Observing Stock Market Fluctuation in Networks of Stocks

C.K. Tse[1], J. Liu[2], F.C.M. Lau[1], and K. He[2]

[1] The Hong Kong Polytechnic University, Kowloon, Hong Kong
encktse@polyu.edu.hk, encmlau@polyu.edu.hk
[2] State Key Laboratory for Software Engineering, Wuhan University, China
j_liu@whu.edu.cn, hekeqing@public.wh.hb.cn

Abstract. In this paper we study the structural variation of the network formed by connecting Standard & Poor's 500 (S&P500) stocks whose closing prices (or price returns) are highly correlated. Specifically we consider S&P500 stocks that were traded from January 1, 2000 to December 31, 2004, and construct complex networks based on cross correlation between the time series of the closing prices (or price returns) over a fixed period of time. A simple threshold approach is used for establishing connections between stocks. The period over which the network is constructed is 20 trading days, which should be long enough to produce meaningful cross correlation values, but sufficiently short in order to avoid averaging effects that smooth off the salient fluctuations. A network is constructed for each 20-trading-day window in the entire trading period under study. The window moves at a 1-trading-day step. The power-law exponent is determined for each window, along with the corresponding mean error of the power law approximation which reflects how closely the degree distribution resembles a scalefree-like distribution. The key finding is that the scalefreeness of the degree distribution is disrupted when the market experiences fluctuation. Thus, the mean error of the power-law approximation becomes an effective indicative parameter of the volatility of the stock market.

1 Introduction

Complex network models have been constructed recently for studying the interdependence of stock prices [1]–[8]. In our earlier work [9], we have made an attempt to construct a full network of US stocks, without applying any specific filtering procedure to reduce complexity. This method has been used to produce complex networks from time series of closing prices, price returns and trading volumes [9]. Such stock networks have been used to study how stocks are connected and the structure of the interconnections. However, the dynamics of the networks has not been exploited for detailed study of the way the stock market varies as time elapses, and in particular the relationship between the market fluctuation and the time-varying structures of the stock networks.

In this paper we study the variations of the network parameters and attempt to relate such variations with the market fluctuation. In particular, we will base

J. Zhou (Ed.): Complex 2009, Part II, LNICST 5, pp. 2099–2108, 2009.

our study on the Standard & Poor's 500 (S&P500) stocks such that the networks constructed from these stocks can be consistently compared with the fluctuation of the S&P500 index [11,12]. Basically we consider cross correlation between the closing prices (or price returns) of the S&P500 stocks over a period of 20 days, and construct networks by connecting stocks that are highly correlated. The networks generated have been found to exhibit a scalefree-like degree distribution. In this work, we construct networks for each 20-day window over the entire period from January 1, 2000 to December 31, 2004. A snapshot of the network is taken for each window ($T = 20$ days), and the window moves along the time scale. Thus, effectively, we are taking snapshots of the network of stocks at 1-trading-day intervals, and the variation of the network can thus be studied in terms of the variation of the parameters as time elapses.

We will focus on the degree distribution of the network. By evaluating the mean error of the power law approximation, we quantify the resemblance of the degree distribution to a scalefree-like distribution, and we compare this property with the market fluctuation in terms of stock index volatility which is defined as the incremental change of the average stock index. Our main objective is to study how the scalefreeness of the network is related to the performance of the stock market. As will be shown in this work, the scalefreeness of the degree distribution gives a very strong indication of market fluctuation. This fundamental finding was not reported previously.

In Section 2, we give a quick review of the construction of complex networks based on cross correlations of the time series of stock prices. In Section 3, we illustrate the construction of market variation time series. In Section 3.2, we examine the dynamics of the networks by examining the variations of the network parameters. In Section 4, we examine on the variation of scalefreeness of the network. Finally we give some conclusions in Section 5.

2 Construction of Complex Networks of Stocks

We consider a network of US stocks that were traded between a given period of time. For each pair of stocks (nodes), we will evaluate the cross correlation of the time series of their *daily closing prices* and *daily price returns.* Thus, two networks can be constructed, one corresponding to closing prices and the other to price returns.

