Physica A 389 (2010) 604–622

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Effects of friend-making resources/costs and remembering on acquaintance networks

Chung-Yuan Huang[a,b,*], Yu-Shiuan Tsai[c]

[a] Department of Computer Science and Information Engineering, 259 Wen Hwa 1st Road, 333 Taoyuan, Taiwan

[b] Research Center for Emerging Viral Infections, Chang Gung University, 259 Wen Hwa 1st Road, Taoyuan 333, Taiwan

[c] Department of Computer Science, National Chiao Tung University, 1001 Ta Hsueh Road, Hsinchu 300, Taiwan

ARTICLE INFO

Article history:
Received 10 April 2009
Received in revised form 17 September 2009
Available online 26 September 2009

Keywords:
Acquaintance networks
Social networks
Small-world phenomenon
Network-oriented simulations

ABSTRACT

We consider two overlooked yet important factors that affect acquaintance network evolution and formation—*friend-making resources* and *remembering*—and propose a bottom-up, network-oriented simulation model based on three rules representing human social interactions. Our proposed model reproduces many topological features of real-world acquaintance networks, including a small-world phenomenon and a sharply peaked connectivity distribution feature that mixes power-law and exponential distribution types. We believe that this is an improvement over fieldwork sampling methods that fail to capture acquaintance network node connectivity distributions. Our model may produce valuable results for sociologists working with social opinion formation and epidemiologists studying epidemic dynamics.

1. Introduction

The primary features of any social network are its small-world phenomena of low degree of separation (indicating how small the world is) and high degree of clustering (explaining why our friend's friend is often our friend). Watts and Strogatz's [1] research has been the basis for building social simulation models, many of which mix regular and random networks (see, for example, [2–4]). However, these models have a significant drawback: since they neglect local rules that affect interactions and relationships between individuals, they are good for explaining small-world phenomena but not for explaining the evolutionary mechanisms and essential topological features of social networks. Furthermore, researchers have shown a tendency to focus on the final products of network topologies and a bias toward top-down approaches that facilitate theoretical analyses [2,5,3,4,6,1,7,8]. Actual social network evolution involves bottom-up processes that entail local human interactions, which explains this study's focus on human social interaction and intercourse rules that affect acquaintance network evolution and formation.

Examples of other network models and applications that can be used as references when researching social networks are (a) Barabási and Albert's [9] (hereafter referred to as BA's) use of growth and preferential attachment mechanisms to establish a scale-free network model in which node degree distribution follows a power-law connectivity feature, and (b) Albert et al.'s [10] proposal that the Internet possesses scale-free features. In BA's scale-free model, nodes and edges are added continuously to a network over time (an example of a growth mechanism); edges are more likely to link with high than with low connectivity nodes (a preferential attachment mechanism). To realistically simulate this kind of complex network, Li and Chen [11] added the concept of *local-world connectivity*, which is found in many physical networks. Their

* Corresponding author at: Department of Computer Science and Information Engineering, 259 Wen Hwa 1st Road, 333 Taoyuan, Taiwan. Tel.: +886 3 2118800x3474; fax: +886 3 2118700.

E-mail address: gscott@mail.cgu.edu.tw (C.-Y. Huang).

0378-4371/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.physa.2009.09.038

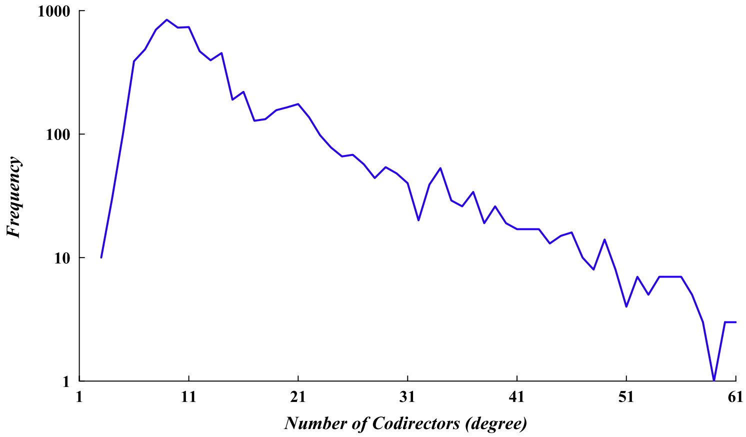

Fig. 1. Node degree distribution of acquaintance network for board directors of 1999 Fortune 1000 companies.

local-world model acknowledges transitions between power-law and exponential scaling; BA's scale-free model is one of the several special (limiting) cases.

As with other complex networks, social networks exhibit the small-world phenomenon, but they differ in terms of detail [12]. We believe there are three reasons why Internet or Web based growth models are inappropriate for describing growth in social networks: (a) in many social networks the node degree distribution does not reflect power-law connectivity features (Fig. 1); (b) social networks usually have high degrees of clustering, but Internet or Web based growth models reflect low degrees of clustering; and (c) the presence of preferential attachment mechanisms is not a crucial feature of social networks. Accordingly, Jin et al.'s [13] model of social network evolution entails a sharply peaked distribution supported by observations for many real-world social networks. Real-world social network features they have identified in their simulation models include the formation of closely-knit communities and high degrees of network transitivity.

Acquaintance networks are social networks that represent the processes and results of people meeting people. Befriending someone in an acquaintance network is an important daily activity, one that more often than not involves a friend or acquaintance who serves as an intermediary. However, a significant number of acquaintances occur by chance, with two or more individuals coming together at a specific time in a specific place because of a shared interest or activity. Davidsen et al. [14] have proposed a model based on two local and interactive rules for building acquaintance networks—one for introductions and meeting by chance, and one for the effects of aging on acquaintance networks. Their model is considered more realistic because it acknowledges the simple observation that everyone's friends regularly introduce them to additional friends. However, their model is still inadequate for explaining changes that occur once acquaintances evolve into friendships (i.e., strengthening, weakening, or separation prior to death), and are therefore of limited use when modeling actual acquaintance networks.

To better represent real-world social interactions and human intercourse, we have added an essential friendship updating rule to our model, one that considers two overlooked yet important factors: *friend-making costs* and *remembering*. Friend-making costs are both recurring and substantial in terms of time and effort spent on maintaining a friendship; given limited friend-making resources, an individual can only maintain a certain number of friendships. Remembering is a memory factor through which individuals maintain or neglect their friendships. To establish a bottom-up, network-oriented simulation model, we have taken into account friend-making resources, remembering, shared friendships (strengthening, weakening, and dissolution), meetings by chance, and joining and leaving. The model supports analyses of how social interaction and intercourse rules and their related parameters affect the evolution of topological features required to obtain an acquaintance network structure, as well as how correlations between parameters and topological features of acquaintance networks are measured under different circumstances.

Based on a series of experiments and sensitivity analyses, our four main findings are (a) the proposed model reproduces many of the topological features of real-world acquaintance networks (e.g., American corporate director data), including a small-world phenomenon and a sharply peaked connectivity distribution feature that mixes power-law and exponential distribution types; (b) acquaintance network topological features are affected by initial average parameter values but not by their statistical distributions; (c) resources, remembering, and initial friendships positively influence increases in average numbers of friends and decreases in degrees of clustering and separation; and (d) widely used fieldwork sampling methods do not capture the node degree distributions of acquaintance networks. We believe our model will produce valuable results for sociologists, computer scientists, epidemiologists, public health professionals, and others who regularly work with social opinion formation and epidemic transmission dynamics involving acquaintance networks.

2. Complex networks and topological features

Milgram [15] was the first to design experiments involving small-world phenomena; since then, many attempts have been made to depict those phenomena in human interactions. To overcome practical barriers, Watts and Strogatz [1] restated the issue as "general conditions under which the world can be small" (see also [2,6,7]), and provided empirical network data to show that low degree of separation (measured using average shortest path length) and high degree of clustering (measured as a clustering coefficient) are common to all social networks. For any given network, average shortest path length $L_{Network}$ is defined as

$$L_{Network} = \langle d_{u,v} \rangle = \frac{1}{\frac{1}{2}N(N-1)} \sum_{u \neq v} d_{u,v} = \frac{2}{N(N-1)} \sum_{u \neq v} d_{u,v}, \tag{1}$$

where N is the number of nodes (i.e., network size) and $d_{u,v}$ the shortest path between u and v. If node v has k_v neighbors and a total number of E_v edges between k_v nodes, the clustering coefficient $C_{Network}$, average node degree $\langle k \rangle$, and average node square degree $\langle k^2 \rangle$ are respectively defined as Eqs. (2)–(4).

$$C_{Network} = \langle C_v \rangle = \left\langle \frac{2E_v}{k_v(k_v-1)} \right\rangle = \frac{1}{N} \sum_{v}^{N} \frac{2E_v}{k_v(k_v-1)} \tag{2}$$

$$\langle k \rangle = \frac{1}{N} \sum_{v}^{N} k_v \tag{3}$$

$$\langle k^2 \rangle = \frac{1}{N} \sum_{v}^{N} (k_v)^2 . \tag{4}$$

To make sense of the network separation dimension, average shortest path length in a network must be compared to that in a random network that is created using the same numbers of nodes and edges per node. To grasp the magnitude of network clustering, a network's clustering coefficient must be compared to that of a regular network, once again with the same numbers of nodes and edges per node.

