A Network Perspective of World Stock Markets

Michael Tse, Hong Kong Polytechnic University

February 2013
Acknowledgments

Dr Xiaofan Liu
Former PhD Student
Now with Southeast University

Prof. Francis Lau
HK Polytechnic University

Dr Jing Liu
Wuhan University

Research Group Members and their Families at Tai Tam Trail
Questions

- How do stocks interact within a market?
- How do different stock markets interact?
Key words for today

- **Scalefree network** of stocks
- **Synchronization** of stock markets
- **Volatility** and **fluctuation** of stocks and stock markets
Stock Market as Network
The US Stock Case Study
Stock Market as Network
A Simple View

- Node = Stock
- Edge = Connection of a pair of stocks having a “correlation”
- Depending on the way “correlation” is defined, different networks can be constructed for different contexts.

- But definitions of nodes and edges can be abstract to produce networks for specific applications.
Network

- Time series can be:
  - Closing price $p_i(t)$
  - Price return $r_i(t)$
  - Trading volume

**Edge Definition:**
Cross correlations are used to determine connectivity.

$$c_{ij} = \frac{\sum_t (x_i(t) - \bar{x}_i)(x_j(t) - \bar{x}_j)}{\sqrt{\sum_t (x_i(t) - \bar{x}_i)^2} \sqrt{\sum_t (x_j(t) - \bar{x}_j)^2}}$$

where $x_i$ is the stock price of stock $i$. If $c_{ij} > \rho$, for example, we connect stock $i$ and stock $j$, i.e., **winner-take-all connection criterion.**
Example: closing price

\[ c_{ij} = \frac{\sum_t (x_i(t) - \bar{x}_i)(x_j(t) - \bar{x}_j)}{\sqrt{\sum_t (x_i(t) - \bar{x}_i)^2} \sqrt{\sum_t (x_j(t) - \bar{x}_j)^2}} \]
US Stock Market Network

* We consider full network. No trimming or reduction.
  * **Data Set 1**: All US stocks that are traded between July 1, 2005 to August 30, 2007. Total = 19,807, out of 51,835 US stocks.
  * **Data Set 2**: All US stocks that are traded between June 1, 2007 to May 30, 2009.
* Closing prices, price returns and trading volumes are considered.
* Time series are analyzed.

**Edge Definition:**
Cross correlations are used to determine connectivity. If the time series of two stocks are “highly correlated”, they are connected.
It’s scalefree!

Degree distribution: $p(k) \text{ vs } k$

Scalefree: $p(k) = \alpha e^{-\gamma k}$

For $\rho = 0.9$

or just regular

* The power-law degree distribution holds better for large cut-off (e.g., \(\rho = 0.9\)) and becomes blur as \(\rho\) decreases, which is again consistent with the fact that the network becomes effectively more fully connected as \(\rho\) decreases. The fitting error is a useful parameter to measure how “scalefree” the distribution is.

\[
\epsilon_{\text{fitting}} = \sum_k \left| p(k) - \alpha e^{-\gamma k} \right|
\]
Network Parameters

- Network from Closing Price data
- Network from Price Return data
- Network from Trading Volume

