Journal of the Korean Physical Society, Vol. 40, No. 6, June 2002, pp. 1105∼1108

Scale-Free Network in Stock Markets

H.-J. Kim and I.-M. Kim
Department of Physics, Korea University, Seoul 138-701

Y. Lee
Yanbian University of Science and Technology, Beishan St. Yanji, Jilin 133000, China

B. Kahng*
School of Physics and Center for Theoretical Physics, Seoul National University, Seoul 151-747

(Received 10 October 2001)

We study the cross-correlations in stock price changes among the S&P 500 companies by introducing a weighted random graph, where all vertices (companies) are fully connected via weighted edges. The weight of each edge is distributed in the range of $[-1, 1]$ and is given by the normalized covariance of the two modified returns connected, where the modified return means the return minus the mean over all companies. We define an influence-strength at each vertex as the sum of the weights on the edges incident upon that vertex. Then we find that the influence-strength distribution in its absolute magnitude $|q|$ follows a power-law, $P(|q|) \sim |q|^{-\delta}$, with exponent $\delta \approx 1.8(1)$.

PACS numbers: 89.75.Da, 89.65.Gh, 05.10.-a
Keywords: Econophysics, Stock-market

Recently complex systems such as biological, economic, physical, and social systems have received considerable attention as an interdisciplinary subject [1]. Such systems consist of many constituents, such as individuals, companies, substrates, spins, *etc.*, exhibiting cooperative and adaptive phenomena through diverse interactions between them. In particular, in economic systems, adaptive behaviors of individuals, companies, or nations, play a crucial role in forming macroscopic patterns such as commodity prices, stock prices, exchange rates, *etc.*, which are formed mostly in a self-organized way [2]. Recently, in physics communities, much attention and many studies have been focused and performed on the fluctuations and the correlations in stock price changes between different companies by applying physics concepts and methods [3,4].

Stock price changes of individual companies are influenced by others. Thus, one of the most important quantities in understanding the cooperative behavior in the stock market is the cross-correlation coefficients between different companies. Since stock-price changes depend on various economic environments, it is very hard to construct a dynamic equation and to predict the evolution of the stock-price change in the future. Recently, there have been many efforts to understand the correlations in stock-price changes between different companies by using a random matrix theory, where large eigenvalues are located far away from a bulk part predicted by the random matrix theory, reflecting the collective behavior of the entire market [5–7].

Let $Y_i(t)$ be the stock price of company i $(i = 1, \ldots, N)$ at time t. Then the return of the stock price after a time interval Δt is defined as

$$S_i(t) = \ln Y_i(t + \Delta t) - \ln Y_i(t), \tag{1}$$

which is the geometrical change of $Y_i(t)$ during the interval Δt. We take Δt as one day throughout the following analysis in this paper. The cross-correlations between individual stocks are considered in terms of a matrix $\mathbf{C}$, whose elements are given by

$$c_{i,j} \equiv \frac{\langle S_i S_j \rangle - \langle S_i \rangle \langle S_j \rangle}{\sqrt{(\langle S_i^2 \rangle - \langle S_i \rangle^2)(\langle S_j^2 \rangle - \langle S_j \rangle^2)}}, \tag{2}$$

where the brackets mean a temporal average over the period we studied. Then $c_{i,j}$ can vary between [-1,1], where $c_{i,j} = 1$ (-1) means that two companies i and j are completely correlated (anti-correlated), while $c_{i,j} = 0$ means that they are uncorrelated. Since the matrix $\mathbf{C}$ is symmetric and real, all eigenvalues are real, and the largest eigenvalue is not degenerate. It was found that the eigenvector corresponding to the largest eigenvalue is strongly localized at a few companies which strongly influence other companies in the stock price changes [5,6]. Moreover, an ultrametric hierarchical tree structure was

*E-mail: kahng@phya.snu.ac.kr

constructed among those companies, by using the concept of the minimum spanning tree in the graph theory, and that arranged those companies in order following their strengths of influence [8,9].

In this paper, we study further properties of the cross-correlations in stock-price changes by using a random graph theory [10]. The study of complex systems through a random graph was initiated by Erdös and Rényi (ER) [11]; however, the ER model is too random to describe real-world complex systems. Recently, Barabási and Albert (BA) [12] introduced an evolving network where the number of vertices increases with time rather than remaining fixed, and a newly introduced vertex is connected to m already existing vertices, following the so-called preferential attachment rule. The number of edges, k, incident upon a vertex is called the degree of the vertex. Then the degree distribution $P_D(k)$ follows a power-law $P_D(k) \sim k^{-\gamma}$ with $\gamma = 3$ for the BA model, while for the ER model, it follows a Poisson distribution. The network following a power-law in the degree distribution is called a scale-free (SF) network. Although natural phenomeno, such as river network, is based on a regular network [13], recently it was found that a variety of real world networks can be explained by such random SF networks.