Lattice graphs (also called *d-lattices* and *nearest-neighbor coupled networks*) represent thoroughly researched regular networks in which individual nodes can only link with immediately adjacent neighbors. Two-neighbor ($K = 2$) periodic 1-lattices have ring shapes, while four-neighbor ($K = 4$) 2-lattices appear as 2D grid graphs. Average shortest path length for a $K > 1$ periodic 1-lattice is calculated as

$$L_{NC} = \frac{N(N+K-2)}{2K(N-1)} \to \infty \quad (N \to \infty). \tag{5}$$

In this equation the average shortest path length is too large to be considered representative of a low degree of separation. For *d-lattices* with larger numbers of neighbors, clustering coefficients are calculated as

$$C_{NC} = \frac{3(K-2)}{4(K-1)} \approx \frac{3}{4}. \tag{6}$$

Random networks are generally constructed using nodes with random connection edges. Another method is to determine the p probability of an edge between node pairs; the result is an [16] (hereafter referred to as ER) random network consisting of N nodes and $p \times [N \times (N-1)/2]$ edges. In their analysis of random networks, ER estimated the average node degree as

$$\langle k \rangle_{ER} = \frac{pN(N-1)}{N} = p(N-1) \approx pN. \tag{7}$$

Average shortest path length for an ER random network is estimated as

$$L_{ER} \propto \frac{\ln N}{\ln \langle k \rangle}. \tag{8}$$

As seen in Eq. (8), the average shortest path length of an ER random network (L_{ER}) increases logarithmically with network size N. This agrees with the low degree of separation observed in social networks—in other words, even large-scale ($N \gg 1$) networks can have very short path lengths.

Clustering coefficients for ER random networks are estimated as

$$C_{ER} = \frac{\langle k \rangle}{N}. \tag{9}$$

In cases where the average node degree $\langle k \rangle$ is fixed and the number of nodes N is sufficiently large, additional edges result in a Poisson distribution [17] of an ER random network according to the equation

$$P_{ER}(k) = \binom{N}{k} p^k (1-p)^{N-k} \approx \frac{\langle k \rangle^k e^{-\langle k \rangle}}{k!}. \tag{10}$$

Even though they exhibit high degrees of clustering, regular networks do not exhibit low degrees of separation; random networks have low degrees of separation but do not exhibit high degrees of clustering. Watts and Strogatz [1] (hereafter referred to as WS) introduced a small-world network model with a construction algorithm that starts from a 1-lattice ring in which each node has K neighbors and each side $K/2$ edges. The algorithm determines whether rewiring with a probability β ($0 \leq \beta \leq 1$) maintains a maximum of a single edge between two arbitrary nodes without any self-links. Newman and Watts [2] simplified the WS small-world model by replacing random edge rewiring with random edge additions. The NW and WS small-world models are equivalent in cases where the probability β of edge rewiring or additions is small and the number of nodes N is sufficiently large. The two models retain the high clustering feature of regular networks. Adding a small amount of randomness dramatically reduces average shortest path length.

While the two models capture the small-world phenomena of complex networks, they are insufficient for social networks because of the additional characteristics involved. For instance, clustering in social networks is relatively unmixed, meaning that many social networks are divided into distinct groups [12]. Leenders [18] has observed that friendships can be categorized as acquaintances, just friends, good friends, best friends, true friends, confidants, and so forth. Furthermore, social networks have *skewed node degree distributions* on top of small-world phenomena [10,19]. For instance, node degree distribution in a collaboration network of actors obeys power-law features for part of its range, but has what appears to be an exponential cutoff for very high node degrees. This kind of skewed distribution is also observed in coauthoring networks such as those found for collaborations between biologists and physicists [20].

However, lower skewness is noted in the Fig. 1 node degree distribution for an acquaintance network consisting of directors from the boards of the top 1000 companies identified by Fortune magazine in 1999. In the Figure, the X-axis is linear horizontal and the Y-axis logarithmic. Any edge between two director nodes indicates that they sit on the same board. In addition to a sharp peak, the figure shows exponential decay with some random rippling and a small peak in the tail; accordingly, its right-skewness does not noticeably contrast with those representing actor collaboration networks. Newman et al. [20] argue that the maintenance of ties in director acquaintance networks carries a substantial and recurring cost in terms of continuous time and effort. In contrast, collaboration among actors carries one-time costs (i.e., the time and effort put into a film or stage production), but the resulting ties can endure for an indefinite period. The difference has the potential of placing a stronger limitation on the number of friendships an individual can hold instead of on the number of collaborators.

3. Friendship evolution and the three-rule model

Davidsen et al. [14] have proposed a two-rule model of acquaintance network evolution, with the first rule addressing how people make new friends via introductions or meetings by chance, and the second addressing how friendships are broken when one party dies. Here we will introduce a third rule that considers two factors that we believe have been overlooked: *resources* and *remembering*. Both allow for the strengthening and weakening of friendships, and are therefore helpful for sociologists, computer scientists, epidemiologists, public health professionals, and other researchers who regularly work with social opinion formation and epidemic transmission dynamics involving acquaintance networks.

The three rules are stated as:

1. *Friend making* (Fig. 2). A randomly chosen individual introduces two friends to each other. If this is their first meeting, a new edge is formed between them, but if the individual only has one friend, he makes a self-introduction to one other randomly chosen individual. Note that we use the verb "introduce" to describe both chance meetings and introductions by common friends.
2. *Joining and leaving* (Fig. 3). At p probability, one randomly chosen individual and all associated edges are removed from an acquaintance network and replaced by a new individual who is then introduced to another randomly chosen individual. According to this rule, groups of acquaintances are viewed as "friend circles" whose members join voluntarily and leave for reasons other than death.
3. *Friendship update* (Fig. 4). According to this rule, $b \times M$ friendships are selected for updating. Probability b is a proportion used for determining how many friendships are updated, and M the total number of friendships (or edges) that exist at any given time step. Further details on friendship updating will be discussed in Section 3.1.

A simulation flowchart of our proposed model[1] is shown in Fig. 5, and a list of parameters is presented in Table 1. Simulations based on our three-rule model begin with parameter initialization. During the network construction stage, an acquaintance network is formulated with a fixed number of N nodes and undirected edges between pairs of randomly chosen nodes. Nodes represent individuals and undirected edges represent connections between individuals who know each other. During a simulation, N remains constant—that is, N in our proposed model lacks variation in the number of individuals belonging to an acquaintance network, whereas BA's scale-free model uses a continuous growth mechanism. During each time step, the three rules are executed in turn until the simulated acquaintance network reaches a statistically stationary state.

[1] We used Python with NetworkX and matplotlib packages to encode our three-rule simulation model and to collect topological data during a simulation run. For a detailed description of our simulation model and related packages, please contact the corresponding author.

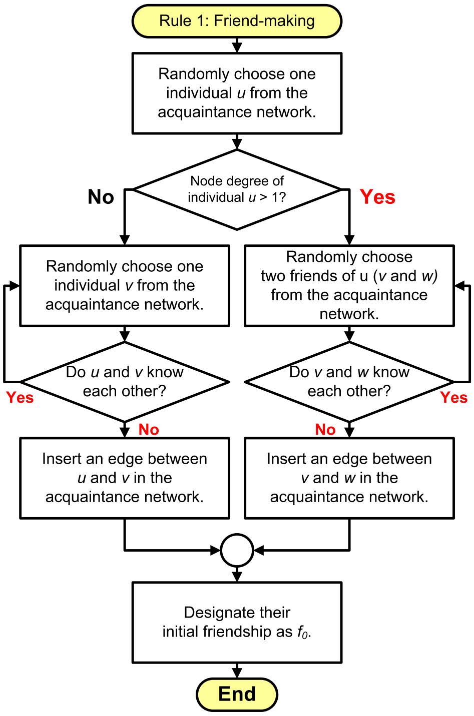

Fig. 2. Friend-making rule flowchart.

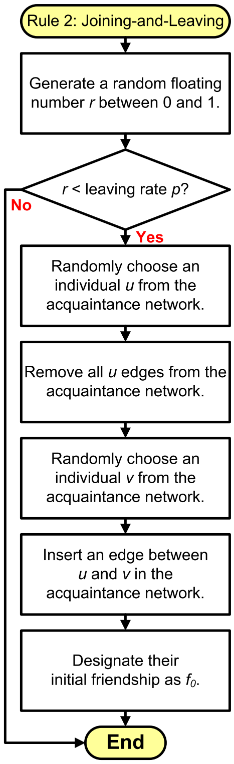

Fig. 3. Joining-and-leaving rule flowchart.

Statistically stationary states are identified when two factors both approach 0 for sufficiently long time periods: (a) average rate of change, and (b) the corresponding standard deviation for each of the four topological properties average

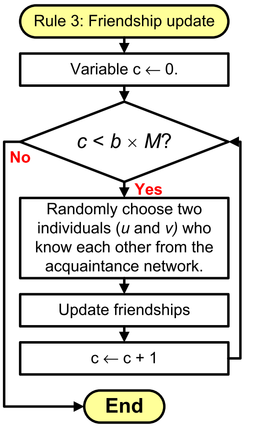

Fig. 4. Friendship-update rule flowchart.

