Application of Network Analysis Techniques for Japanese Corporate Transaction Network

Kouhei Sugiyama Osamu Honda Hiroyuki Ohsaki Makoto Imase

Graduate School of Information Science and Technology, Osaka University, Japan
{k-sugi,o-honda,oosaki,imase}@ist.osaka-u.ac.jp

Abstract

In recent years, several social activities have been shifting into networked virtual society called *Cybersociety* because of rapid advancement of network technologies. For designing and realizing a new service in Cybersociety, it is crucial to fully investigate desired structure of commercial transactions and circulation activities in Cybersociety. For this purpose, it must be fruitful to investigate the structure of business relationships among existing corporations in real society, which are believed to be working effectively. In this paper, by utilizing a corporate information database, we first build a corporate transaction network based on dealings information among Japanese corporations. We then investigate structure of the corporate transaction network by applying several network-analysis techniques, which have been used for analyzing various social networks. We show that the average distance of the corporate transaction network is approximately 3.4–5.4, and that it has a scale-free property with a power-law index of 2.1. We also show that our proposed structural superiority index of the corporate transaction network is greatly related to several stock price indices and profit financial indices.

1 Introduction

In recent years, several social activities have been shifting into networked virtual society because of rapid advancement of network technologies. It is expected that virtual society called *Cybersociety* will be emerged on a network. Since several commercial transactions and circulation activities will be performed in Cybersociety, realization of new services that can improve efficiency and productivity of those activities is strongly demanded.

For designing and realizing a new service in Cybersociety, it is crucial to fully investigate desired structure of commercial transactions and circulation activities in Cybersociety. We expect that business transactions and circulation activities in Cybersociety will be performed between several types of entities, such as individuals and corporations.

For this purpose, it must be fruitful to investigate the structure of business transactions among corporations in real society, which are believed to be working effectively. Once the structure of such business transactions becomes clear, it is expected that several useful services, such as offering business transactions to talented individuals and predicting growth of corporations in Cybersociety, can be realized in Cybersociety. However, it has not been sufficiently investigated how such business transactions should be developed in Cybersociety.

For clarifying the desired structure of business transactions in Cybersociety, it is necessary to clarify the structure of business transactions in real society. For instance, analyzing the structure of corporate transactions in real society will give us knowledge on the desired structure of business transactions in Cybersociety.

In recent years, researches on social networks, which are networks of human communications, have been actively performed [1–3]. Moreover, analyses on the structure of real networks (*network analysis*), such as a WWW hyperlink network and a co-authorship network of journal papers, have been actively performed [4–6]. For instance, [4] analyzes the structure of a movie actors network. It is reported in [4] that the movie actors network has a small network distance and clustered structure. Also, [5] analyzes the structure of a mathematics co-authorship network. It is reported in [5] that the mathematics co-authorship network has a small average degree and clustered structure.

These social networks have common characteristics that their network structure are considerably different from that of a random network. For instance, these networks have clustered structure; i.e., the network distance is small (*small-world network*) [4] and their degree distributions follow a power-law (*scale-free network*) [7]. Researches on revealing reason that unlike a random network, those social networks have a small-world property and/or a scale-free property have been performed [6, 8–10]. For instance, a simple network model for reproducing characteristics of social network is proposed in [8]. This model mimics the following characteristics of social networks: (1) a person with many existing human relations is likely to acquire a new human relation, (2) since it is costly for a person to maintain existing human relations, there is an upper-bound in the number of human relations that every person can have, and (3) the strength of human relations decays with time (i.e., they will be diminished after a certain period).

As explained above, researches on structure of several social networks and models reproducing the characteristics of social networks have been actively performed. However, when business transactions among individuals and corporations are viewed as a network, the structure of such business transactions has not been sufficiently investigated.

In this paper, by utilizing a corporate information database, we first build a corporate transaction network based on dealings information among Japanese corporations. We then investigate structure of the corporate transaction network by applying several network-analysis techniques, which have been used for analyzing various other social networks.