Let $p_i(t)$ be the *closing price* of stock i on day t. Then, the *price return* of stock i on day t, denoted by $r_i(t)$, is defined as

$$r_i(t) = \ln \left[\frac{p_i(t)}{p_i(t-1)} \right] \tag{1}$$

Suppose $x_i(t)$ and $x_j(t)$ are the daily prices or price returns of stock i and stock j, respectively, over the period $t = 0$ to $N - 1$. We now compare the two time series with no relative delay. In other words, x_i and x_j are compared from

$i = 0$ to $N - 1$ with no relative time shift. The cross correlation between x_i and x_j is given by [10]

$$c_{ij} = \frac{\sum_t [(x_i(t) - \overline{x_i})(x_j(t) - \overline{x_j})]}{\sqrt{\sum_t (x_i(t) - \overline{x_i})^2} \sqrt{\sum_t (x_j - \overline{x_j})^2}} \tag{2}$$

where $\overline{x_i}$ and $\overline{x_j}$ are the means of the time series and the summations are taken over $t = 0$ to $N - 1$.

In defining our criterion for connecting a pair of nodes, we need a threshold value for the cross correlation. Since cross correlation is a measure of similarity and its value is between 0 and 1, we simply choose a positive fractional value as the threshold. Suppose the threshold is ρ. Then, the connection criterion for stock i and stock j is

$$c_{ij} > \rho. \tag{3}$$

2.1 Degree Distribution

We have constructed the closing price network and the price return network. Scalefree-like degree distribution has been found and the approximated power-law exponent is about 0.6 to 0.87 for $\rho = 0.8$ to 0.9. Fig. 1 shows the distribution for the closing price network for $\rho = 0.9$. A comprehensive set of results for the entire US stock market can be found in Tse *et al.* [9].

At this point, we should note that power-law distributions can be more formally argued for via suitable statistical tests, such as the Kolmogorov-Smirnov test (KS-test), and a tailored procedure for testing the goodness-of-fit of given data to the power-law distribution has been proposed by Goldstein *et al.* [13]. However, application of the method to practical data requires some attention. Since the procedure involves computing the cumulative distributions before

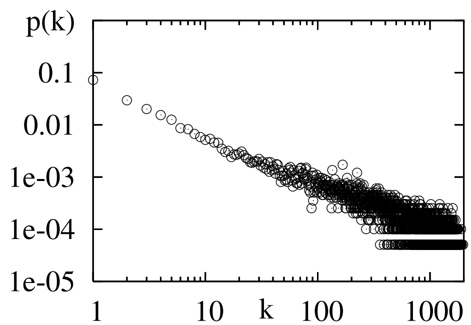

Fig. 1. Degree distribution of closing price network formed by connection criterion based on cross correlation for $\rho = 0.9$. Data are all US stocks traded from July 1, 2005 to August 30, 2007.

application of the KS-test, practicality is restricted to networks with power-law exponents greater than 1. Furthermore, real data often contains irregular "heads" and "tails" (segments corresponding to very low and very high degrees) where the power-law distribution is not interested or practically invalid. Thus, some pre-screening of data is needed, and the practical implementation of the method thus becomes quite tedious and time consuming if the number of distributions to be fitted is very large. In this work we choose a simpler procedure to measure the fitness via the total fitting error.

2.2 Dynamics of Networks

Now suppose we construct a network over a period of $N = 20$ days, initially from $t = 0$ to $t = 19$. As we advance in time, we can construct networks for all 20-day periods, i.e., from $t = 0$ to $t = 19$, from $t = 1$ to $t = 20$, from $t = 2$ to $t = 21$, from $t = 3$ to $t = 22$, etc. until all data are exhausted. Essentially, we are taking snapshots of the network at 1-day intervals.

3 Market Fluctuation from S&P500 Index

The network constructed from the time series within a particular 20-day window basically reflects the stock market internal structure for the 20-day period concerned. The series of networks as time elapses thus provides information about the structural change of the network over time. In other words, we are able to capture how the network parameters and structure change as time elapses, and in the following we attempt to compare these changes with the way the market fluctuates. Our metric for market fluctuation is the *average index volatility,* which will be defined in the following subsection.