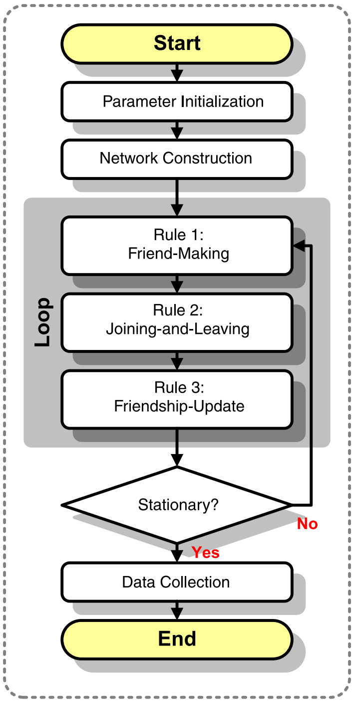

Fig. 5. Three-rule model flowchart.

node degree $\langle k\rangle$ (Eq. (3)), average node square degree $\langle k^2\rangle$ (Eq. (4)), clustering coefficient C (Eq. (2)), and average shortest path length L (Eq. (1)) (Table 2). Furthermore, each of the four properties eventually converges to a stable value that exhibits slight rippling (Fig. 6). Once an acquaintance network reaches a statistically stationary state, $\langle k\rangle$, $\langle k^2\rangle$, C, L, and node degree distribution properties can be calculated to characterize the resulting acquaintance network. Since the node degree distributions in our simulations involve some random rippling (especially for smaller populations), we applied Bruce's [22] ensemble average:

$$\bar{P}(k) = \frac{1}{M}\sum_{i=1}^{M} P_i(k), \tag{11}$$

where M is the number of curves to be averaged, k the node degree, $P(k)$ the proportion of nodes in the acquaintance network having node degree k, and $P_i(k)$ the proportion of network nodes with node degree k in the ith curve to be averaged.

Fig. 6 presents an example of a statistically stationary state for an acquaintance network initialized at $N = 1000$, $p = 0$, $b = 0.001$, $q = 0.9$, $\theta = 0.1$, r with a fixed value of 0.5, and beta14 $f_0(\mu = 0.9)$. Solid lines indicate the acquaintance network. The dashed lines in Fig. 6(c), (d) (average shortest path length L_{ER} and clustering coefficient C_{ER} calculated using

Table 1
Parameters and distribution functions of our proposed three-rule model for simulating acquaintance network evolution.

Parameter	Description	Default value of sensitivity analysis experiments			
		Text section 4.1.1	Text section 4.1.2	Text section 4.1.3	Text section 4.3
N	Initial number of nodes in acquaintance network.	1000	1000	1000	1000
p	Rule 2 joining-and-leaving probability.	0 and 0.0025	0	0	0
b	Proportion of updated Rule 3 friendships.	0.001	0.001	0.001	0.001
f	Friendships between two individual.	×	×	×	×
f_0	Initial friendship distribution.	0.5	1	×	0.5
q	Friend remembering.	0.6	×	×	0.4
r	Friend-making resource.	0.5	×	0.5	0.5
θ	Breakup threshold.	×	0.1	0.1	0.1
μ	Mean.				
α	Shape.				

Function	Description
fixed_value (μ)	A distribution with a random variable as a fixed value μ.
beta14 (μ)	A beta distribution instance in [0,1] that satisfies $\alpha+\beta=14$ and has an average of $\mu=\alpha/(\alpha+\beta)$. We chose a beta distribution because critical parameters such as initial friendships, friends remembering, resources, and breakup thresholds have ranges of 0–1. Beta distributions belong to a two-parameter family of continuous probability distributions defined according to the interval [0, 1] with a probability density function of $f(x;\alpha,\beta)=[1/B(\alpha,\beta)]x^{\alpha-1}(1-x)^{\beta-1}$, where B is the beta function and where α and β must be greater than zero [21].
pareto (α)	Return a random floating point in [0, 1] with a Pareto distribution. α specifies the shape parameter.

Table 2
Topological features for characterizing acquaintance networks.

Parameter	Description	Complex networks	ER's (1960) random networks	1D regular networks
$\langle k\rangle$	Average node degree.	Eq. (3)	Eq. (7)	Eq. (3)
$\langle k^2\rangle$	Average node degree squared.	Eq. (4)	Eq. (4)	Eq. (4)
C	Average clustering coefficient.	Eq. (2)	Eq. (9)	Eq. (6)
L	Average shortest path length.	Eq. (1)	Eq. (8)	Eq. (5)

Eqs. (8) and (9), respectively) are for an ER random network at the same numbers of nodes and edges per node as the acquaintance network.

3.1. Friendship update

A selected friendship connecting individuals u and v is updated using Eqs. (12) and (13).

$$f_{u,v}^{new}=q\cdot f_{u,v}^{old}+(1-q)\cdot J\left(D\left(\frac{r_u}{k_u}\right),D\left(\frac{r_v}{k_v}\right)\right) \tag{12}$$

$$\text{if } (f_{u,v}^{new}<\theta) \text{ then remove } edge(u,v) \text{ from acquaintance network} \tag{13}$$

where $f_{u,v}^{new}$ represents the new friendship between u and v; $f_{u,v}^{old}$ the original friendship; q the friend remembering; θ the breakup threshold; r_u individual u's friend-making resources; k_u individual u's number of friends; r_u/k_u individual u's average friend-making resources spent on each friend; and r_v, k_v, and r_v/k_v individual v's friend-making resources, friend numbers, and average resources spent on each friend, respectively. J is a joint function and D a distribution function. For convenience, the q, r, and θ parameters are normalized between 0 and 1.

Eq. (12) is divided into two parts by the friend-remembering q. The first part $f_{u,v}^{old}$ represents the effect of old friendships and the second part $J(D(r_u/k_u),D(r_v/k_v))$ the effect of limited resources. Newly updated friendships may grow weaker or stronger. The friendship between u and v may cease to exist and the edge between them removed from the acquaintance network if $f_{u,v}^{new}$ is lower than the θ breakup threshold (Eq. (13)). To prevent the loss of generality of the friendship-update rule and to maintain accurate simulation results, we used $D(x)=\text{fixed\_value}(x)$ (Table 1) as the distribution function and $J(a,b)=\text{fixed\_value}((a+b)/2)$ as the joint function. We implemented the updated friendship equation as Eq. (14) for use in subsequent simulation experiments.

$$f_{u,v}^{new}=q\cdot f_{u,v}^{old}+(1-q)\cdot\left(\frac{r_u}{k_u}+\frac{r_v}{k_v}\right)\cdot\frac{1}{2}. \tag{14}$$

3.2. Model validation

We replicated Davidsen et al.'s [14] simulations using their original $N=7000$ and p values of 0.04, 0.01, and 0.0025. According to our proposed three-rule model, a joining-and-leaving probability of 0 means that Rule 2 is inactive, and a θ

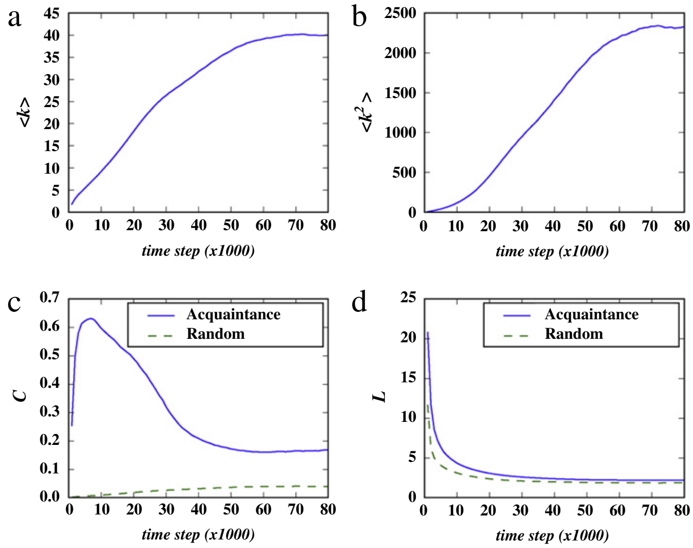

Fig. 6. Examples of a statistically stationary state using our proposed three-rule model.

breakup threshold equal to 0 means that Rule 3 is inactive. An inactive Rule 3 makes our three-rule model the equivalent of Davidsen et al.'s two-rule model. In addition to the $P(k)$ node degree distribution, we presented $\langle k\rangle$, C, and L values (all at various p levels) and analyzed their correlations to observe the effects of p on the acquaintance network. The $P(k)$ node degree distribution is presented in Fig. 7. All $\langle k\rangle$, C, and L values with parameter initializations for various probability p values are shown in Table 3 and Fig. 8. Correlations among $\langle k\rangle$, C, and L are shown in Fig. 9. In Fig. 8, solid lines reflect the application of Davidsen et al.'s two-rule model; the dashed lines (clustering coefficient C_{ER} in Fig. 8(b) and average shortest path length L_{ER} in Fig. 8(c) calculated respectively using Eqs. (8) and (9)) reflect the application of the ER random network using the same numbers of nodes and edges per node. Differences between the two lines indicate small-world phenomena in the acquaintance network.

According to Fig. 8(a), an individual's average number of friends sharply increases at a small joining-and-leaving probability p. As shown in Fig. 9(d), the clustering coefficient closely follows average node degree but not average shortest path length. A larger p indicates a higher chance of leaving and a lower p a longer activity period. To reflect the amount of time required to make friends, Davidsen et al. narrowed their focus to a $p \ll 0.1$ regime. We also explored the $p \gg 0.1$ regime in an effort to satisfy model validation needs, and found that average node degree $\langle k\rangle$ decreased for p values between 0 and 0.5. This decrease slowed once $p > 0.1$ (Fig. 8(a)).