Specifically, for building a corporate transaction network, we first obtain information on existence of corporate transactions between real Japanese corporations from the corporate information database [11]. We then investigate the fundamental characteristics of the corporate transaction network, such as the number of nodes, the number of links, and average degree. By applying network-analysis techniques, we also investigate detailed charac-

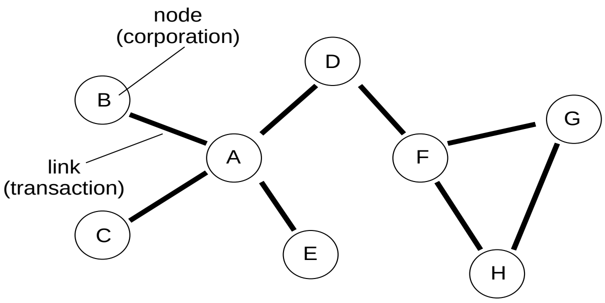

Figure 1: Example of a corporate transaction network

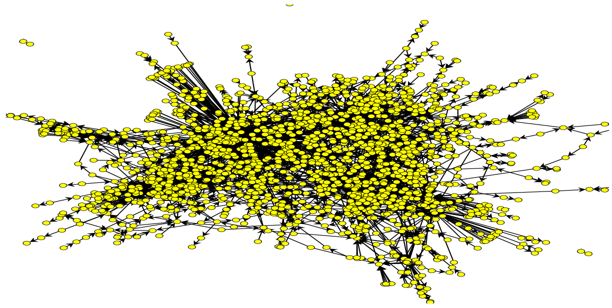

Figure 2: Graphical representaion of the corporate transaction network (1st quater of 2004)

teristics of the corporate transaction network: average distance, clustering coefficient, degree distribution, spectrum density, and structural superiority. We show that the average distance of the corporate transaction network is approximately 3.36–5.36, and that it has a scale-free property with a power-law index of 2.06. We also show that our proposed structural superiority index of the corporate transaction network is greatly related to several stock price indices and profit financial indices.

The organization of this paper is as follows. First, in Section 2, we define the corporate transaction network. We explain how the corporate transaction network is built from the corporate information database. In Section 3, average distance, clustering coefficient, degree distribution, spectrum density, and structural superiority index of the corporate transaction network are analyzed. By comparing with a random network or other social networks, we discuss the structure of corporate transactions. Finally, Section 4 concludes this paper and discusses future works.

2 Corporate Transaction Network

In this paper, a corporate transaction network is defined as a network that represents existence of transactions among corporations. Generally, each corporation performs several commercial transactions and circulation activities by dealing in products and services with other corporations. In a corporate transaction network, every node represents a corporation, and a link between nodes represents existence of transactions between nodes (i.e., dealings of products or services) between two corporations.

An example of a corporate transaction network is shown in Fig. 1. As shown in Fig. 1, a corporation is represented by a node, and a transaction between corporations is by a link between nodes. By representing corporate transactions as a network, it becomes possible to analyze the structure of corporate transactions.

It is an important problem how information on nodes (i.e., corporations) and links (i.e., existence of transactions) from real business transactions is obtained. It is easy to obtain information on nodes from, for example, a publicly available corporation database. On the other hand, it might be possible to obtain information on links from, for example, some documents or Web pages introducing the profile of a corporation, which sometimes includes a list of major customers. However, it is unrealistic to obtain information on links in such a way since profiles of all corporations are not available and those corporation profiles are difficult to be processed electronically.

In this paper, by utilizing the corporate information database [11], we obtain information of transactions among real corporations and build a corporate transaction network based on this information. This corporate information database is published quarterly and includes data for approximately 3,700 Japanese corporations. In addition to fundamental information, such as the corporate name, total assets, capital, stock price, the date of foundation, and the number of employees, this database contains a list of major customers.

