Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Physica A 370 (2006) 151–155
PHYSICA A
www.elsevier.com/locate/physa

Correlation in business networks

Wataru Souma[a,*,1], Hideaki Aoyama[b], Yoshi Fujiwara[a,1], Yuichi Ikeda[c], Hiroshi Iyetomi[d], Taisei Kaizoji[e]

[a]ATR Network Informatics Laboratories, Kyoto 619-0288, Japan
[b]Department of Physics, Kyoto University, Kyoto 606-8501, Japan
[c]Hitachi Research Institute, Tokyo 101-8010, Japan
[d]Department of Physics, Niigata University, Niigata 950-2181, Japan
[e]Division of Social Science, ICU, Tokyo 181-8585, Japan

Available online 12 May 2006

Abstract

This paper considers business networks. Through empirical study, we show that business networks display characteristics of small-world networks and scale-free networks. In this paper, we characterize firms as sales and bankruptcy probabilities. A correlation between sales and a correlation between bankruptcy probabilities in business networks are also considered. The results reveal that the correlation between sales depends strongly on the type of network, whereas the correlation between bankruptcy probabilities does so only weakly.

Keywords: Business networks; Small-world; Scale-free; Correlation

1. Introduction

Recently, many studies have revealed the true structure of real-world networks. This development also holds true in the field of econophysics. Such studies have investigated business networks, world trade networks, and corporate board networks (see Ref. [1]). Though investigating topological properties of networks is important, we consider the study of a correlation between nodes in networks to also be important. Hence in this paper, the subjects we address are topological characteristics of business networks and correlation between firms in business networks.

This paper is organized as follows. Section 2 explains characteristics of data sets analyzed in this article. In Section 3, we calculate the degree distribution, the averaged path length, and the clustering coefficient for business networks. In this paper, business networks indicate a transaction network, a shareholding network,

*Corresponding author.
E-mail address: souma@atr.jp (W. Souma).
[1]W. Souma and Y. Fujiwara are supported in part by the National Institute of Information and Communications Technology. We are also supported in part by a Grant-in-Aid for Scientific Research (#15201038) from the Ministry of Education, Culture, Sports, Science and Technology.

doi:10.1016/j.physa.2006.04.040

and an intersection of them. Following that, in Section 4 we consider the correlation between sales and the correlation between bankruptcy probabilities in business networks. The last section is devoted to a summary.

2. Data sets

In this paper, we use transaction data, shareholding data, and securities data. These data are established in 2004. Transaction data is sold by Teikoku Databank Ltd. [2], and contains lists of suppliers and customers for firms. We obtained this data for 1405 firms, which are listed on the first section of the Tokyo Stock Exchange and belong to the non-financial industry. If we draw arrows according to the physical distribution, we can see the transaction network as a directed one. This network features 17,302 nodes and 54,345 links.

In this paper we use shareholding data sold by Toyo Keizai Inc. [3]. This data includes lists of shareholders of firms listed on the stock market or on the over-the-counter market. If we draw arrows from shareholders to stock-issuing corporations, we can represent the shareholding network as a directed one. Of course, the lengths of shareholder lists vary from firm to firm, and the most comprehensive lists contain the top 30 shareholders. Therefore the incoming degree has an upper bound, while the outgoing degree has no bound. This network includes approximately 4000 nodes and 30,000 links.

The securities data contain lists of securities held by firms listed on the stock market or on the over-the-counter market. These lists are available to the public, and we can obtain them from the electronic disclosure for investors' network (EDINET) [4]. The lengths of the securities lists vary according to the firms, though a typical list contains information on the top 10 or 20 securities. Securities are approximately equivalent to shares of stock-issuing corporations. Hence, the outgoing degree has an upper bound, while the incoming degree has no bound. The size of this network is approximately equal to the network constructed from shareholding data.

The transaction network is constructed from the transaction data only, while the shareholding network is constructed from both shareholding data and securities data. We can also obtain the intersection of these networks, and simply call it the intersection network. In this article, we ignore the directions of edges in the intersection network. Though entities in a network vary with the network type, one of the purposes of this article is to consider the effect of business networks on the correlation between firms. It is therefore preferable to match entities in these networks. Here we adopt the 1405 firms mentioned above as a minimal set of nodes.