As shown in Fig. 7, the smaller the joining-and-leaving rate p, the better the fit between the acquaintance network node degree and a power-law distribution. However, a node degree distribution in a social network usually has exponential cutoff features for very high degree values, since friend-making resources such as activity time, age, and energy are limited—it is impossible for individuals to continuously make new friends. This explains the low number of "rich" individuals (i.e., nodes with high degrees, many friends, or large amounts of social resources) in social networks.

In addition, Davidsen et al. believe that a small joining-and-leaving probability p (e.g., $p = 0.0025$ for the short-dash curve in Fig. 7) is a more reasonable explanation for why individuals join existing acquaintance networks (e.g., schools, corporations). We agree with their observation that a small p joining-and-leaving probability in Rule 2 is the best match for actual social activity scenarios. However, we also acknowledge that the node degree in the short-dashed curve in Fig. 7 has none of the cutoff features commonly found in social network curves, even though it has a sharp peak on the right side. Moreover, as seen in Fig. 8(a) and Table 3, as the p joining-and-leaving probability decreases from 0.04 to 0.0025, $\langle k\rangle$ increases from 12.64 to 115.34 and $\langle k^2\rangle$ increases from 467.23 to 29652.69—in other words, at $p = 0.0025$, the acquaintance network has a large number of "rich" nodes. Fig. 10(a), (b) show node degree distributions and network structures for $N = 100$ and $p = 0.0025$ without using Rule 3 (i.e., consistent with Davidsen et al.'s two-rule model). Fig. 10(c), (d) show node degree distributions and network structures for $N = 100$, $p = 0.0025$, $b = 0.01$, $r = 1.0$, $q = 0.5$, $\theta = 0.1$, and $f_0 = 1.0$. As seen in Fig. 10(a), an exceptionally high peak on the right side of the node degree distribution indicates that high connectivity individuals in the acquaintance network represent the majority, which is inconsistent with the cutoff features associated with the age limitation (Fig. 10(c), right tail of the node degree distribution).

In Fig. 11 we present a comparison of node degree distributions for one of our acquaintance networks at a statistically stationary state (solid curve) versus American corporate director data (dashed curve). Network parameters were initialized at $N = 7680$, $p = 0$, $b = 0.001$, $q = 0.4$, $\theta = 0.1$, $r = 0.5$ (fixed) and $f_0 = 0.5$ (fixed). American corporate director data are for connections among nearly 7680 directors on the boards of Fortune 1000 companies in 1999 [23]. The corresponding node degree distribution has a strong peak around node degree 10 and rapid exponential tail decay—much faster than for a

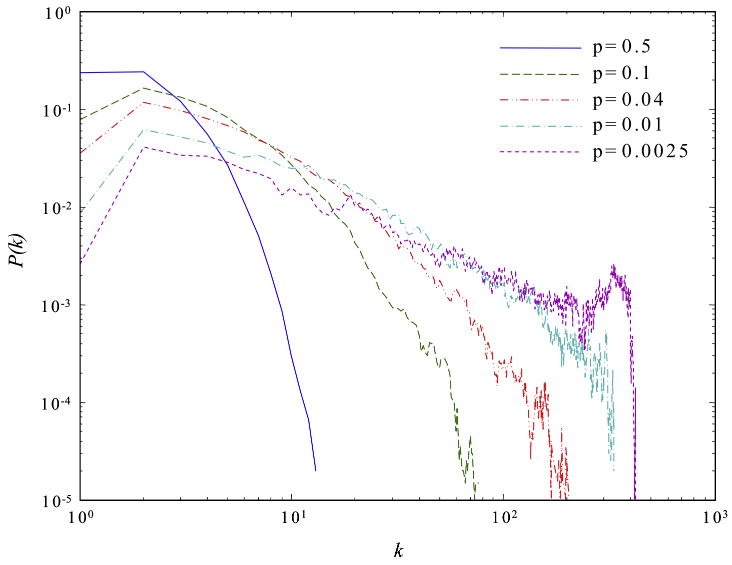

Fig. 7. Node degree distributions for a two-rule model.

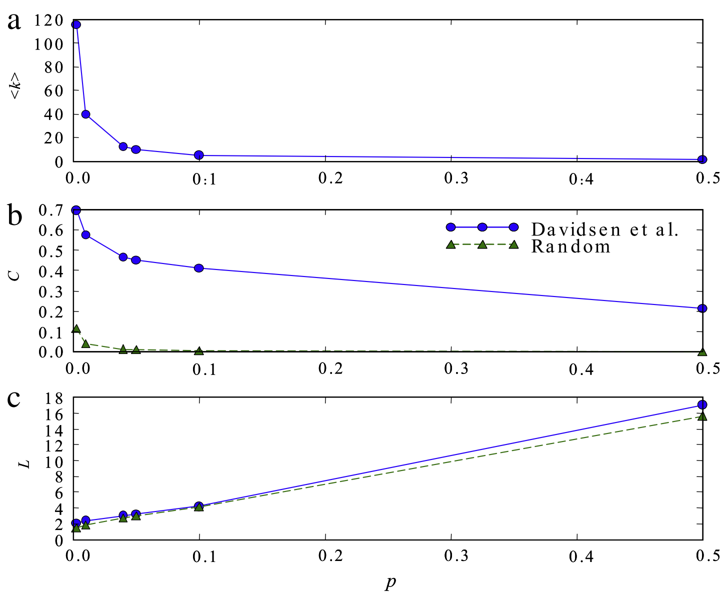

Fig. 8. Value of topological features according to various joining-and-leaving probability values.

Table 3
Topological feature values according to various joining-and-leaving probabilities.

p	0.5	0.1	0.05	0.04	0.01	0.0025
$\langle k\rangle$	1.55	5.246	10.02	12.64	39.71	115.34
$\langle k^2\rangle$	4.88	60.13	314.18	467.23	4708.25	29652.69
C	0.213	0.413	0.453	0.465	0.577	0.697
L	17.076	4.226	3.262	3.053	2.416	2.111

power-law distribution, but slower than for a Poisson or normal distribution (see also [20,6]). Interlocking networks of boards and directors reveal a remarkable node degree distribution that differs significantly from either BA's scale-free model or ER's random networks [24]. The nearly overlapping node degree distribution curves exhibit similar peaks and tails that do not decay smoothly (i.e., no statistically significant differences; correlation coefficient $c \approx 0.983822$).

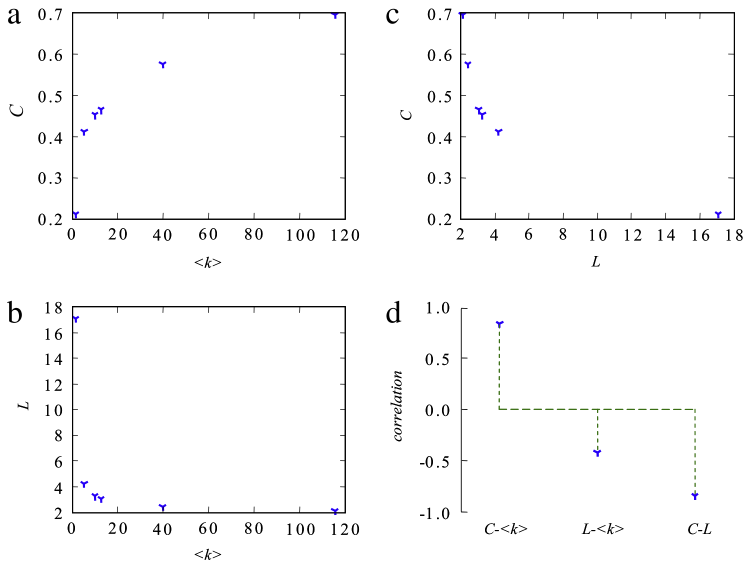

Fig. 9. Correlations among topological features according to various joining-and-leaving probability values.