Each corporation included in the corporate information database is assigned as a node in the corporate transaction network. Based on the list of major customers in the database, undirected links between nodes in the corporate transaction network are generated.

Note that, the list of major customers in the database does not contain information on trading direction of products and/or services (e.g., direction of cash flow). Therefore, we build the corporate transaction network as an undirected graph. This is because the corporate transaction from a corporation A to a corporation B should exist when the corporation B considers the corporation A to be a major customer. In the corporate transaction network, corporations having no transaction with other corporations in the database (i.e., isolated nodes) are removed.

In the corporate information database, for each corporation, not all customers but major customers are listed. For this reason, links of the corporate transaction network represent not all corporate transactions among corporations, but major corporate transactions among corporations. In reality, it is virtually impossible to obtain lists of all customers for all corporations. We believe that lists of major customers in the corporate information database is useful to reveal fundamental structure of the corporate transaction network since those lists represents *typical* transactions among corporations.

An example of corporate transaction network that we have built is graphically shown in Fig. 2. This figure represents the network built from the corporate information database [11] in the first quarter of 2004. This figure shows that the corporate transaction network forms a large cluster. In Tab. 1, we summarize fundamental characteristics of corporate transaction networks built from the corporate information database from the first quarter of 2004 to the first quarter of 2005. This table shows that the number of nodes N are 2,066–2,099, and average degrees $\overline{k}$ are 4.52–4.58. The network distances ℓ and clustering coefficients C are also shown in the table, which will be explained in Section 3. The network distance ℓ_{rand} and clustering coefficient C_{rand} of a random network, which has the same number of nodes N and the same average with the corporate transaction network, are also included in the table. Fundamental characteristics of the movie actors network [4] and the mathematics co-authorship network [5] are also included.

Table 1: Characteristics of corporate transaction networks and other social networks

	N	$\overline{k}$	ℓ	ℓ_{rand}	C	C_{rand}
corporate transaction network (1st quarter of 2004)	2,096	4.57	5.36	5.03	0.076	0.0022
corporate transaction network (2nd quarter of 2004)	2,089	4.56	5.19	5.04	0.075	0.0022
corpprate transaction network (3rd quarter of 2004)	2,099	4.58	4.02	5.03	0.073	0.0022
corporate transaction network (4th quarter of 2004)	2,066	4.55	3.76	5.04	0.073	0.0022
corporate transaction network (1st quarter of 2005)	2,078	4.52	3.36	5.07	0.072	0.0021
movie actors network [4]	225,226	61	3.65	2.99	0.79	0.00027
mathematics co-authorship network [5]	70,975	3.9	9.5	8.2	0.59	0.000054

3 Network Analysis

In this section, the corporate transaction network built in Section 2 is analyzed, and its characteristics are investigated. In what follows, for analyzing the structure of the corporate transaction network, we employ several metrics such as average distance, clustering coefficient, degree distribution, spectrum density, and structural superiority index, which have been developed in network-analysis techniques.

3.1 Average Distance

We first focus on the average distance ℓ of the corporate transaction network. We investigate distance between corporations from a network topological viewpoint. In other words, we investigate how close corporate transactions exist among corporations.

By letting N be the number of nodes, the average distance ℓ of a network is defined as [12]

$$\ell \equiv \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{\substack{j=1 \\ i \neq j}}^{N-1} \ell_{ij}, \tag{1}$$

where ℓ_{ij} is the shortest path length from node i to node j.

Let N be the number of nodes, and $\overline{k}$ be the average degree. It is known that the average distance ℓ_{rand} of a random network is approximately given by [6]

$$\ell_{rand} \quad \sim \quad \frac{\ln N}{\ln \overline{k}}. \tag{2}$$

Hence, if the corporate transaction network was a random network, its average distance would be 5.07 (in the case of the first quarter of 2005) (see Tab. 1). The average distance ℓ of the corporate transaction network is shown in Tab. 1. For instance, the average distance of the corporate transaction network in the first quarter of 2005 is $\ell = 3.36$.