3. Characteristics of business networks

Though many quantities have been proposed to characterize networks (for a review, see Ref. [6]), we simply consider the degree distribution, the averaged path length, and the clustering coefficient, which are the basic quantities needed to understand the characteristics of networks. Fig. 1(a)–(c), respectively, show log–log plots

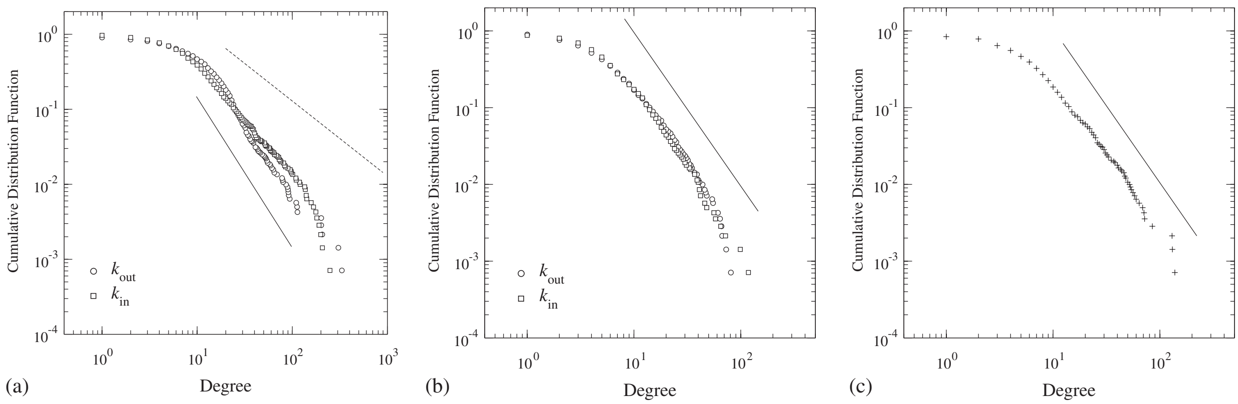

Fig. 1. A log–log plot of the cumulative degree distribution (a) in the transaction network, (b) in the shareholding network, and (c) in the intersection network.

of the cumulative degree distribution in the transaction network, in the shareholding network, and in the intersection network. In these figures the horizontal axis is the degree and the vertical axis represents the cumulative distribution. Open circles and open squares in Fig. 1(a) and (b) denote the distribution of the outgoing degree and the incoming degree. Because we ignore the edge direction in the intersection network, there is one type of distribution in Fig. 1(c). The dashed line in Fig. 1(a) corresponds to the power-law function with the exponent $\gamma = 2$, and the solid lines in Fig. 1 correspond to that with $\gamma = 3$. From these diagrams, we can understand that the tail part of these distributions approximately follow the power-law distribution with $\gamma = 2$–3, though the exponent depends on the fitting range. Hence business networks display characteristics of scale-free networks.

The averaged path length and the clustering coefficient are $L^{\mathrm{T}} = 2.563$ and $C^{\mathrm{T}} = 0.209$ in the transaction network. In a regular network the same size as the transaction network, $L^{\mathrm{reg.}} = 29.92$ and $C^{\mathrm{reg.}} = 0.719$, while in a random network the same size as the transaction one, $L^{\mathrm{rand.}} = 2.228$ and $C^{\mathrm{rand.}} = 0.019$. Thus we find that $L^{\mathrm{T}} \approx L^{\mathrm{rand.}} \ll L^{\mathrm{reg.}}$ and $C^{\mathrm{T}} \approx C^{\mathrm{reg.}} \gg C^{\mathrm{rand.}}$, indicating that the transaction network has characteristics of small-world networks. In the shareholding network $L^{\mathrm{S}} = 2.563$ and $C^{\mathrm{S}} = 0.209$; the size of this network is approximately the same as that of the transaction network. Hence, the shareholding network also has characteristics of small-world networks. This result is applicable to the case of the intersection network.

4. Correlation in business networks

We characterize firms as sales and bankruptcy probabilities, and consider the correlation between sales and the correlation between bankruptcy probabilities in business networks. In this paper, we use the simple analysis of failure proposed in 2002 (SAF2002) as bankruptcy probabilities [5]. The definition of SAF2002 will be provided later.