3.1 Index Volatility

In order to measure the stock market fluctuation over a time interval, we define an *average index volatility* based on the variation of the average index value over an interval of T days. A time series of the stock market fluctuation can thus be obtained as the window moves. To be comparable with the market internal structure variation, the same window size T and step of movement $T - \delta T$ should be taken, where δT is the overlapping period between two consecutive time windows.

Consider a stock index whose value is $I(t)$ at time t. The original time series is divided into M windows: $W_1, W_2, ..., W_M$.

Average Index Volatility (AIV) is defined as the fractional change of the average index values of two consecutive time windows:

$$\mathrm{AIV}(t) = \frac{|\,\langle I(t)\rangle_{i+1} - \langle I(t)\rangle_i\,|}{\langle I(t)\rangle_i} \tag{4}$$

where $\langle I(t)\rangle_i$ is the average index value in window W_i, i.e.,

$$\langle I(t) \rangle_i = \frac{\sum_{k=0}^{T-1} I(t_i + k \cdot \Delta t)}{T} \tag{5}$$

where t_i is the starting time point of window W_i.

The absolute value of the difference is used here because significant fluctuations of the stock index are usually caused by synchronized stock price movement in one direction, upward or downward.

Fig. 2 shows the time series of AIV in the studied period. To smooth out the spikes in the AIV time series, we calculate the value of each point as the average of its neighboring nodes, resulting in AIV′, which is effectively the low-pass filtered version of AIV.

3.2 Correlation of Market Fluctuation and Network Parameters

In this subsection, we construct networks for the S&P500 stocks that were traded from January 1, 2000 to December 31, 2004. The time variations of some network parameters are captured using the moving 20-day window network described in the foregoing. Specifically, closing price and price return time series are analyzed with window size $T = 20$ and $\delta T = 19$. Time series of network parameters such as average degree (K), average cluster coefficient (C), average shortest length (d) and diameter (D) are computed and shown in Figs. 3 and 4. Their correlations with the AIV′ time series are also calculated and given in Table 1.

From Table 1, Figs. 3 and 4, we see that K and C are closely correlated with AIV′, whereas d and D are much less correlated with AIV′. This indicates that

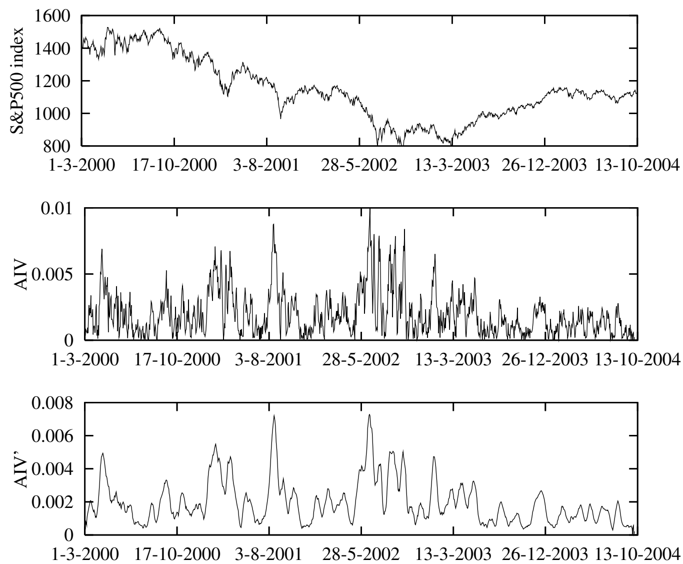

Fig. 2. Time series of S&P500 index, average index volatility (AIV), and low-pass filtered average index volatility (AIV′)

Table 1. Cross correlations between AIV′ and network parameters

Cross correlation	AIV′ and K	AIV′ and C	AIV′ and d
Closing price network	0.512	0.572	-0.470
Price return network	0.303	0.339	-0.309

Cross correlation	AIV′ and D	AIV′ and γ	AIV′ and fitting error
Closing price network	-0.264	-0.475	0.605
Price return network	-0.265	-0.308	0.358