For the problem of the correlations in stock-price changes, each vertex (edge) in the random graph corresponds to a company (the cross-correlation in stock price changes between the companies connected via that edge). The random graph generated in this way is different from a typical one in the following way: While the edge in a typical random graph has weight either 1 or 0, depending on if the edge is present, the edge in the random graph we will introduce has weight $w_{i,j}$, which is rather distributed between [-1,1]. For further studies, the random graph for the former (latter) case is called a binary random graph (BRG) (weighted random graph (WRG)). While the WRG can be found easily in real-world networks such as neural networks, cardiovascular networks, and respiratory networks in biological systems, acquaintance networks in social systems, etc, it has been studied less than the BRG [14,15]. In the WRG, one may wonder if the edge with less weight, called the weak edge, can be ignored in considering the correlations; however, there have been ongoing discussions about the importance of weak edges, for example, in a social acquaintance network, the scientific collaboration web, and ecosystems [16,17].

Recently, Yook, Jeong and Barabási (YJB) introduced a WRG [17]. In that model, a vertex i is newly introduced at each time step, connecting to the m vertices already existing according to the so-called preferential attachment rule. The edge connecting from the vertex i to an existing vertex j is assigned a weight $w_{i,j}$, depending on the degree of the vertex j. The weight at each vertex is assigned as the sum of the weights on the edges incident upon that vertex, which follows a power-law distribution, $P(q) \sim q^{-\delta}$, where q means the weight at a vertex. Then the exponent δ is different from the degree exponent γ and turns out to strongly depend on the mean degree m. While the YJB model is meaningful as a first step towards understanding diverse WRGs in the real world in a simple way, it still remains a theoretical network. The YJB graph is different from ours; While the YJB graph is weighted but sparsely connected, our WRG are fully connected and weighted.

We consider the cross-correlations in stock-price changes between the S&P 500 companies during 5-year period 1993-1997. Thus, the $N = 500$ companies correspond to 500 vertices, which are fully connected to each other with $N(N-1)/2$ edges. Each edge is assigned a weight, $w_{i,j}$ $(i, j = 1, \ldots, N)$, which is slightly modified from the cross-correlation coefficient $c_{i,j}$ defined in Eq. (2). Before defining $w_{i,j}$ specifically, we first recall some properties exhibited by $c_{i,j}$. It is known that the distribution of the coefficients $\{c_{i,j}\}$ forms a bell-shaped curve and that the mean value of the distribution is slowly time-dependent while the standard deviation is almost constant [18]. The time-dependence of the mean value might be caused by external economic environments such as bank interest, the inflation index, the exchange rate, *etc.*, which fluctuates from time to time. To extract intrinsic properties of the correlations in stock price changes, we introduce a quantity,

$$G_i(t) = S_i(t) - \frac{1}{N}\sum_i S_i(t), \tag{3}$$

where $G_i(t)$ indicates the relative return of a company i to its mean value over the entire 500 companies at time t. The cross-correlation coefficients are redefined in terms of G_i,

$$w_{i,j} \equiv \frac{\langle G_i G_j\rangle - \langle G_i\rangle\langle G_j\rangle}{\sqrt{(\langle G_i^2\rangle - \langle G_i\rangle^2)(\langle G_j^2\rangle - \langle G_j\rangle^2)}}. \tag{4}$$

The cross-correlation coefficients $\{w_{i,j}\}$ are assigned to each edge of the WRG as its weight. In order to check if the distribution is time-independent, we take the temporal average in Eq. (4) over each year from 1993 to 1997. In Fig. 1, we plot the distributions of $\{w_{i,j}\}$ obtained for each year. The data for different years are indeed overlapped, and time-independent. Therefore we think that the cross-correlation coefficients $\{w_{i,j}\}$ we introduced are appropriate to study intrinsic properties of the cross-correlations among the 500 S&P companies.