A correlation matrix R_{ij} is given by $R_{ij} = [\langle x_i x_j \rangle - \langle x_i \rangle \langle x_j \rangle]/\sigma_i \sigma_j$, where $\langle \cdot \rangle$ represents a time average and σ_i represents a standard deviation. A correlation matrix corresponds to a complete graph in which every node is connected to every other node including itself. For example, if we regard $x_i(t)$ as sales for ith firms in the last five years, we can obtain the distribution of correlation coefficients shown in Fig. 2(a). In this figure the horizontal axis is the correlation coefficient and the vertical axis is the probability density. To obtain this figure, we ignore self-loops, i.e., diagonal components in the correlation matrix. If sales are completely random, the distribution of correlation coefficients will have a peak at $R_{ij} = 0$ and decay rapidly. However, as shown in Fig. 2(a), the empirical fact is contrary to this. In fact, the distribution has two peaks around $R_{ij} = \pm 1$ and a valley around $R_{ij} = 0$. This means that a strong and positive correlation or a strong and negative correlation are frequently observed in the correlation between sales in the real economy. The average correlation coefficient is $\overline{R} = 0.079$.

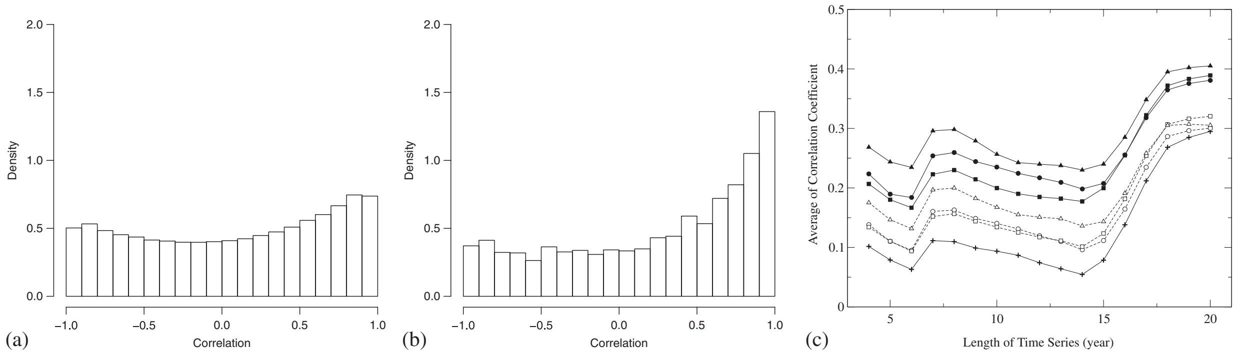

Fig. 2. (a) A distribution of correlation coefficients in the correlation matrix, and (b) in the one-link correlation in the intersection network. (c) Changes in the average values of correlation coefficients.

Now we define an n-link correlation matrix as

$$R_{ij}^{\mathrm{K}(n)} = \begin{cases} R_{ij} & \text{if } i \text{ and } j \text{ are connected by } n\text{-link}, \\ \text{null} & \text{otherwise}, \end{cases}$$

where $\mathrm{K} \in \{T, S, T \cap S\}$. Here T denotes the transaction network, S is the shareholding network, and $\mathrm{T} \cap \mathrm{S}$ represents the intersection network. In this article, we assume that the network is static. If we regard $x_i(t)$ as sales for ith firms in the last five years and calculate $R_{ij}^{\mathrm{T}\cap\mathrm{S}(1)}$, we obtain Fig. 2(b). This figure shows that the distribution has one peak around $R_{ij}^{\mathrm{T}\cap\mathrm{S}(1)} = 1$, and gradually decreases as the correlation decreases. The average correlation coefficient is $\overline{R^{\mathrm{T}\cap\mathrm{S}(1)}} = 0.244$. Distributions of $R_{ij}^{\mathrm{T}(1)}$ and $R_{ij}^{\mathrm{S}(1)}$ interpolate two extreme cases, i.e., Fig. 2(a) and (b), and their respective averages are $\overline{R^{\mathrm{T}(1)}} = 0.190$ and $\overline{R^{\mathrm{S}(1)}} = 0.180$. From these empirical facts, we can conclude that business networks increase the correlation between sales, and that the correlation in an intersection network is stronger than that in other networks.