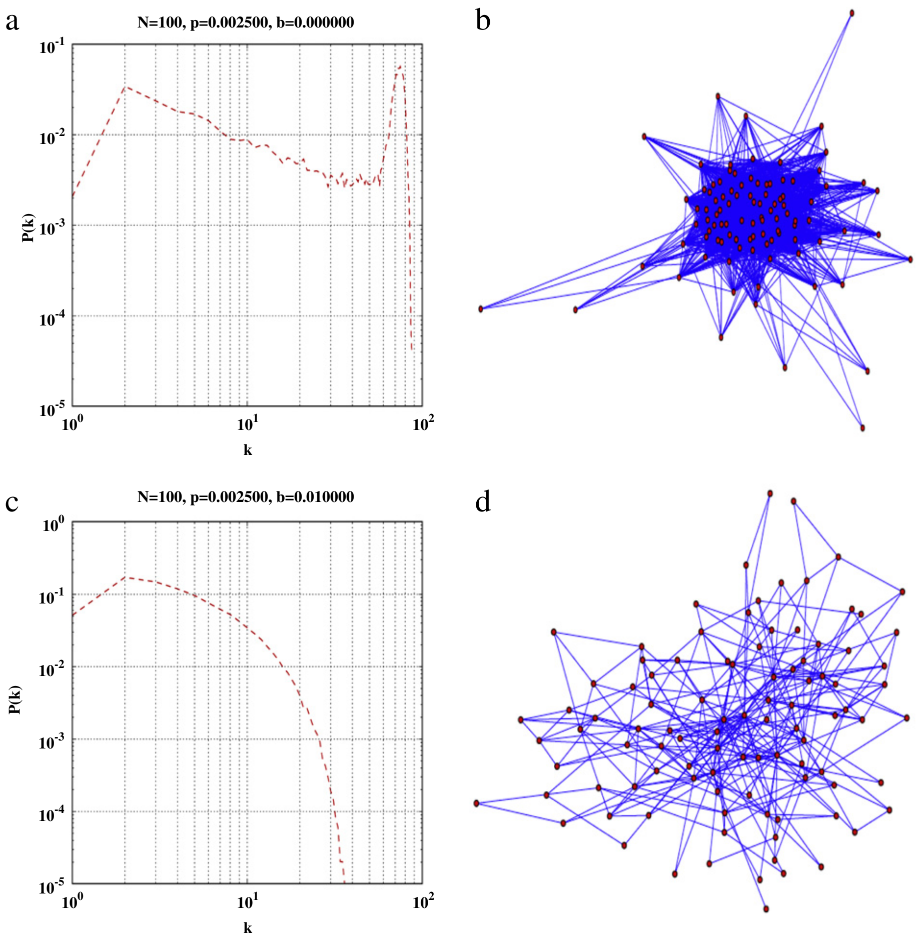

Fig. 10. Comparison of node degree distributions and network structures between Davidsen et al.'s two-rule model and our proposed three-rule model.

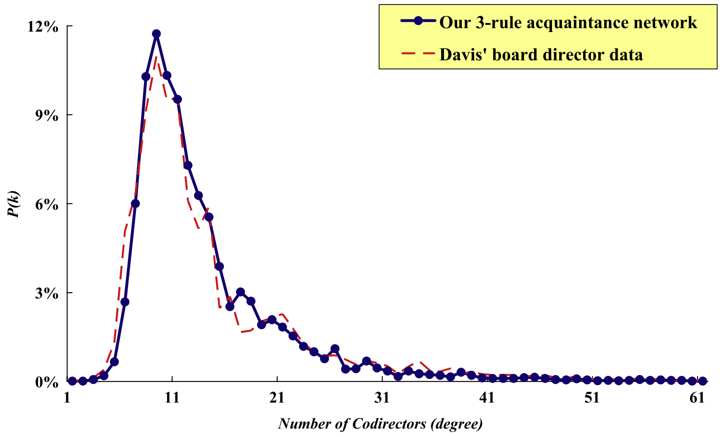

Fig. 11. Comparison of node degree distributions between one instance of our acquaintance network (solid line) and Davis's [23] board director data (dashed line).

4. Results

The three rules in our proposed model influence two aspects of acquaintance networks: the friend-making rule adds edges (thereby increasing the average number of friends), and both the joining-and-leaving and friendship-update rules are capable of removing edges, thereby reducing the average number of friends. In addition, increases in $\langle k\rangle$ lead to decreases in L and C.

As opposed to the large number of factors associated with the friendship-update rule, the joining-and-leaving rule has a single parameter p. The q factor denotes an individual's ability to remember friends, thus increasing the average number of friends. The resource factor r determines an individual's resources for making friends, thereby setting an upper limit for average number of friends. The breakup threshold θ determines the difficulty of cutting off a friendship—a negative influence on average number of friends. The initial friendship factor f_0 reflects the amount of attention an individual gives when making a new acquaintance—a positive contribution to friend making. We assume that parameters q, r, and f_0 exert positive influences on $\langle k\rangle$, and parameters p and θ exert negative influences on $\langle k\rangle$.

4.1. Sensitivity analyses: Friendship-update rule parameters

To determine the effects of the breakup threshold on acquaintance networks, we ran a series of simulations with parameters initialized at different levels of the breakup threshold θ. In all of the sensitivity analysis experiments described in this section, N was initialized at 1000 and b at 0.001. Other initialized parameters are listed in Table 1. The solid lines in Fig. 12 represent data for average node degree $\langle k\rangle$, clustering coefficient C, and average shortest path length L without Rule 2 ($p = 0$); dashed lines represent the same data with Rule 2 ($p = 0.0025$) enforced. The data indicate that Rule 2 (acting as an aging factor) slightly reduced $\langle k\rangle$ and C while slightly increasing average shortest path length L.

According to the data presented in Table 4 and Fig. 12, a larger breakup threshold θ results in significant decreases in $\langle k\rangle$ and increases in both C and L. As seen in Fig. 12, our results for the acquaintance network in this experiment are inconsistent with those for a small-world network. When θ is small (e.g., $\theta = 0.0125$ in Fig. 12), even though the average shortest path length becomes extremely small, the network's clustering coefficient also becomes extremely small due to the significantly higher average node degree. On the other hand, when θ is large (e.g., $\theta = 0.2$ in Table 4), even though the average shortest path length becomes extremely large, the network's clustering coefficient also becomes extremely large due to the significantly lower average node degree. The Fig. 13 data indicate negative $C - \langle k\rangle$ and $L - \langle k\rangle$ correlations and a positive C–L correlation, and the threshold θ reflects the ease with which friendships are broken. As expected, a higher θ results in a smaller number of "average friends" and greater separation between individuals.

To determine the effects of friend-making resources and remembering on acquaintance networks, we ran a series of simulations using parameters initialized with different r and q values. According to our simulation results, a larger r resulted in an increased $\langle k\rangle$ and decreased C and L (Table 5 and Fig. 14). The influences of different resource distributions on topological features were unclear, but the opposite was true for resource averages. The Fig. 14 data also indicate that increases in q boosted $\langle k\rangle$ and reduced both C and L. Furthermore, $C - \langle k\rangle$ and $L - \langle k\rangle$ correlations were identified as negative and the C–L correlation as positive (Fig. 15).

As seen in Fig. 14, networks consisting of cherished friends easily add more friends, and their friendships are less likely to deteriorate before ending. In contrast, networks consisting of neglected friends easily lose friends, and their friendships are more likely to deteriorate. Results from the sensitivity analysis experiment for q remembering values are similar to those for breakup threshold θ—that is, the q values in acquaintance networks are inconsistent with those in small-world networks.

Table 4
Topological feature values according to various breakup thresholds and joining-and-leaving probabilities.

θ			0.2	0.1	0.05	0.025	0.0125
p	0.	$\langle k \rangle$	5.99	9.94	14.39	22.05	38.21
		$\langle k^2 \rangle$	58.84	153.62	256.06	533.08	1528.89
		C	0.3778	0.3146	0.1899	0.1284	0.1085
		L	4.0986	3.2369	2.8368	2.5431	2.1414
	0.0025	$\langle k \rangle$	6.00	9.61	14.30	21.89	34.84
		$\langle k^2 \rangle$	60.48	146.94	257.08	531.31	1365.39
		C	0.3876	0.3112	0.1532	0.1108	0.1484
		L	4.0543	3.3064	2.8975	2.6230	2.3075

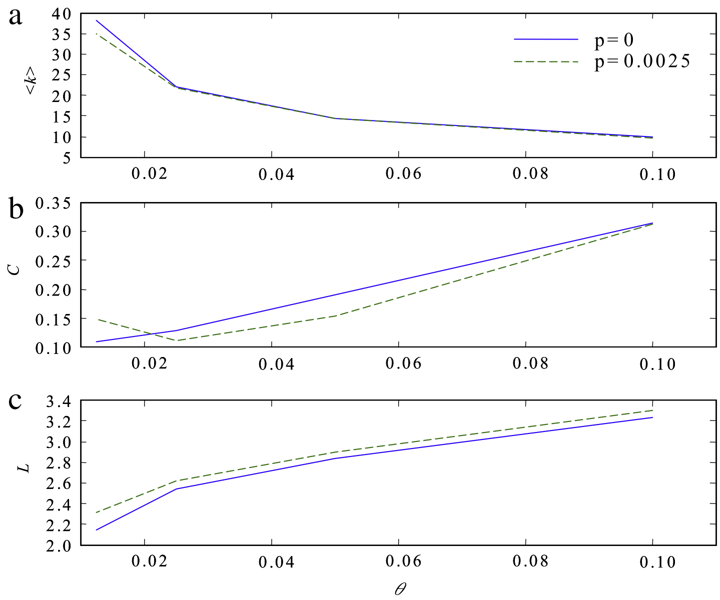

Fig. 12. Values of topological features according to various breakup thresholds and joining-and-leaving probabilities.

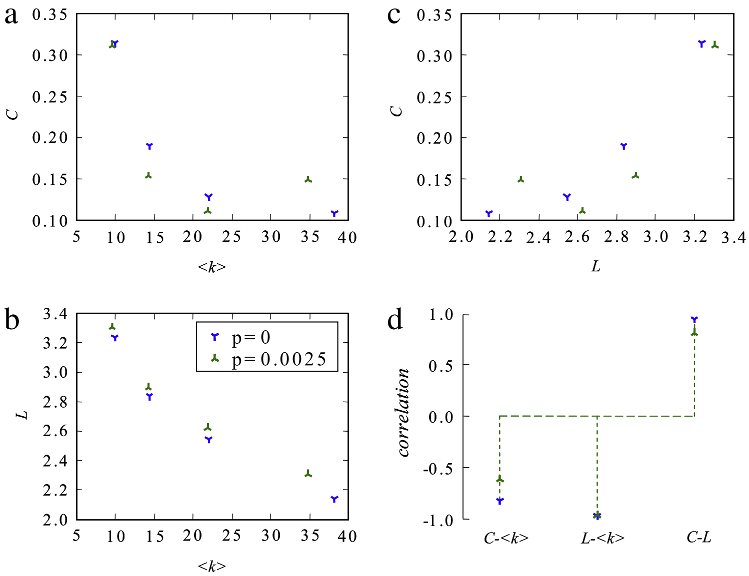

Fig. 13. Correlations among topological features according to various breakup threshold values.