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$\rho = 0.70$</th>
<th>$\rho = 0.80$</th>
<th>$\rho = 0.90$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Nodes $N$</td>
<td>19807</td>
<td>19807</td>
<td>19807</td>
</tr>
<tr>
<td>Number of connections $L$</td>
<td>15785</td>
<td>6675</td>
<td>2359</td>
</tr>
<tr>
<td>Average shortest length $s$</td>
<td>2.946</td>
<td>2.290</td>
<td>2.043</td>
</tr>
<tr>
<td>Diameter $D$</td>
<td>20</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>Average clustering coefficient $C$</td>
<td>0.104</td>
<td>0.058</td>
<td>0.238</td>
</tr>
<tr>
<td>Average degree $K$</td>
<td>1.594</td>
<td>0.674</td>
<td>0.238</td>
</tr>
<tr>
<td>Power-law exponent $\gamma$</td>
<td>2.019</td>
<td>3.067</td>
<td>2.920</td>
</tr>
<tr>
<td>Mean fitting error</td>
<td>1.607e-5</td>
<td>8.29e-5</td>
<td>2.78e-6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$\rho = 0.85$</th>
<th>$\rho = 0.90$</th>
<th>$\rho = 0.95$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Nodes $N$</td>
<td>19807</td>
<td>19807</td>
<td>19807</td>
</tr>
<tr>
<td>Number of connections $L$</td>
<td>4652650</td>
<td>1495250</td>
<td>143181</td>
</tr>
<tr>
<td>Average shortest length $s$</td>
<td>3.375</td>
<td>3.954</td>
<td>4.995</td>
</tr>
<tr>
<td>Diameter $D$</td>
<td>16</td>
<td>18</td>
<td>30</td>
</tr>
<tr>
<td>Average clustering coefficient $C$</td>
<td>0.421</td>
<td>0.302</td>
<td>0.148</td>
</tr>
<tr>
<td>Average degree $K$</td>
<td>469.80</td>
<td>150.98</td>
<td>14.46</td>
</tr>
<tr>
<td>Power-law exponent $\gamma$</td>
<td>0.778</td>
<td>1.075</td>
<td>0.992</td>
</tr>
<tr>
<td>Mean fitting error</td>
<td>6.26e-7</td>
<td>4.26e-7</td>
<td>1.65e-7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$\rho = 0.70$</th>
<th>$\rho = 0.80$</th>
<th>$\rho = 0.90$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Nodes $N$</td>
<td>19807</td>
<td>19807</td>
<td>19807</td>
</tr>
<tr>
<td>Number of connections $L$</td>
<td>256046</td>
<td>167340</td>
<td>96203</td>
</tr>
<tr>
<td>Average shortest length $s$</td>
<td>4.542</td>
<td>4.927</td>
<td>7.165</td>
</tr>
<tr>
<td>Diameter $D$</td>
<td>21</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>Average clustering coefficient $C$</td>
<td>0.260</td>
<td>0.194</td>
<td>0.140</td>
</tr>
<tr>
<td>Average degree $K$</td>
<td>25.854</td>
<td>16.897</td>
<td>9.714</td>
</tr>
<tr>
<td>Power-law exponent $\gamma$</td>
<td>1.374</td>
<td>1.285</td>
<td>1.5933</td>
</tr>
<tr>
<td>Mean fitting error</td>
<td>1.33e-5</td>
<td>2.56e-6</td>
<td>2.50e-7</td>
</tr>
</tbody>
</table>
What does it mean by being scalefree?

- Stocks having close resemblance with a large number of other stocks are relatively few.
- Thus, the stock market is essentially influenced by a relatively small number of stocks.
- We may introduce an index that reflects on the performance of the stock market based on a small number of stocks that have a relatively high number of connections. In other words, an index can be defined by the stocks of high degrees.

\[
\text{Index} = \frac{\sum_i \text{price}_i \times \text{number of shares}_i}{\text{total market value of stocks during base period}}
\]

- (Market capitalization formula)
Interim Conclusion: Small is influential

Who are the most influential?
A network perspective of the stock market

Chi K. Tse a,*, Jing Liu a,b, Francis C.M. Lau a

a Department of Electronic and Information Engineering, Hong Kong Polytechnic University, Hong Kong
b State Key Laboratory for Software Engineering, Wuhan University, Hebei, China

ARTICLE INFO
Article history:
Received 31 July 2008
Received in revised form 27 July 2009
Accepted 29 April 2010
Available online 16 May 2010

Keywords:
Stock market
Complex network
Degree distribution
Stock indexes

ABSTRACT
Complex networks are constructed to study correlations between the closing prices for all US stocks that were traded over two periods of time (from July 2005 to August 2007; and from June 2007 to May 2009). The nodes are the stocks, and the connections are determined by cross correlations of the variations of the stock prices, price returns and trading volumes within a chosen period of time. Specifically, a winner-take-all approach is used to determine if two nodes are connected by an edge. So far, no previous work has attempted to construct a full network of US stock prices that gives full information about their interdependence. We report that all networks based on connecting stocks of highly correlated stock prices, price returns and trading volumes, display a scalefree degree distribution. The results from this work clearly suggest that the variation of stock prices are strongly influenced by a relatively small number of stocks. We propose a new approach for selecting stocks for inclusion in a stock index and compare it with existing indexes. From the composition of the highly connected stocks, it can be concluded that the market is heavily dominated by stocks in the financial sector.

© 2010 Elsevier B.V. All rights reserved.
Fluctuation and Network Dynamics

Disrupting scalefreeness!
Network Dynamics

- We consider a window of time and take snapshots of the network as time goes.
- Data used in this part of study are from the Standard & Poor's 500 (S&P500) stocks that were traded from January 1, 2000 to December 31, 2004.
Our task is to

* find the link between

  * network’s phenomena:

  * market’s phenomena:

  Scalefree network
  Measure: fitting error

  Market fluctuation
  Measure: volatility
Market Fluctuations

- Market Index: measure of the overall market performance
  - S&P 500
  - Dow Jones
  - Nasdaq and etc.