If we change the length of the time series and calculate the average value of the correlation coefficients, we obtain Fig. 2(c). In this figure, the horizontal axis is the length of the time series and the vertical axis is the average value of the correlation coefficients. The solid line with crosses is for $\overline{R}$, that with filled circles is for $\overline{R^{\mathrm{T}(1)}}$, that with filled squares is for $\overline{R^{\mathrm{S}(1)}}$, and that with filled triangles is for $\overline{R^{\mathrm{T}\cap\mathrm{S}(1)}}$. This figure shows that the one-link correlation in the intersection network is stronger than that in other networks throughout the period. Since we use the nominal value of sales in this article, fluctuations in the business cycle will affect the correlation. This is the origin of the fluctuations observed in Fig. 2(c). We must remove the effect of business fluctuations to determine the correlation without fluctuations, and there are two methods to accomplish it. One is to use the real value of sales. The other is to remove the effect of the primary eigenvalue.

We can also consider two-link correlations. In Fig. 2(c), the dashed line with open circles is for $\overline{R^{\mathrm{T}(2)}}$, that with open squares is for $\overline{R^{\mathrm{S}(2)}}$, and that with open triangles is for $\overline{R^{\mathrm{T}\cap\mathrm{S}(2)}}$. This figure shows that the two-link correlation in the intersection network is also stronger than that in other networks throughout the period. This figure also shows that two-link correlations are weaker than one-link correlations throughout the period. The presence of competitors will reduce correlation, and the probability of finding competitors two links ahead is higher than that for one link ahead. This is one reason why two-link correlations are weaker than one-link ones.

Now we consider the correlation between SAF2002s. SAF2002 is given by

$$\mathrm{SAF2002} = 0.0104x_1 + 0.0268x_2 - 0.0661x_3 - 0.0237x_4 + 0.7077.$$

Here x_1 denotes a retained earnings to total assets, x_2 is net income before tax to total assets, x_3 represents an inventory turnover period, and x_4 is interest expenses to sales.

Though it is not clear as to what length of time SAF2002 can be applied, we consider it for the last five years. In this case, the average value of the correlation matrix is $\overline{R} = 0.0863$. On the other hand, average values for the one-link correlation are $\overline{R^{\mathrm{T}(1)}} = 0.1345$, $\overline{R^{\mathrm{S}(1)}} = 0.1614$, and $\overline{R^{\mathrm{T}\cap\mathrm{S}(1)}} = 0.1780$. In the case of two-link correlation, these are $\overline{R^{\mathrm{T}(2)}} = 0.1136$, $\overline{R^{\mathrm{S}(2)}} = 0.1334$, and $\overline{R^{\mathrm{T}\cap\mathrm{S}(2)}} = 0.1677$. From these results, we understand that the correlation in the transaction network is stronger than that in other networks. However, in comparison to the case of sales, the difference of correlation between different networks is small. In addition, in comparison to the case of sales, the difference of correlation between one-link correlations and two-link correlations is small. There are two perspectives from which to view these results: one relates to the characteristics of SAF2002, which is constructed through analysis of financial statements; therefore, correlation between firms is not considered. The other relates to the characteristics of the bankruptcy chain. Our findings suggest that bankruptcy will spread regardless of the type of business network, and without distinguishing one link ahead from two links ahead.

5. Summary

In this article, we showed that business networks possess the characteristics of small-world networks and scale-free networks. The correlation between sales and the correlation between bankruptcy probabilities in business networks were considered. The results showed that the correlation between sales strongly depends on the type of network, whereas the correlation between bankruptcy probabilities does so only weakly. Though the study of economics from the viewpoint of network science has just begun, we believe that it will bring to light new knowledge to economics and econophysics.

References

[1] W. Souma, Y. Fujiwara, H. Aoyama, in: Practical Fruits of Econophysics, Springer, Tokyo, 2006, pp. 307–311.
[2] Teikoku Databank Ltd., URL: ⟨http://www.tdb.co.jp/⟩ (in Japanese) or ⟨http://www.tdb.co.jp/english/index.html⟩ (in English).
[3] Toyo Keizai Inc., URL: ⟨http://www.toyokeizai.co.jp/⟩ (in Japanese) or ⟨http://www.toyokeizai.co.jp/english/index.html⟩ (in English).
[4] EDINET. URL: ⟨http://info.edinet.go.jp/EdiHtml/main.htm⟩ (in Japanese).
[5] Y.C. Shirata, The bankruptcy prediction model-SAF2002 Model, Chuo-Keizai Sha, 2003 (in Japanese).
[6] L. da F. Costa, et al., arXiv:cond-mat/0505185.