted complete graph, with a given threshold $\theta$ ($0 \leq \theta$) we remove the directed edges from node $J$ to $l$ if $w_{j \to l} < \theta$.

To analyze the essential properties of ITN, the dependencies of the average clustering coefficient, $\langle C \rangle$, and of the average distance between nodes, $\langle l \rangle$, on $\theta$ are studied as shown in Fig. 2. The clustering coefficient of node $l$ in a directed network is defined as [43]

$$C_l = \frac{(A + A^T)^3_{ll}}{2\left[d_{l}^{in}(d_{l}^{out} - 1) - 2d_l\right]}.$$

Here $A$ is an adjacency matrix, $A^T$ is a transpose of $A$, $d_{l}^{in} \equiv (A^T + A)_l \mathbf{1}$ and $d_{l} \equiv A^T_0$, $(A^T + A)_l$ represents the $l$th row of the matrix $(A^T + A)$, and $\mathbf{1}$ stands for the $N$-dimensional column vector $(1, 1, \ldots, 1)$. $\langle C \rangle$ is obtained by averaging $C_l$ over all nodes $l$. $\langle l \rangle$ is obtained by averaging the distance $l_{i \to j}$ over all possible pairs of nodes $l$ and $J$. $l_{i \to j}$ is defined by the smallest number of directed links from $l$ to $J$. Since ITN is a directed network, $l_{i \to j} \neq l_{j \to i}$. When there is no connected directed path from $l$ to $J$, we assign $l_{i \to j} = 100$. As shown in Fig. 3, there exists the interval of $\theta$ in which $\langle C \rangle$ decreases and $\langle l \rangle$ increases for all periods. The similar behavior of $\langle C \rangle$ and $\langle l \rangle$ was observed by Watts and Strogatz [44] which is known as small-world (SW) phenomena. Thus, we will call this interval of $\theta$ as the SW regime. More precisely, SW regime in Fig. 2 means that the corresponding ITN shows the behavior of SW network with $0 < \langle C \rangle < 1$. The intervals of $\theta$ for SW regimes are shaded by gray (yellow) rectangles in Fig. 2. SW regimes for ordinary periods (Fig. 2(a) and (b)) are smaller than those for periods 4–6 (Fig. 2(d)–(f)) and the value of $\theta$ at which $\langle C \rangle \to 0$ for periods 4–6 is larger than that for periods 1 and 2. The periods 4–6 correspond to bear market or global crisis. This indicates that the information transfer between some of the market indices significantly increases when the market is a bear market or in the global crisis, which increases the fluctuation of TE between each pair of indices. As a result, those indices are strongly connected to each other. On the other hand, for the bull market (period 3) the interval for SW regime does not significantly change from that of period 2 as shown in Fig. 2(c). More interestingly, it is found that the average volatility, $\langle |R| \rangle$, shows the behavior consistent with that of $\langle C \rangle$ in SW regime. As an
example, in Fig. 3, we compare \(|R|\) of S&P 100 with \(C\) obtained from ITN with \(\theta = \theta_{SW} = 0.126\) in each two-year period. With \(\theta = \theta_{SW}\) ITNs for all periods are in SW regime [45].

Each two-year period is marked with vertical grids in Fig. 3. The data clearly show the synchronized behavior of \(|R|\) with \(C\) for ITN, where \(C\) for ITN increases (decreases) as \(|R|\) increases (decreases).

For a comparison, we also construct the correlation network (CN) based on the asymmetric cross-correlation,

\[
G_{i,j} = \frac{\langle R_i(t)R_j(t') \rangle - \langle R_i(t) \rangle \langle R_j(t') \rangle}{\sqrt{\langle R_i^2(t) - \langle R_i(t) \rangle^2 \rangle \langle R_j^2(t') - \langle R_j(t') \rangle^2 \rangle}}. \tag{4}
\]

Here \(\langle \cdots \rangle\) represents the average over the given period. If \(I\) and \(J\) are in the same continent, then \(t = t'\). If \(I\) belongs to the continent different from that for \(J\), then \(t'\) represents the first closing day of market \(J\) after market \(I\) is closed at \(t\). CN is constructed by assigning weight \(w_{i,j} = G_{i,j}\) to a directed edge from \(I\) to \(J\). Then we also apply the same threshold method as for ITN. The measured \(\langle C \rangle\)'s of CN at \(\theta_{SW} = 0.42\) are displayed in Fig. 3. As shown in Fig. 3, the behavior of \(|R_{S&P100}|\) is hardly inferred from \(C\)'s for CN. All the markets except for China show the same behavior (which is not shown).