- Average Index Volatility (AIV): fractional change of the average index values between two consecutive time windows

\[
AIV(t) = \frac{|\langle I(t) \rangle_{i+1} - \langle I(t) \rangle_i|}{\langle I(t) \rangle_i} \quad \langle I(t) \rangle_i = \frac{\sum_{k=0}^{T-1} I(t_i + k \cdot \Delta t)}{T}
\]
Average Index Volatility (AIV)

S&P500 index

Standard deviation of S&P500 index (StdI)

Average index volatility of S&P500 index (AIV)

Low passed average index volatility of S&P500 index (AIV')
Standard deviation is highly correlated with AIV'.
Network properties vs market fluctuation
• AIV' and fitting error
Interim Conclusion: Fluctuation lessens scalefreeness!

- Fitting error is a measure of “scalefreeness”
- Volatility is highly correlated to the loss of scalefreeness.
- Details of statistical analysis: Liu, Tse and Ke, Quant. Finan. 2009
Fierce stock market fluctuation disrupts scalefree distribution

JING LIU†, CHI K. TSE*† and KEQING HE‡

†Department of Electronic and Information Engineering,
The Hong Kong Polytechnic University, Hong Kong
‡State Key Laboratory for Software Engineering, Wuhan University,
Hubei, 430072, People’s Republic of China

(Received 11 August 2008; in final form 21 April 2009)

1. Introduction

Recently, in the finance discipline, cross correlations of stock price return time series have been increasingly used in the study of the internal structure of stock markets. The data obtained from calculating the cross correlations of stocks’ times series forms a correlation matrix which provides information about the interdependence of the stocks. From a network viewpoint, the stocks form a complex network that describes how the individual stocks are related. Some proposed models for studying

In this paper, we employ a network model similar to the asset graph to study the correlation between the time series of stock prices and the time series of stock price returns. Data used in this study are from the Standard and Poor’s 500 (S&P500) stocks that were traded from 1 January 2000 to 31 December 2004 (see Historical Data for S&P500 Stocks 2009). For an observing window of 7 days, we construct a complex network for the stocks. A simple connection rule is used to construct networks. Essentially, any two stocks whose correlation is larger than a threshold, \( p \), are considered to be connected. It has been found that when \( p = 0.4 \), the resulting
Connecting the Markets

Synchronize or get panic!
Markets as “nodes”

**Node:** Market

**Edge:** Similarity between markets

Take 20 working days of index closing values
Approximately one month of data
Calculate Pearson’s correlation between markets
This is defined as the EDGE WEIGHT

\[ \rho_{i,j}(m) = \frac{\text{cov}(P_i(m), P_j(m))}{\sigma_{P_i(m)} \sigma_{P_j(m)}} \]
Network of Markets

- 32 countries (each represented by an index)
Market Benchmark: Index

- Each market is benchmarked by its index, which represents collective behavior of the stocks within the market.

  - US: Standard & Poor 500
  - Hong Kong: Hang Seng
Distribution of Correlations

- 496 correlation values between each pair of 32 stock market indices
- Correlation values range from -1 to 1
- Right skewed Pearson’s distribution
Snapshot of Global Network

- Window starting on Aug 28, 2008 to Sept 17, 2008
- Correlation threshold = 0.85 (connect only if \( \geq 0.85 \))
Dynamics of Network

- To capture the dynamics of the network
- 1000+ windows with size of 20 days
- Window shifts one day per movement
- Starting from March 2004 until April 2009
Dynamics of Correlation Distribution

- A plot of correlation distribution in 280 consecutive windows
- Transform between right skewed Pearson’s distribution and random distribution
Dynamics of Node Connectivity

- Node Strength = Average edge weight of a node

\[ s_i(m) = \frac{1}{31} \sum_{j=1, i \neq j}^{32} \rho_{ij}(m). \]

Red: HSI, green: DJI
Network Synchronization

- Network Synchronization = Average edge weight of all nodes

\[ s_{\text{NET}}(m) = \frac{1}{496} \sum_{i=1}^{32} \sum_{j=1, i \neq j}^{32} \rho_{ij}(m) \]
Challenge is to find the link

Average index closing value

\[ \mu_i(m) = \frac{\sum_{t=t_m}^{t_m+w-1} p_i(t)}{w} \]