When the statistical distribution of friend-making resources follows a long-tailed power-law pattern (also referred to as a heterogeneous property for individuals in an acquaintance network), the majority of those individuals have very limited

Table 5
Topological feature values according to various friend-remembering levels and friend-making resource distributions.

q			0.0	0.2	0.4	0.6	0.8	0.9
r	beta14 ($\mu = 0.1$)	$\langle k \rangle$	4.00	4.05	6.00	8.17	17.96	39.95
		$\langle k^2 \rangle$	24.62	25.60	68.48	135.10	616.24	2338.91
		C	0.3588	0.3816	0.3582	0.3590	0.2969	0.1780
		L	5.5254	5.4701	3.9116	3.4143	2.7518	2.2644
	beta14 ($\mu = 0.5$)	$\langle k \rangle$	4.96	6.03	7.99	11.43	21.20	41.25
		$\langle k^2 \rangle$	30.59	49.21	105.52	237.55	753.64	2438.32
		C	0.4554	0.3814	0.3642	0.3245	0.2382	0.1802
		L	6.0173	4.4481	3.4747	3.0483	2.5915	2.2003
	beta14 ($\mu = 0.9$)	$\langle k \rangle$	8.18	9.73	11.60	14.00	23.97	43.99
		$\langle k^2 \rangle$	71.30	109.03	161.57	252.06	821.56	2625.60
		C	0.1905	0.2246	0.1737	0.1778	0.1705	0.1461
		L	3.8626	3.3340	3.1260	2.8912	2.4840	2.1564
	fixed_value ($\mu = 0.5$)	$\langle k \rangle$	4.68	6.02	8.01	11.37	19.97	40.30
		$\langle k^2 \rangle$	26.03	46.23	95.12	226.93	659.40	2356.56
		C	0.486	0.386	0.348	0.319	0.275	0.189
		L	6.954	4.611	3.575	3.067	2.624	2.201

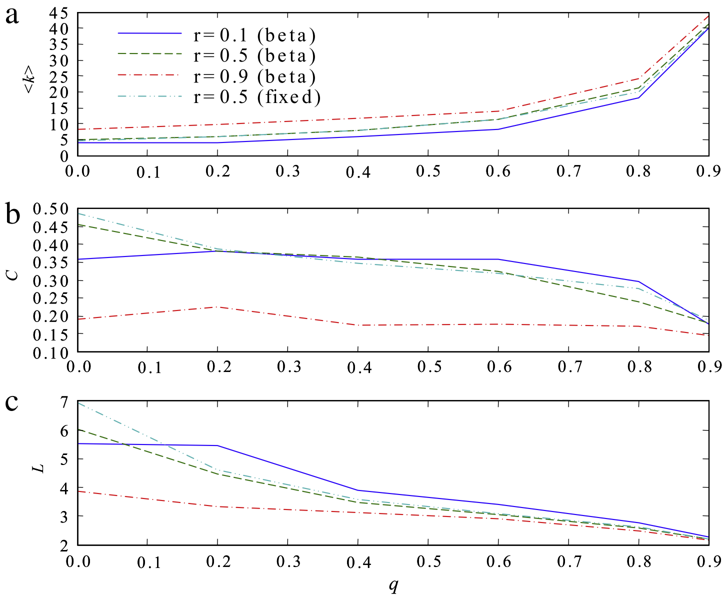

Fig. 14. Values of topological features according to various friend-remembering levels and friend-making resource distributions.

friend-making resources and a small number have large amounts (Table 6 and Figs. 16 and 17). In this scenario, the power-law distribution pattern exerts only a slight influence on the topological features of the simulated acquaintance network. Regardless of whether the friend-making resources follow a delta (fixed value), beta, or power-law distribution pattern, simulated acquaintance networks that reach a statistically stationary state have approximately the same topological features (i.e., no statistically significant differences) as long as the average individual friend-making resource values for the three distributions are equal.

According to the simulation results, joint function J in Eq. (14) is a dominant factor in that it requires both time and effort for friendship maintenance. If one or two sides lack friend-making resources, their friendship will not last, even if they both value it (i.e., the q remembering value is very high). In a simulated acquaintance network in which individual friend-making resources are heterogeneous, if one side has a small amount of friend-making resources while the other has a large amount, according to joint function J the average friend-making resource values from both sides should be close (or approximately equal) to the average value of all individuals in the acquaintance network. In other words, according to the effect of joint function J, even though the statistical distribution of individual friend-making resources follows a heterogeneous power-law distribution pattern, the topological features of the simulated acquaintance network that reach a statistically stationary state should be very close to that of an acquaintance network with a homogeneous friend-making resource distribution pattern (e.g., delta or beta).

We ran a series of simulations using parameters initialized at different initial friendship f_0 and remembering q values to determine the effects of each on acquaintance networks. Our results indicate that a larger f_0 increased $\langle k \rangle$ but lowered

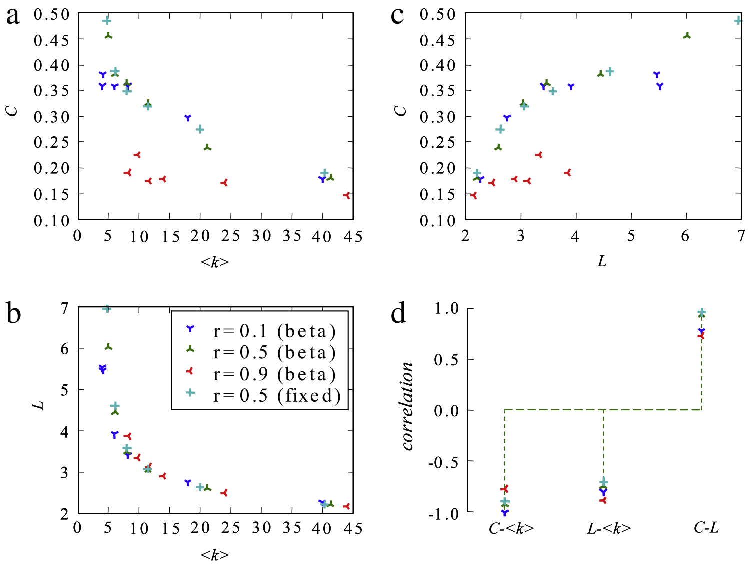

Fig. 15. Correlations among topological features according to various friend-remembering values and friend-making resource distributions.

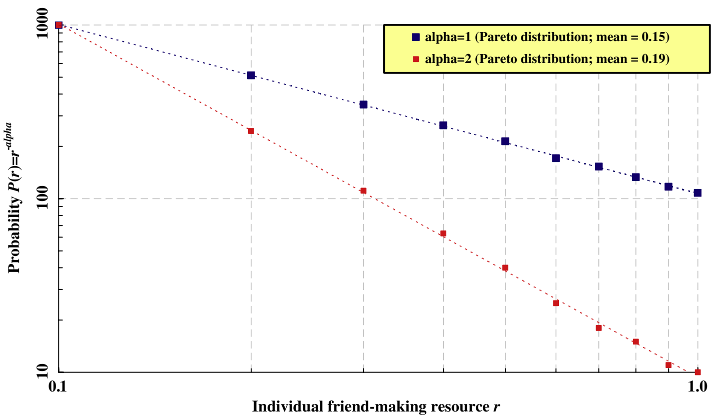

Fig. 16. Two power-law distributions of individual friend-making resources.

Table 6
Values for topological features in terms of various friend-remembering and friend-making resource levels. Two power-law (Pareto) and two delta (fixed value) distributions are shown.

q			0.0	0.2	0.4	0.6	0.8	0.9
r	pareto ($\alpha = 1.0$) mean = 0.15	$\langle k \rangle$	2.71	4.00	6.00	7.99	17.83	38.254
		$\langle k^2 \rangle$	10.88	23.68	60.07	116.72	571.24	2142.708
		C	0.3146	0.3613	0.3518	0.3520	0.2906	0.1777
		L	8.7266	5.6882	4.0422	3.4821	2.7740	2.2561
	pareto ($\alpha = 2.0$) mean = 0.19	$\langle k \rangle$	2.08	4.00	5.99	8.00	17.94	37.30
		$\langle k^2 \rangle$	6.56	24.41	63.62	128.22	585.16	2080.39
		C	0.2759	0.3747	0.3511	0.3625	0.2814	0.1989
		L	12.4231	5.5760	4.0443	3.4263	2.7553	2.3004
	fixed_value ($\mu = 0.15$)	$\langle k \rangle$	2.01	4.00	4.55	7.99	17.696	37.98
		$\langle k^2 \rangle$	5.78	24.47	31.34	116.99	603.72	2107.99
		C	0.2065	0.3473	0.3575	0.3499	0.2973	0.1753
		L	21.0000	5.6215	5.1167	3.4962	2.7743	2.2778
	fixed_value ($\mu = 0.19$)	$\langle k \rangle$	2.01	4.00	5.99	8.01	17.71	36.96
		$\langle k^2 \rangle$	5.78	24.35	60.93	117.17	577.51	2106.60
		C	0.2158	0.3492	0.3548	0.3478	0.2879	0.1999
		L	17.9657	5.5376	4.0378	3.4779	2.7680	2.3114