For a quantitative analysis of the correlation between \(|R|\) and \(C\) in SW regime, we calculate the Pearson coefficient defined as,

\[
G_{\langle |R(T)| \rangle, \langle C(T) \rangle} = \frac{\langle \langle |R(T)| \rangle \langle C(T) \rangle \rangle_T - \langle \langle |R(T)| \rangle \rangle_T \langle \langle C(T) \rangle \rangle_T}{\sqrt{\langle \langle |R(T)| \rangle^2 \rangle_T - \langle \langle |R(T)| \rangle \rangle_T^2} \langle \langle C(T) \rangle^2 \rangle_T - \langle \langle C(T) \rangle \rangle_T^2} \rangle_T, \tag{5}
\]

where \(T\) represents the periods marked by vertical grids in Fig. 3. \(\langle |R(T)| \rangle\) and \(\langle C(T) \rangle\) are the averages over a certain two-year period \(T\). \(\langle \cdots \rangle_T\) denotes the average over all two-year periods. In Fig. 4(a) we display the obtained Pearson coefficient for each network with 10 major market indices. The data in Fig. 4(a) clearly show that \(G_{\langle |R(T)| \rangle, \langle C(T) \rangle} > G_{\langle |R(T)| \rangle, \langle C(T) \rangle}\) for all indices except for China. This implies that the analysis of the topological property of ITN provides much better insight to understand the volatility changes in global market indices. In Fig. 4(b) the obtained Pearson coefficients for 14 indices traded from 01/09/1995 to 31/12/2013 are shown. Even though the estimated errors are large because of the shortened traded interval, the data shows the same behavior as in Fig. 4(a). For 14 indices, we also find that the value of \(G_{\langle |R| \rangle, \langle C(T) \rangle} \) for Brazil, China, and Russia are relatively smaller than the other countries. These countries are classified as emerging markets. This implies that the volatility for emerging market is not relatively well synchronized with the average clustering coefficient. This can be understood from the characteristic feature for the emerging markets such as antileverage effect [46].

6. Summary

Based on the measured TE we investigate the topological properties of ITN composed of 10 global market indices. Since TE is asymmetric, the resulting ITN is a weighted and directed graph. By using the threshold method, we find that ITN undergoes from a directed weighted complete graph to a SW network. When ITN is in SW regime, it is found that the average clustering coefficient of ITN is synchronized with volatility, except for China. However, the behavior of volatility is hardly synchronized with the clustering coefficient of CN. We also find that SW regimes are more enhanced for bear market or global crisis than for ordinary periods, because the interval of threshold \(\theta\) for SW regime is much wider for bear markets. This result clearly shows that the information flow between markets through ITN significantly increases when there occurs a world-wide financial shock. In contrast, the information flow is hardly captured by CN. Therefore, understanding the topological properties of ITN is essential to study various properties of time series in global financial...
systems. It is also expected that ITN plays an important role as a underlying interaction network in the theoretical modeling of economic systems.

Acknowledgment

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science ICT and Future Planning (NRF-2013R1A1A2057791 and NRF-2012R1A1A2007430).

References

The ordinary periods are defined as the periods during which there is no significant financial event. For a quantitative verification, we measure $\langle |R| \rangle$, the average volatility of each index $I$ over the period from 01/01/1991 to 31/12/2012. We also measure the average volatility of each index $I$ over each two-year period $T$, $\langle |R| \rangle_T$. Since it is widely believed that the volatility during the period for a significant financial event is generally larger than that during the ordinary period \cite{12, 13}, we compare $\langle |R| \rangle_T$ with $\langle |R| \rangle$. For a two-year period $T$, if $\langle |R| \rangle > \langle |R| \rangle_T$, then the period is regarded as an ordinary period. Otherwise the period is regarded as a period for a significant financial event. The result coincides with Table 1.