This table shows that the average distance ℓ of the corporate transaction network is close to the average distance ℓ_{rand} of a random network (Fig. 3). Note that the average distance of 3.36–5.36 is smaller than that of the mathematics co-authorship network, which has a comparable average degree (Tab. 1). These results suggest that in the corporate transaction network, corporations are more closely related than nodes in other social networks such as the mathematics co-authorship network.

3.2 Clustering Coefficient

We investigate cluster structure of the corporate transaction network. The objective of this section is to answer the following questions: does each corporation belong to a specific group? If so, are the corporate transactions closed within those groups?

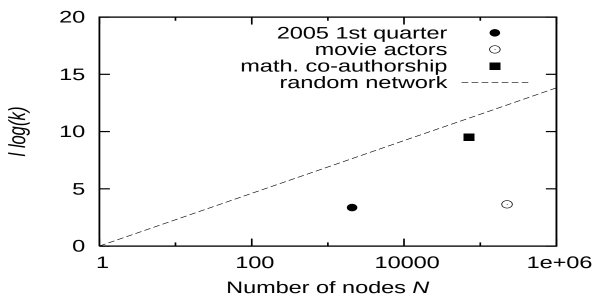

Figure 3: Average distance of the corporate transaction network (1st quarter of 2005) (dotted line is the case of random network)

For this purpose, we derive the clustering coefficient of the corporate transaction network. Clustering coefficient is a metric that measures cluster structure of nodes in a network. Specifically, clustering coefficient is defined as a probability that triangular cycle exists between each node pair. The clustering coefficient C of a network is defined as

$$C \equiv \frac{1}{N} \sum_{i=1}^{N} C_i, \tag{3}$$

where C_i is the clustering coefficient of node i, which is defined as

$$C_i \equiv \frac{2E_i}{k_i(k_i-1)}. \tag{4}$$

In the above equation, k_i is the degree of node i, and E_i is the total number of triangular cycles, which start from and return to node i.

Let N be the number of nodes, and $\overline{k}$ be the average degree. It is known that the clustering coefficient C_{rand} of a random network is given by [6]

$$C_{rand} \quad = \quad \frac{\overline{k}}{N}. \tag{5}$$

Hence, if the corporate transaction network was a random network, the clustering coefficient would be 0.0021 (in the case of the first quarter of 2005) (see Tab. 1). The clustering coefficient C of the corporate transaction network is shown in Tab. 1. For instance, the clustering coefficient of the corporate transaction network in the first quarter of 2005 is $C = 0.072$.

This table shows that the clustering coefficient C of the corporate transaction network is larger than the clustering coefficient C_{rand} of a random network by an order of magnitude (Fig. 4). This means that corporations tend to have more grouped structure in the corporate transaction network than in a random network. However, if compared with other social networks such as the movie actors network and the mathematics co-authorship network,

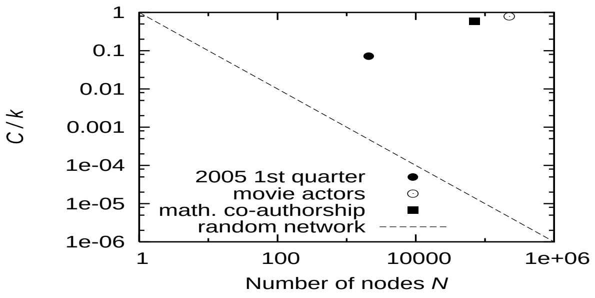

Figure 4: Clustering coefficient of the corporate transaction network (1st quarter of 2005) (dotted line is the case of random network)