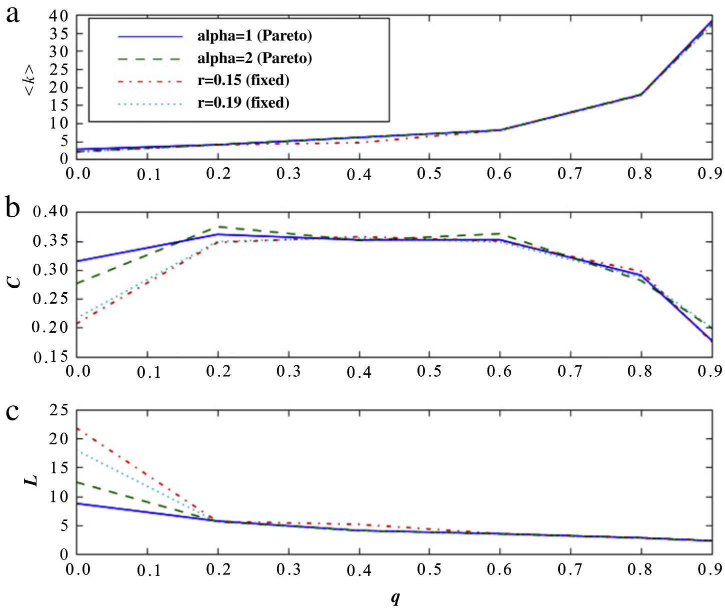

Fig. 17. Values of topological features according to various friend-remembering and friend-making resource levels. Two power-law (Pareto) and two delta (fixed value) distributions are shown.

Table 7
Topological feature values according to various friend-remembering levels and initial friendship distributions.

q			0.0	0.2	0.4	0.6	0.8	0.9
f_0	beta14 ($\mu = 0.1$)	$\langle k\rangle$	4.01	4.14	4.58	5.05	6.00	8.00
		$\langle k^2\rangle$	19.00	19.59	24.33	30.8	44.01	84.36
		C	0.5007	0.4788	0.4539	0.4512	0.3547	0.2886
		L	9.2204	9.7683	7.7325	6.6268	4.7783	3.7074
	beta14 ($\mu = 0.5$)	$\langle k\rangle$	4.74	6.00	7.03	9.52	15.96	29.98
		$\langle k^2\rangle$	25.41	47.06	69.95	141.19	442.85	1342.34
		C	0.4892	0.4023	0.3652	0.3232	0.2939	0.2065
		L	8.1108	4.5738	3.8898	3.2932	2.7749	2.3597
	beta14 ($\mu = 0.9$)	$\langle k\rangle$	4.66	5.99	7.99	10.59	20.04	40.04
		$\langle k^2\rangle$	25.24	45.92	92.82	179.43	625.14	2330.23
		C	0.5064	0.4068	0.3370	0.3031	0.2313	0.1691
		L	8.2640	4.7841	3.6030	3.1511	2.6172	2.2277
	fixed_value ($\mu = 0.5$)	$\langle k\rangle$	4.66	6.00	7.37	9.94	16.01	30.46
		$\langle k^2\rangle$	24.85	46.29	76.41	153.62	395.63	1352.15
		C	0.4950	0.4299	0.3670	0.3146	0.2312	0.1862
		L	7.9614	4.8324	3.8830	3.2369	2.7998	2.3541

both C and L (Table 7 and Fig. 18). Different initial friendship averages clearly influenced network topological features, but different initial friendship distributions did not. The Fig. 18 data also show that increases in q raised $\langle k\rangle$ and reduced C and L. Both $C - \langle k\rangle$ and $L - \langle k\rangle$ correlations were negative, while the C–L correlation was positive (Fig. 19). According to our results for the sensitivity analysis experiment in Fig. 18, we concluded that when acquaintance network members are more enthusiastic and make a stronger impression when meeting other members for the first time, the resulting friendships are larger in number and harder to break.

4.2. Discussion: Effects of rule parameters

All simulations were relationally cross-classified to analyze the effects of different parameters on our proposed three-rule model. The $+/-$ signs in Table 8 denote positive/negative relations between parameters and statistics; the $+/-$ signs and statistics in Table 9 denote direction and correlation strength. As shown in Table 8, increasing the values of Rule 3 parameters q, r, and f_0 triggered increases in $\langle k\rangle$ and decreases in C and L. However, the effects of increasing the value of the Rule 3 parameter θ were exactly opposite. Increasing $\langle k\rangle$ results in decreases in C and L, and decreasing $\langle k\rangle$ results in increases in C and L. Acquaintance network results from increasing the four parameters are inconsistent with the social network characteristics of high degree of clustering and low degree of separation—that is, acquaintance networks cannot maintain the two small-world properties simultaneously.

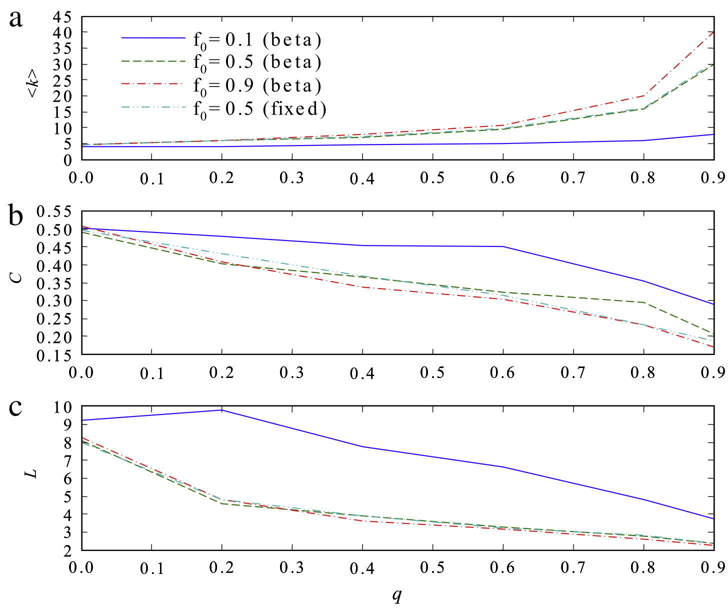

Fig. 18. Values of topological features according to various friend-remembering levels and initial friendship distributions.

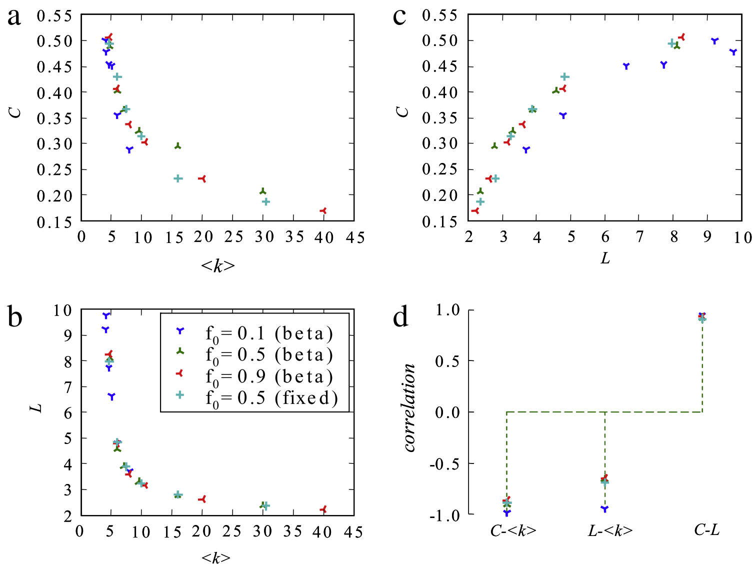

Fig. 19. Correlations among topological features according to various friend-remembering values and initial friendship distributions.

The affection triggered by increased p in Rule 2 is inconsistent with the affection triggered by q, r, f_0 and θ in Rule 3; moreover, an increase in p triggered decreases in $\langle k \rangle$ and C, as well as an increase in L. Taking into consideration the affection associated with the five parameters in Tables 8 and 9, we conclude that using only one of the three rules is insufficient for increasing C and decreasing L in an acquaintance network—that is, the small-world phenomenon can only be generated by interaction among the three rules.

4.3. Effects of sampling on acquaintance networks

Surveys, questionnaires, and sampling techniques stand at the center of traditional social science research and are considered cheaper and more practical than collecting and organizing large amounts of census data. With the exception of a few ingenious indirect studies (e.g., Milgram's [15] small-world experiments), social network data collection is usually carried out by querying participants directly (e.g., using questionnaires or conducting interviews). Such labor-intensive methods limit the size of networks that can be observed and measured. Survey data are influenced by subjective biases

Table 8
Correlations between Rules 2 and 3 parameters and acquaintance network topological features.

Statistics	Rule 2	Rule 3			
	p	q	θ	r	f_0
$\langle k \rangle$	−	+	−	+	+
C	−	−	+	−	−
L	+	−	+	−	−

Table 9
Correlations among acquaintance network topological features from sensitivity analysis experiments.