Volatility

\[ \sigma_i(m) = \sqrt{\frac{\sum_{t=t_m}^{t_m+w-1} (p_i(t) - \mu_i(m))^2}{w-1}} \]

Average return (incremental change in window)

\[ r_i(m) = \frac{p_i(t_{m+w-1}) - p_i(t_m)}{p_i(t_m)} \times 100\% \]

network synchronization

node strength
Domestic Comparison

<table>
<thead>
<tr>
<th>Index</th>
<th>Country</th>
<th>$\rho_{s,r}$</th>
<th>$\rho_{s,\mu}$</th>
<th>$\rho_{s,\sigma}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATX</td>
<td>Austria</td>
<td>-0.23</td>
<td>-0.10</td>
<td>0.64</td>
</tr>
<tr>
<td>All Ordinaries</td>
<td>Australia</td>
<td>-0.16</td>
<td>0.09</td>
<td>0.59</td>
</tr>
<tr>
<td>FTSE 100</td>
<td>United Kingdom</td>
<td>-0.25</td>
<td>-0.14</td>
<td>0.57</td>
</tr>
<tr>
<td>CAC 40</td>
<td>France</td>
<td>-0.26</td>
<td>-0.16</td>
<td>0.57</td>
</tr>
<tr>
<td>DAX</td>
<td>Germany</td>
<td>-0.24</td>
<td>0.00</td>
<td>0.56</td>
</tr>
<tr>
<td>S&amp;P Mib</td>
<td>Italy</td>
<td>-0.14</td>
<td>-0.25</td>
<td>0.55</td>
</tr>
<tr>
<td>TA-100</td>
<td>Israel</td>
<td>0.03</td>
<td>0.17</td>
<td>0.55</td>
</tr>
<tr>
<td>AEX</td>
<td>Netherlands</td>
<td>-0.24</td>
<td>-0.18</td>
<td>0.54</td>
</tr>
<tr>
<td>OMX Stockholm 30</td>
<td>Sweden</td>
<td>-0.22</td>
<td>-0.11</td>
<td>0.53</td>
</tr>
<tr>
<td>Swiss Market</td>
<td>Switzerland</td>
<td>-0.11</td>
<td>-0.08</td>
<td>0.53</td>
</tr>
<tr>
<td>Hang Seng</td>
<td>Hong Kong</td>
<td>-0.13</td>
<td>0.10</td>
<td>0.52</td>
</tr>
<tr>
<td>OMX Copenhagen 20</td>
<td>Denmark</td>
<td>-0.26</td>
<td>0.00</td>
<td>0.52</td>
</tr>
<tr>
<td>S&amp;P 500</td>
<td>USA</td>
<td>-0.16</td>
<td>-0.08</td>
<td>0.52</td>
</tr>
<tr>
<td>BEL-20</td>
<td>Belgium</td>
<td>-0.18</td>
<td>-0.16</td>
<td>0.51</td>
</tr>
<tr>
<td>Nikkei 225</td>
<td>Japan</td>
<td>-0.11</td>
<td>-0.26</td>
<td>0.49</td>
</tr>
<tr>
<td>Total Share</td>
<td>Norway</td>
<td>0.25</td>
<td>-0.06</td>
<td>0.49</td>
</tr>
<tr>
<td>NZSE 50</td>
<td>New Zealand</td>
<td>0.11</td>
<td>-0.04</td>
<td>0.48</td>
</tr>
<tr>
<td>Bovespa</td>
<td>Brazil</td>
<td>-0.01</td>
<td>0.17</td>
<td>0.48</td>
</tr>
<tr>
<td>Straits Times</td>
<td>Singapore</td>
<td>-0.08</td>
<td>-0.02</td>
<td>0.46</td>
</tr>
<tr>
<td>Madrid General</td>
<td>Spain</td>
<td>-0.13</td>
<td>-0.24</td>
<td>0.45</td>
</tr>
<tr>
<td>IPC</td>
<td>Mexico</td>
<td>-0.02</td>
<td>-0.08</td>
<td>0.45</td>
</tr>
<tr>
<td>MerVal</td>
<td>Argentina</td>
<td>-0.04</td>
<td>-0.06</td>
<td>0.44</td>
</tr>
<tr>
<td>KLSE Composite</td>
<td>Malaysia</td>
<td>-0.05</td>
<td>0.16</td>
<td>0.44</td>
</tr>
<tr>
<td>BSE 30</td>
<td>India (BSE)</td>
<td>-0.22</td>
<td>-0.14</td>
<td>0.43</td>
</tr>
<tr>
<td>S&amp;P TSX Composite</td>
<td>Canada</td>
<td>-0.11</td>
<td>-0.16</td>
<td>0.42</td>
</tr>
<tr>
<td>Seoul Composite</td>
<td>South Korea</td>
<td>-0.05</td>
<td>0.03</td>
<td>0.41</td>
</tr>
</tbody>
</table>

s: node strength  
r: price return  
$\mu$: average closing value  
$\sigma$: volatility