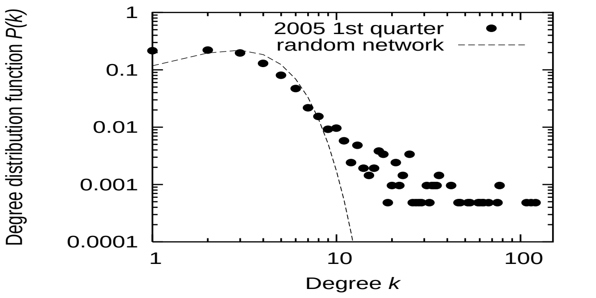

Figure 5: Degree distribution of the corporate transaction network (1st quarter of 2005) (dotted line is the case of random network)

the value of clustering coefficient is not so large (Tab. 1). These results indicate that the corporate transaction network tends to have less grouped structure than other social networks.

3.3 Degree Distribution

We then focus on degree distribution of the corporation transaction network. Namely, we investigate the distribution of the number of transactions of each node (i.e., corporation) in the corporation transaction network. The objective of this section is to answer the following questions: does any hub corporation, which has a very large number of business transactions with other corporations, exist in the corporate transaction network? does the corporate transaction network have a scale-free property?

The degree distribution is determined by the degree-distribution function $P(k)$. $P(k)$ is a probability that the degree of a randomly chosen node is k in a network.

When the number of nodes N is large, it is known that the degree-distribution function $P(k)$ of a random network with the average degree $\overline{k}$ approximately follows the Poisson distribution [6]; i.e.,

$$P(k) \simeq e^{-\overline{k}} \frac{\overline{k}^k}{k!}. \tag{6}$$

It is reported that many real social networks have a scale-free property unlike a random network [13]; i.e., the tail of the degree-distribution function $P(k)$ is governed by a power-law function

$$P(k) \sim k^{-\gamma}, \tag{7}$$

where γ is a constant and called a *power-law index*. It is known that many real scale-free networks have a power-law index γ of 2.0–3.4 [6].

Figure 5 shows degree distribution of the corporate transaction network (in the case of the first quarter of 2005). This figure shows the estimated degree distribution function $P(k)$ as a function of node degree k. It can be found from this figure that the degree distribution of the corporate transaction network is considerably different from the Poisson distribution (i.e., a random network). Also can be found that the tail of the degree distribution function $P(k)$ follows a power law. For instance, in the corporate transaction network, the average degree $\overline{k}$ is around 4.52–4.58. We found that the power-law index γ of the corporate transaction network was $\gamma = 2.06$, which clearly suggested existence of scale-free structure in the corporate transaction network.

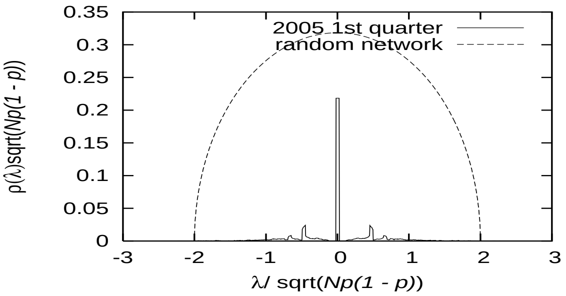

Figure 6: Spectrum density of the corporate transaction network (1st quarter of 2005) (dotted line is the case of random network)

3.4 Spectrum Density

We focus on spectrum density of the corporation transaction network. Namely, topological characteristics of the corporation transaction network are clarified.

Spectrum density is a metric used for analyzing the topology structure of a network [14]. Let $\mathbf{A}$ be the incidence matrix of a network with the number N of nodes. Namely, $\mathbf{A}$ is a square matrix of $N \times N$, and each element a_{ij} is 1 if a link exists between node i and node j, and is 0 otherwise. Eigenvalues of the incidence matrix $\mathbf{A}$ are denoted by λ_i ($1 \leq i \leq N$). The spectrum density $\rho(k)$ is defined as [6]

$$\rho(\lambda) \equiv \frac{1}{N} \sum_{j=1}^{N} \delta(\lambda - \lambda_j). \tag{8}$$