We define the influence-strength q_i at a vertex i as the sum of the weights on the edges incident upon the vertex i; that is,

$$q_i = \sum_{j\neq i} w_{i,j}, \tag{5}$$

where j denotes the vertices connected to the vertex i. Here, $\{w_{i,j}\}$ was obtained numerically by temporal averaging over the 5 years in Eq. (4). Then the weight q_i means the net amount of influence-strength for

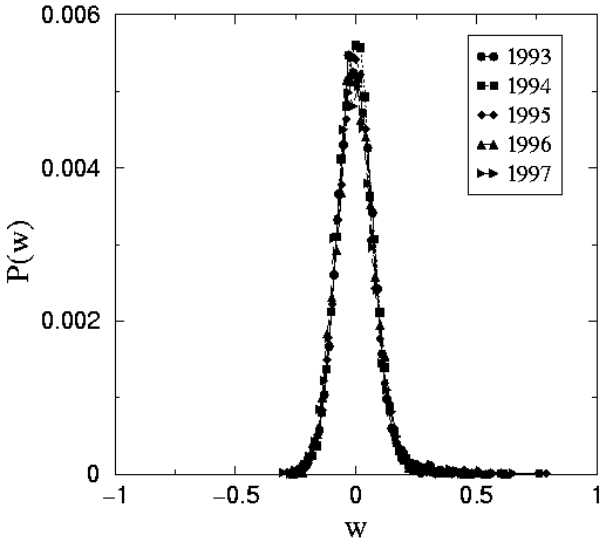

Fig. 1. Plot of the distribution of the cross-correlation coefficients $\{w_{i,j}\}$. The data are obtained by temporal averaging over each year from 1993 to 1997.

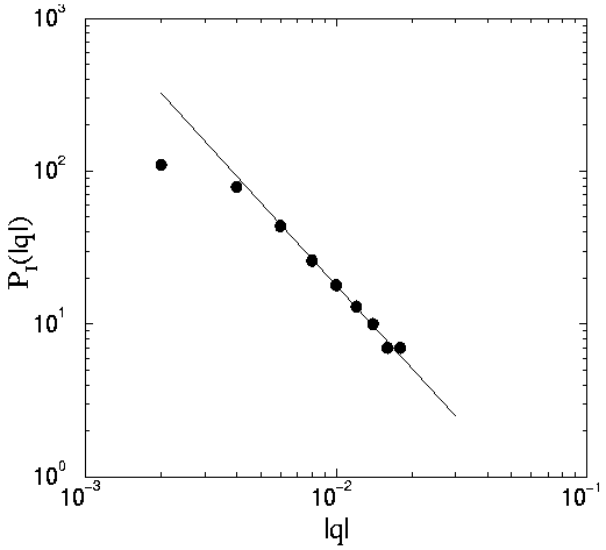

Fig. 2. Log-log plot of the influence-strength distribution $P_I(|q|)$ versus the absolute magnitude of the influence-strength $|q|$. The solid guideline has a slope -1.8.

the company i to affect other companies in stock-price changes. Since the weight $w_{i,j}$ is distributed between [-1,1], the influence-strength at a certain vertex could be negative. Thus, we deal with the absolute magnitude of the influence-strength for each vertex. In Fig. 2, we plot the influence-strength distribution $P_I(|q|)$ as a function of $|q|$, which turns out to follow a power-law, $P_I(|q|) \sim |q|^{-\delta}$. The exponent δ is estimated to be $\delta \approx 1.8(1)$. Thus the cross-correlations in stock-price changes forms a SF network, in particular a weighted SF network. To our knowledge, this is the first observation of SF WRG emerging in real-world economic systems. The presence of a scaling in the cross-correlations through the WRG could be related to the ultrametric hierarchical tree structure, implying a few companies exist having strong influence in stock price changes. On the other hand, it is easy to see that as the degree exponent in SF networks is smaller, the connectivity to the hub, the vertex with the largest degree, is higher, and the network is much centralized. This fact is also applicable to the SF WRG we introduced. Since the influence-strength exponent δ is smaller than 2 in the WRG, the vertex having largest influence-strength plays a much more important role in affecting stock-price changes of other vertices, compared with the role of the hub in the Internet [19], the world-wide web [20] and the metabolic network [21], where the degree exponent is greater than 2. We think that this result reflects economic systems being much more correlated and adaptive to achieve high profits. Thus the rich-get-richer phenomenon appears much strongly in economic systems than in the information systems. In contrast, we can expect that a simple drop in the stock price occurring in one of most influential companies could lead to a crash in the entire stock market.