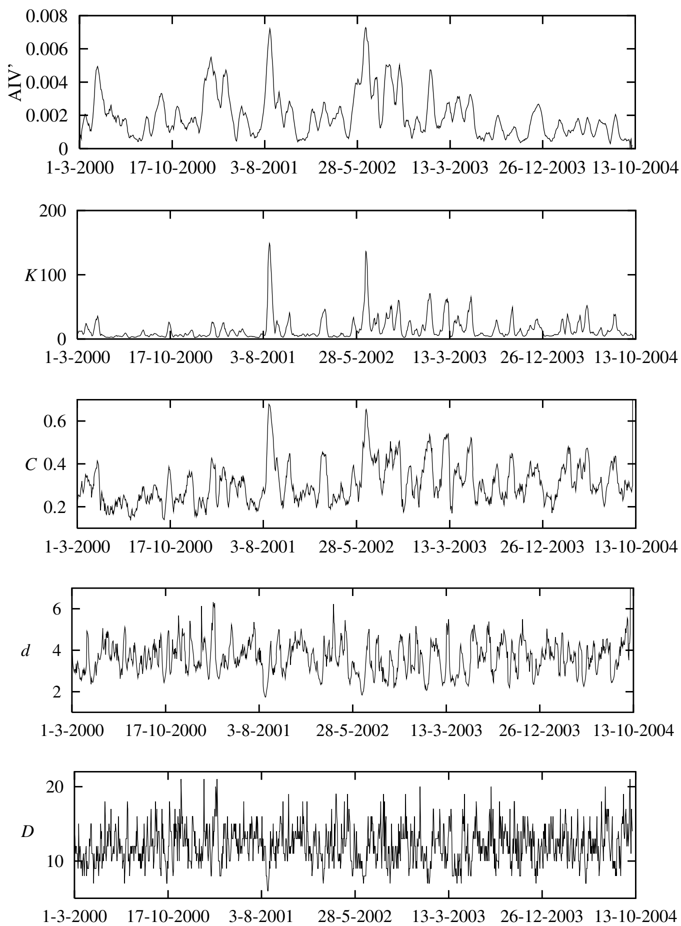

Fig. 3. Times series of AIV′ and parameters of the closing price network. K is average degree, C is average clustering coefficient, d is average shortest distance and D is diameter. Their cross correlations are given in Table 1.

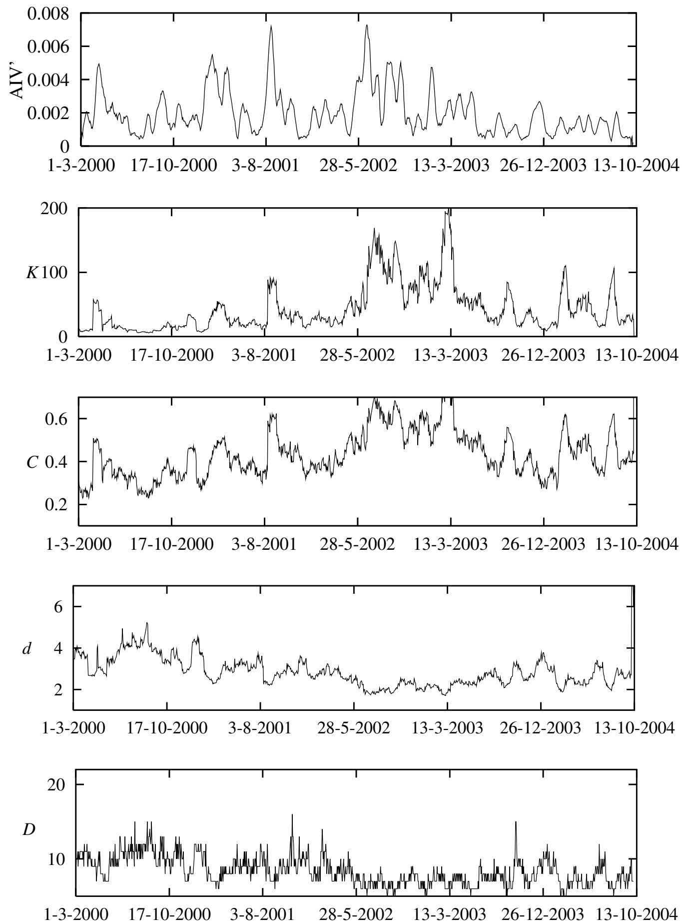

Fig. 4. Time series of AIV′ and parameters of the price return network. K is average degree, C is average clustering coefficient, d is average shortest distance and D is diameter. Their cross correlations are given in Table 1.