Let N be the number of nodes, and $\overline{k}$ be the average degree. It is known that the spectrum density $\rho(k)$ of a random network is given by [6]

$$\rho(\lambda) = \begin{cases} \frac{\sqrt{4Np(1-p)-\lambda^2}}{2\pi Np(1-p)} & \text{if } |\lambda| < 2\sqrt{Np(1-p)} \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

The spectrum density of the corporate transaction network for different values of λ (in the case of the first quarter of 2005) is shown in Fig. 6. . This figure shows that the spectrum density of the corporate transaction network significantly differs from that of a random network. Spectrum density of the corporate transaction network has a large peak at $\lambda = 0$ and a small peak at $\lambda = 1$. However, it takes rather small values in other regions.

3.5 Structural Superiority

In [15], an index called *structural superiority index* is defined in the inter-industry relation network, where a node corresponds to an industry and a link to relation between industries. In the inter-industry relation network, the structural superiority index shows how advantageous a certain industry is as compared with other industries from a network structural viewpoint. In [15], the inter-industry relation network is built from Japanese inter-industry relations table [16], and the relation between structural superiority index and profit of each industry is investigated. Consequently, it is shown that the industry with a high structural superiority index has tendency to obtain more profits than that with a low structural superiority index.

In this paper, we propose a structural superiority index for the corporate transaction network by extending the idea of the structural superiority index proposed in [15]. We clarify relations between the structural superiority index obtained from the corporate transaction network and several corporate indices such as stock price indices and profit financial indices. As stock price indices, we focus on price earnings ratio (PER) and earnings per share. As profit financial indices, we focus on return on equity (ROE), return on asset (ROA), and market capitalization.

In Section 2, the corporate transaction network was built as an undirected graph. In what follows, for obtaining structural superiority index, the corporate transaction network is built as a directed graph. This is because the structural superiority index is determined by whom a corporation is considered to be a major customer. For instance, if a corporation A considers a corporation B to be a major customer, and if the corporation B does not consider the corporation A to be a major customer, the corporation B is in a superior position to the corporation A.

Hence, the definition of a link in the corporate transaction network is changed as follows. When the list of major customers of a corporation A includes a corporation B (i.e., transaction exists from the corporation A to the corporation B), we have a link from the corporation A to the corporation B.

In this case, the structural superiority index C_i of a corporation i in the corporate transaction network is defined as

$$C_i \equiv \sum_{j \neq i} \zeta_{ji} \frac{C'_j}{\left(k_j^O\right)^{\alpha}}, \tag{10}$$

where C'_j is an approximate value of the structural superiority index of a corporation j, and is defined as

$$C'_j \equiv \sum_{k \neq j} \zeta_{kj} \frac{k_k^I}{(k_k^O)^{\alpha}}, \tag{11}$$

where k_i^I is the input degree of node i, k_i^O is the output degree of node i, and α is a parameter ($0 \leq \alpha \leq 1$). ζ_{ij} is 1 if a link exists from node i to node j, and 0 otherwise.

Equation (10) indicates that, when k_j^I is large (i.e., a corporation j, which considers a corporation i to be a major customer, is considered to be a major customer by many other corporations), the structural superiority index is large. Moreover, Eq. (10) also indicates that, when k_j^O is small (i.e., the number of major customers of a corporation j, which includes a corporation i is small), the structural superiority index is large.

The relations between the structural superiority index for $\alpha = 0$ and several corporation indices — price earnings ratio (PER), earnings per share, return on equity (ROE), return on asset (ROA), and market capitalization — are shown in Figs. 7 through 11, respectively. In the figures, regression lines obtained from the least-squares method are also plotted.