It would be interesting to compare our WRG with the BRG constructed via the minimum spanning tree of Vandewalle *et al.* [22], To be specific, they considered cross-correlation coefficients defined in Eq. (2) in stock price changes between the 6 358 US companies, and constructed a minimum spanning tree structure. They measured the degree distribution for this BRG, which follows a power-law, $P_D(k) \sim k^{-\gamma}$, with $\gamma \approx 2.2$, which is obviously different from our results, $\delta \approx 1.8(1)$.

In conclusion, we have considered the cross-correlations in stock-price changes between the S&P 500 companies by introducing a weighted random graph (WRG). The vertices of the WRG representing the 500 companies are fully connected to each other through weighted edges. The edge connecting vertices i and j have weights given by the correlation coefficient $w_{i,j}$, defined as the normalized covariance of the modified returns of the two companies i and j. Here the modified return of a company i means the deviation of the return of the company i from its average over all 500 companies. This modification yields the effect of excluding the overall behavior of the entire stock prices fluctuating from time to time. The distribution of the correlation coefficients obtained using the modified return is time-independent, and the coefficients themselves describe generic correlations between different companies without considering the effect of external environments. We defined the influence-strength at each vertex as the sum of the weights assigned to the edges incident upon that vertex. We found that the influence-strength distribution follows a power-law $P_I(|q|) \sim |q|^{-\delta}$ with $\delta \approx 1.8(1)$, where q means influence-strength. The fact that the exponent δ is smaller than 2 implies that the stock-price changes of the 500 companies are much strongly corre-

lated than they are with the Internet topology or the world-wide web, reflecting that cooperative and adaptive phenomena appear dominantly in economic systems.

ACKNOWLEDGMENTS

This work is supported by Korean Research Foundation grant (KRF-2001-015-DP0120) and by the Ministry of Education through the BK21 project in KU, and by grants No.R01-2000-00023 from the BRP program of the KOSEF through SNU.

REFERENCES

[1] N. Goldenfeld and L. P. Kadanoff, Science **284**, 87 (1999).
[2] W. B. Arthur, Science **284**, 107 (1999).
[3] R. N. Mantegna and H. E. Stanley, *An Introduction to Econophysics : Correlations and Complexity in Finance* (Cambridge Univ., Cambridge, 2000).
[4] J-P. Bouchaud and M. Potters, *Theory of Financial Risks: From Statistical Physics to Risk Management*, (Cambridge Univ., Cambridge, 2000).
[5] L. Laloux, P. Cizeau, J.-P. Bouchaud and M. Potters, Phys. Rev. Lett. **83**, 1467 (1999).
[6] V. Plerou, P. Gopikrishnan, B. Rosenow, L. A. N. Amaral and H. E. Stanley, Phys. Rev. Lett. **83**, 1471 (1999).
[7] J. D. Noh, Phys. Rev. E **61**, 5981 (2000).
[8] R. N. Mantegna, Eur. Phys. J. B **11**, 193 (1999).
[9] G. Bonanno, N. Vandewalle and R. N. Mantegna, Phys. Rev. E **62**, R7615 (2000).
[10] B. Bollobás, *Random Graphs* (Academic, London, 1985).
[11] P. Erdös and A. Rényi, Publ. Math. Inst. Hung. Acad. Sci. Ser. A **5**, 17 (1960).
[12] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
[13] H.-J. Kim and In-mook Kim, J. Korean Phys. Soc. **38**, 497 (2001); H.-J. Kim, In-mook Kim and J. M. Kim, Phys. Rev. E **62**, 3121 (2000).
[14] R. Albert and A-L. Barabási, (cond-mat/0106096).
[15] S. N. Dorogovtsev and J. F. F. Mendes, (cond-mat/0106144).
[16] E. L. Berlow, Nature **398**, 330 (1999).
[17] S. H. Yook, H. Jeong, A.-L. Barabási and Y. Tu, Phys. Rev. Lett. **86**, 5835 (2001).
[18] R. N. Mantegna, in *Applied Nonlinear Dynamics and Stochastic Systems near the Millennium*, edited by J. B. Kadtke and A. Bulsara (AIP Press, New York, 1997).
[19] M. Faloutsos, P. Faloutsos and C. Faloutsos, Comp. Comm. Rev. **29**, 251 (1999).
[20] R. Albert, H. Jeong and A.-L. Barabási, Nature **401**, 130 (1999).
[21] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvani and A.-L. Barabási, Nature **407**, 651 (2000).
[22] N. Vandewalle, F. Brisbois and X. Tordoir, (cond-mat/0009245).