when the market fluctuation is fierce, the market internal structure becomes highly interwoven, resulting in an increase in the edge number and clustering coefficient.

4 Disruption of Scalefreeness of Degree Distribution under Fierce Market Fluctuation

Variation of the network structure is particularly interesting. As mentioned earlier, the degree distributions for the networks constructed from the closing prices and price returns have been found to be scalefree-like [9]. In this section we

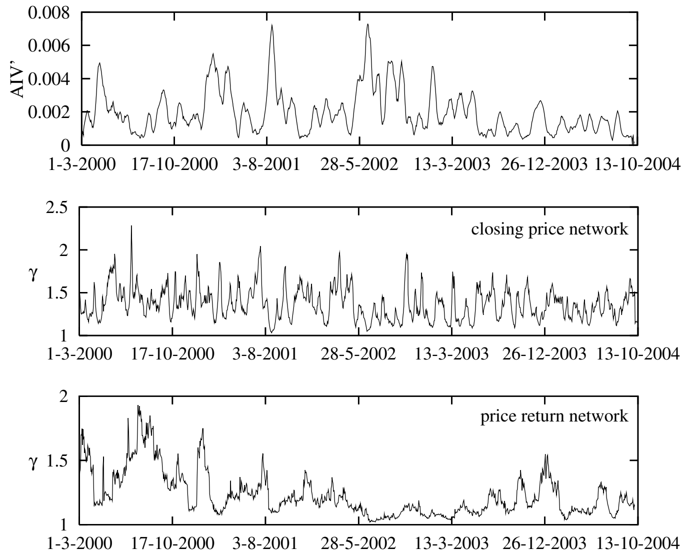

Fig. 5. Time series of AIV′ and power-law exponent γ for closing price network and price return network

examine the variation of the power-law exponents (denoted by γ) and the corresponding fitting error as time elapses, again using the moving 20-day window network, and evaluate their cross correlations with AIV′. See Table 1 for numerical results. Moreover, our study has shown a rather striking phenomenon, which relates to the disruption of the scalefree-like structure of the network under fierce market fluctuation.

By assuming the power law degree distributions of the constructed 20-day window networks, we apply the least mean square fitting method on the cumulative degree distributions of these networks to obtain the power law exponents and the corresponding fitting errors.

Clearly the fitting error is a measure of how close the empirical distribution is to the theoretical power-law distribution. If the degree distribution deviates significantly from the power-law distribution, the fitting error becomes large. Therefore, we expect the fitting error to reflect the scalefreeness of the network whose power-law degree distribution is being approximated.

As can be seen in Fig. 5, the variation of the power-law exponent γ has no observable resemblance to that of AIV′ for both closing price and price return networks. However, from Fig. 6, the variations of the fitting errors are found strongly correlated to the variation of AIV′, especially for closing price network. This clearly shows that the scalefreeness of the network is an important health-check indicator. The occurrence of spikes in fitting error variation corresponds to disruption of scalefreeness of the network, which in turn correlates strongly with fluctuation of the stock market.