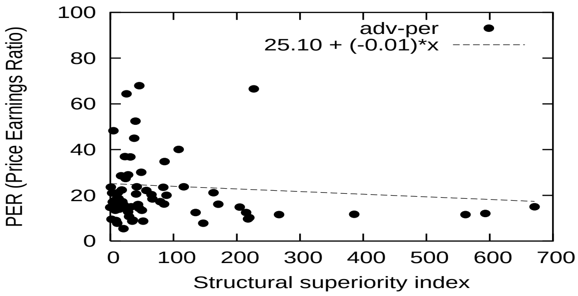

Figure 7: Relation between structural superiority index and price earnings ratio (PER) for $\alpha = 0$

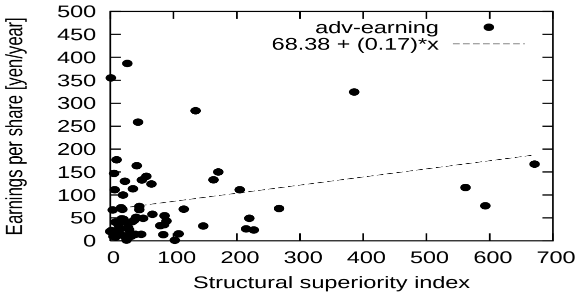

Figure 8: Relation between structural superiority index and earnings per share for $\alpha = 0$

First, we focus on the relation between the structural superiority index of the corporate transaction network and stock price indices. Figure 7 shows existence of a weak negative correlation between the structural superiority index and price earnings ratio (PER). On the contrary, Fig. 8 shows that relatively strong positive correlation exists between the structural superiority index and earnings per share. From these observations, we conclude that the structural superiority index of the corporate transaction network is strongly related with stock price indices.

Next, we focus on the relation between the structural superiority index of the corporate transaction network and profit financial indices. Figures 9 and 10 show that a weak negative correlation exists between the structural superiority index and either return on equity (ROE) or return on asset (ROA). On the contrary, Figure 11 shows that the strong correlation exists between the structural superior-

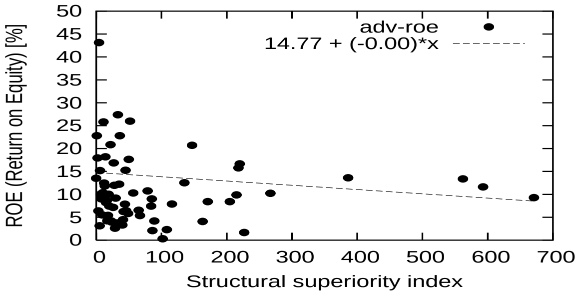

Figure 9: Relation between structural superiority index and return on equity (ROE) for $\alpha = 0$

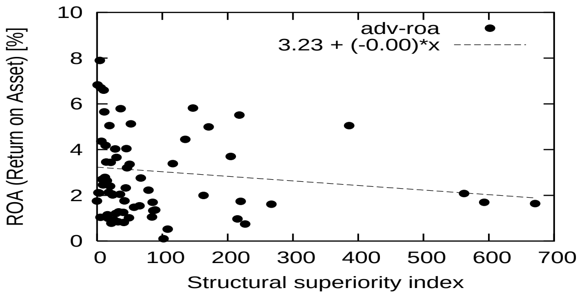

Figure 10: Relation between structural superiority index and return on asset (ROA) for $\alpha = 0$

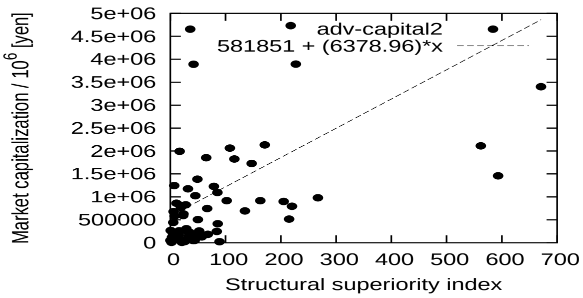

Figure 11: Relation between structural superiority index and market capitalization for $\alpha = 0$

ity index and market capitalization. From these observations, we conclude that the structural superiority index of the corporate transaction network is strongly related with profit financial indices.