* Node strength is highly correlated with the volatility of corresponding market index
Global Comparison

Network synchronization is highly correlated with the volatility of world market index (MSCI AC World Index)

Table 1: Pearson’s correlation coefficients between each pair of dynamics of network synchronization $s$, World Index window return $r$, average $\mu$ and volatility $\sigma$.

<table>
<thead>
<tr>
<th>$w$</th>
<th>$\rho_{s,r}$</th>
<th>$\rho_{s,\mu}$</th>
<th>$\rho_{s,\sigma}$</th>
<th>$\rho_{r,\mu}$</th>
<th>$\rho_{r,\sigma}$</th>
<th>$\rho_{\mu,\sigma}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>-0.18</td>
<td>-0.10</td>
<td>0.65</td>
<td>0.07</td>
<td>-0.49</td>
<td>-0.21</td>
</tr>
<tr>
<td>20</td>
<td>-0.28</td>
<td>-0.13</td>
<td>0.65</td>
<td>0.12</td>
<td>-0.63</td>
<td>-0.21</td>
</tr>
<tr>
<td>40</td>
<td>-0.33</td>
<td>-0.10</td>
<td>0.65</td>
<td>0.24</td>
<td>-0.75</td>
<td>-0.20</td>
</tr>
<tr>
<td>60</td>
<td>-0.34</td>
<td>-0.05</td>
<td>0.61</td>
<td>0.31</td>
<td>-0.83</td>
<td>-0.19</td>
</tr>
<tr>
<td>120</td>
<td>-0.29</td>
<td>-0.38</td>
<td>0.55</td>
<td>0.38</td>
<td>-0.88</td>
<td>-0.33</td>
</tr>
</tbody>
</table>
A COMPLEX NETWORK PERSPECTIVE OF WORLD STOCK MARKETS: SYNCHRONIZATION AND VOLATILITY

XIAO FAN LIU and CHI K. TSE*
Department of Electronic and Information Engineering,
The Hong Kong Polytechnic University, Hong Kong
*encktse@polyu.edu.hk

Received March 15, 2011

This paper studies the cross-correlations of 67 stock market indices in the past 5 years. In order to capture the interaction of the stock markets, we propose to take a complex network approach to analyzing the interdependence of the individual stock markets. Specifically, stock markets are considered as network nodes, and the network links (weights of links) are defined by the cross-correlations between market indices over a period of time (time window). Thus, the resulting network provides information about the interdependence of individual markets, with the network links representing the extents to which the markets are correlated. If we allow the time window to move in forward time and construct a network for each time window over a long period of time, we are able to capture the dynamics of the network. In our study, all networks are constructed from raw data of market indices, and our aim is to investigate how network properties can be used to infer market behavior. By examining the variation of the network parameters as time elapses, we show that stock markets of different countries have time-varying interaction, and that developed markets tend to demonstrate similar behavior while emerging markets are statistically independent of each other. Furthermore, we observe synchronization in the network of stock markets, which is an important universal property observed in many physical and man-made networks. Specifically, we show that stock markets of different countries generally behave in a synchronous manner when they experience fluctuation, which is especially notable in the developed markets. This work exposes the interdependence of stock markets in the world and proposes a complex network approach to identifying some salient global behavior of the interconnecting markets.

Keywords: Complex network; synchronization; assortativity; stock market; volatility.
Additional information

How individual market interact with the world?

Hang Seng

Nikkei
Conclusion

- Using a network perspective can enrich understanding of systems. Challenges:
  - how to identify nodes and connections
  - how to link physical phenomena with the network properties
- For the stock network problems we are dealing with here,
  - network provides useful clue as to the interaction of stocks within a market:
    - scalefree structure implies strong influence of a small group;
  - network structure is related to its dynamical change:
    - scalefreeness disrupted at times of fluctuation
  - it also provides clear connection of the behavior of the different markets:
    - get panic and be synchronized!
It’s always your perspective that determines how much you understand.