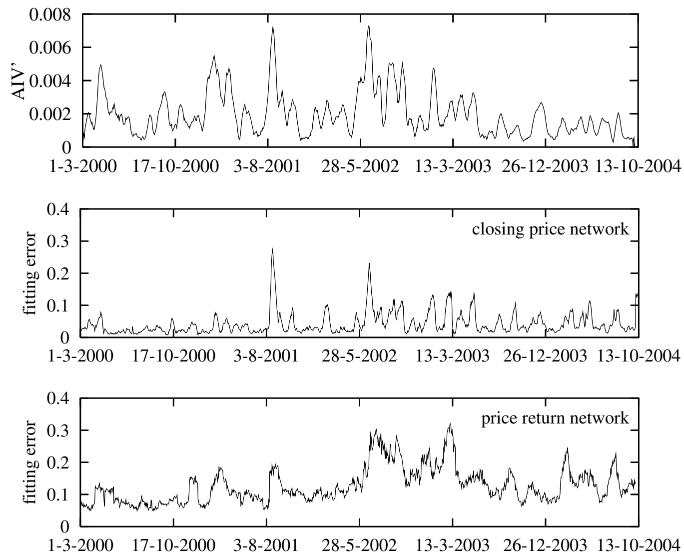

Fig. 6. Time series of AIV′ and power-law exponent fitting error. Large fitting error reflects poor "scalefreeness" approximation of the network structure. Correspondence is evidenced between market fluctuation and the disruption of scalefree-like structure.

5 Conclusion

In this work, we study the structural variation of networks formed by connecting S&P500 stocks based on cross correlation. The network is examined in a 20-day window, and as the window advances in time, we effectively capture the variation of the network properties including some network parameters and the scalefree-like structure. It has been shown that the market fluctuation, measured in terms of average index volatility, is strongly correlated with the scalefree-like structure of the network. Specifically we have shown that the scalefree-like structure, while being the default structure, is disrupted under fierce market fluctuation. It can therefore be concluded that the level of resemblance of scalefree-like structure of the stock network is an indicator of the normality of the market. An appropriate quantitative measure is the fitting error of the power-law exponent whose time series has been found highly correlated with that of the average index volatility.

Acknowledgment

This work was supported by Hong Kong Polytechnic University Research Project 1-BBZA.

References

1. Mantegna, R.N.: Hierarchical Structure in Financial Markets. Euro. Phys. J. B. 11, 193–197 (1999)
2. Vandewalle, N., Brisbois, F., Tordoir, X.: Self-organized Critical Topology of Stock Markets. Quan. Finance 1, 372–375 (2001)
3. Bonanno, G., Caldarelli, G., Lillo, F., Micciché, S., Vandewalle, N., Mantegna, R.N.: Networks of Equities in Financial Markets. Euro. Phys. J. B. 38, 363–371 (2004)
4. Bonanno, G., Lillo, F., Mantegna, R.N.: High-frequency Cross-correlation in a Set of Stocks. Quan. Finance 1, 96–104 (2001)
5. Bonnanno, G., Caldarelli, G., Lillo, F., Mantegna, R.N.: Topology of Correlation-based Minimal Spanning Trees in Real and Model Markets. Phys. Rev. E. 68, 046103 (2003)
6. Onnela, J.P., Chakraborti, A., Kaski, K.: Dynamics of Market Correlations: Taxonomy and Portfolio Analysis. Phys. Rev. E. 68, 056110 (2003)
7. Tumminello, M., Aste, T., Di Matteo, T., Mantegna, R.N.: A Tool for Filtering Information in Complex Systems. Proc. National Academy of Sciences USA 102(3), 10421–10426 (2005)
8. Onnela, J.P., Kaski, K., Kertesz, J.: Clustering and Information in Correlation Based Financial Networks. Euro. Phys. J. B. 38, 353–362 (2004)
9. Tse, C.K., Liu, J., Lau, F.C.M.: Winner-take-all Correlation-based Complex Networks for Modeling Stock Market. In: Proc. Int. Symp. Nonl. Theory and Its Appl., pp. 245–248 (2008)
10. Cohen, J., Cohen, P., West, S.G., Aiken, L.S.: Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences, 3rd edn. Lawrence Erlbaum Associates, New Jersey (2003)
11. Standard and Poor's, `http://www2.standardandpoors.com/`
12. Historical Data for S&P 500 Stocks, `http://biz.swcp.com/stocks/`
13. Goldstein, M.L., Morris, S.A., Yen, G.H.: Problems with Fitting to the Power-law Distribution. Euro. Phys. J. B. 41, 255–258 (2004)