4 Conclusion and Future Works

In this paper, we have built the corporate transaction network based on dealings information among real Japanese corporations. Furthermore, average distance, clustering coefficient, degree distribution, spectrum density, and structural superiority index of the corporate transaction network have been obtained and investigated by applying network-analysis techniques. Consequently, we have found the followings regarding the corporate transaction network: (1) average distance is approximately 3.4–5.4, which indicates that nodes in the corporate transaction network are more closely related than nodes in other social networks, (2) clustering coefficient is approximately 0.07, which indicates that the corporate transaction network tends to have less grouped structure than other social networks, (3) the tail of the degree distribution function follows a power law, which suggests the corporate transaction network has a scale-free property, (4) spectrum density has a small peak at $\lambda = 1$, but it takes rather small values in other regions, and (5) structural superiority index is strongly related to stock price indices and profit financial indices.

As future work, constructing a network generation model that can reproduce the structure of the corporate transaction network is of interest. Moreover, based on our findings on real corporate transaction networks, realization of a new network service that can improve the efficiency and convenience of a commercial transaction in Cybersociety is also interesting.

References

[1] G. F. Davis, M. Yoo, and W. E. Baker, "The small world of the american corporate elite, 1982-2001," *Strategic Organization*, vol. 1, pp. 301–326, Aug. 2003.

[2] E. Holger, M. Lutz-Ingo, and B. Stefan, "Scale-free topology of e-mail networks," *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics)*, vol. 66, pp. 1–4, Sept. 2002.

[3] M. Granovetter, "The strength of weak ties: a network theory revisited," *Sociological Theory*, vol. 1, pp. 201–233, Sept. 1983.

[4] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature (London)*, vol. 393, pp. 440–442, Oct. 1998.

[5] A. L. Barabasi, H. Jeong, Z. Neda, E. Ravasz, A. Schubert, and T. Viesek, "Evolution of the social network of scientific collaborations," *PHYSICA A*, vol. 311, pp. 590–614, Jan. 2002.

[6] R. Albert and A.-L. Barabasi, "Statistical mechanics of complex networks," *Reviews of Modern Physics*, vol. 74, June 2002.

[7] A.-L. Barabasi and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, pp. 509–512, Oct. 1999.

[8] M. Jin, M. Girvan, and M. E. J. Newman, "The structure of growing social networks," *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics)*, vol. 64, pp. 381–399, Oct. 2001.

[9] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, "Structure of growing networks with preferential linking," *Physical Review Letters*, vol. 85, pp. 4633–4636, Nov. 2000.

[10] M. E. J. Newman, "The structure and function of complex networks," *SIAM Review*, vol. 45, pp. 167–256, Mar. 2003.

[11] Nihon Keizai Shimbun, Inc., "CD-ROM Nikkei corporation information (first quarter of 2005)." `http://www.nikkei-bookdirect.com/cd-rom/index.php`, Feb. 2005 (in Japanese).

[12] W. S. Lovejoy and C. H. Loch, "Minimal and maximal characteristic path lengths in connected sociomatrices," *Social Networks*, vol. 25, no. 4, pp. 333–347, 2003.

[13] S. H. Storogatz, "Exploring complex networks," *Nature*, vol. 410, pp. 268–276, Mar. 2001.

[14] I. J. Farkas, I. Derenyi, A.-L. Barabasi, and T. Vicsek, "Spectra of real-world graphs: beyond the semicircle law," *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics)*, vol. 64, pp. 1–12, Aug. 2001.

[15] Y. Yasuda, *Network analysis — What decides activity*. Sinyosha, Feb. 1992. (in Japanese).

[16] National Federation of Statistical Associations, "Japanese inter-industry relations table." `http://www.stat.go.jp/data/io/`(in Japanese).