Variational parameters	$C - \langle k \rangle$	$L - \langle k \rangle$	C–L
p	+ + +	− −	− − −
θ, p	− −	− − −	+ + +
q, r	− − −	− −	+ + +
q, f_0	− − −	− −	+ + +

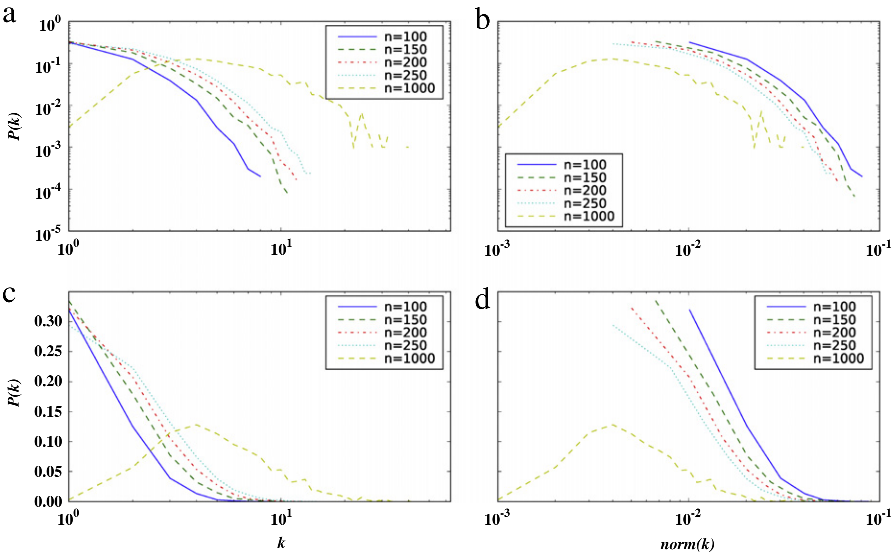

Fig. 20. Node degree distributions after sampling at 100, 150, 200, 250 nodes.

on the part of respondents: how one respondent distinguishes among friends, best friends, and confidants can be quite different from how others make the same distinctions. Although a great deal of effort may be put into eliminating potential inconsistencies, it is generally accepted that large and essentially uncontrolled errors exist in most social network studies. Based on our belief that the effectiveness of social network analytical methods has not been thoroughly examined, we ran the first application of our proposed simulation model in an effort to collect a sufficient node sample once our simulated acquaintance network reached a statistically stationary state. Initialized parameters for this experiment are presented in Table 1.

Node degree distribution results for sampling at 100, 150, 200, 250, 300, 500, 700, 900 nodes are shown in Figs. 20 and 21. Figs. 20(a), (b) and 21(a), (b) are log plots with log scaling on the X- and Y-axes; these were used to determine whether distributions were scale-free. Figs. 20(c), (d) and 21(c), (d) are semi-log plots with log scaling on the X-axis only; these were used to determine whether distributions were exponential. The degrees shown in Figs. 20(b)–(d) and 21(b)–(d) are post-normalization, as required for different numbers of sampled nodes. Each curve in Figs. 20 and 21 represents an ensemble average of 100 sampling repetitions. The solid lines in Fig. 20 reflect a lower sampling ratio of 0.1—considered common for traditional surveys and sampling techniques. The light dotted lines in Fig. 21 reflect a higher sampling ratio of 0.9—considered common for a census. Turns in Y-axis direction were observed for high sampling rates but not for low. As shown, node degree distribution clearly lost its original shape after sampling. According to the node degree distribution results in Figs. 20 and 21 after sampling the acquaintance networks, we conclude that experimental network-oriented simulations are a necessary aspect of social network research, not only due to the expenses and other difficulties involved with fieldwork, but also because widely used sampling approaches cannot capture real-world social network distributions, since distributions for higher sampling rates differ from those for lower sampling rates.

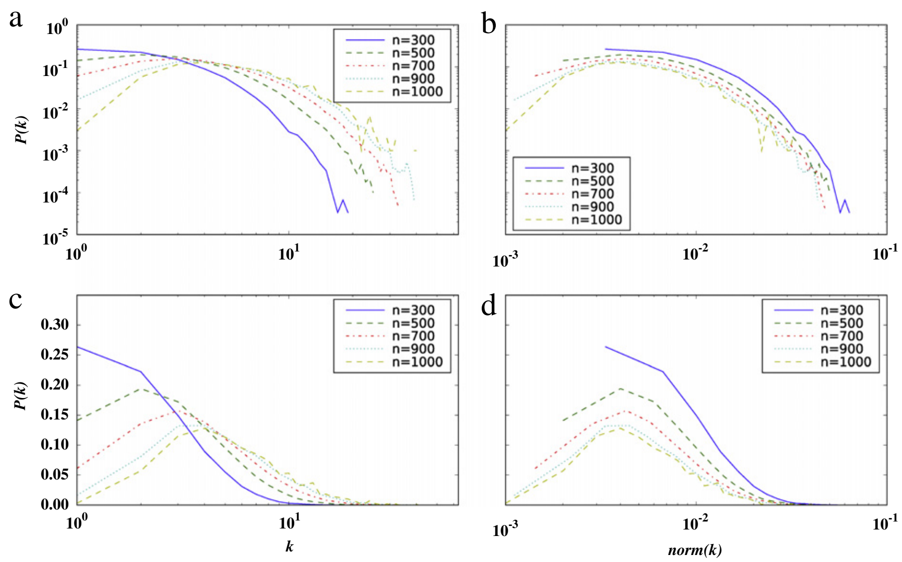

Fig. 21. Node degree distributions after sampling at 300, 500, 700, 900 nodes.

5. Conclusion

Most small-world models of social networks are analyzed using mixes of regular graphs with random networks—a practice based on Watts and Strogatz's [1] small-world model, which mixes one regular and one random network to facilitate theoretical analysis. While acknowledging the important impacts of their model, there is growing awareness that social network research should not be restricted to issues associated with separation and clustering [12].

Exploring how people make new friends is a meaningful and significant research task. In most cases people make new social connections via introductions by their friends, but there are many cases in which strangers become friends through chance meetings with no introductions. With few exceptions, most of us can only give limited attention or spend limited resources on friend making, meaning that friends once considered close can become distant over time. In this paper we proposed a three-rule model of network evolution to build a better understanding of acquaintance networks. In Rule 1, acquaintances are made via introductions and chance meetings. An aging factor is added in Rule 2, and friendships in Rule 3 are altered according to such factors as limited resources, friend remembering, breakup thresholds, and initial friendships. Taking a bottom-up, network-oriented simulation approach to modeling reflects the evolution mechanism of real-world social networks. Building on insights from previous studies, we applied local and interactive rules to acquaintance network evolution. This approach produces findings that can be used to explore human activity in specific social networks—for example, rumor propagation and disease outbreaks.

Acknowledgements

This work was supported in part by the Republic of China National Science Council (Grant No. NSC 98-2314-B-182-043), and Chang Gung Memorial Hospital (No. CMRPD260023).

References

[1] D.J. Watts, S.H. Strogatz, Collective dynamics of 'small-world' networks, Nature 393 (1998) 440–442.
[2] M.E.J. Newman, D.J. Watts, Renormalization group analysis of the small-world network model, Physics Letters A 263 (1999) 346–347.
[3] D.J. Watts, Networks, dynamics, and the small-world phenomenon, American Journal of Sociology 105 (1999) 493–527.
[4] D.J. Watts, Small Worlds: The Dynamics of Networks between Order and Randomness, Princeton University, Princeton, NJ, 1999.
[5] R. Pastor-Satorras, A. Vespignani, Epidemic spreading in scale-free networks, Physical Review Letters 86 (2001) 3200–3203.
[6] D.J. Watts, Six Degrees: The Science of a Connected Age, W.W. Norton, New York, NY, 2003.
[7] D.J. Watts, P.S. Dodds, M.E.J. Newman, Identity and search in social networks, Science 296 (2002) 1302–1305.
[8] D.H. Zanette, Dynamics of rumor propagation on small-world networks, Physical Review E 65 (2002) 041908.
[9] A.L. Barabási, R. Albert, Emergence of scaling in random networks, Science 286 (1999) 509–512.
[10] R. Albert, H. Jeong, A.L. Barabási, Error and attack tolerance of complex networks, Nature 406 (2000) 378–382.
[11] X. Li, G. Chen, A local world evolving network model, Physica A 328 (2003) 274–286.
[12] M.E.J. Newman, J. Park, Why social networks are different from other types of networks, Physical Review E 68 (2003) 036122.
[13] E.M. Jin, M. Girvan, M.E.J. Newman, Structure of growing social networks, Physical Review E 64 (2001) 046132.
[14] J. Davidsen, H. Ebel, S. Bornholdt, Emergence of a small world from local interactions: Modeling acquaintance networks, Physical Review Letters 88 (2002) 128701.
[15] S. Milgram, The small world problem, Psychology Today 2 (1967) 60–67.
[16] P. Erdös, A. Rényi, On the evolution of random graphs, Publications of the Mathematical Institute of the Hungarian Academy of Science 5 (1960) 17–61.

[17] B. Bollobàs, Random Graphs, 2nd ed., Academic, New York, NY, 2001.
[18] R.TH.A.J. Leenders, Evolution of friendship and best friendship choices, Journal of Mathematical Sociology 21 (1996) 133–148.
[19] G. Bianconi, A.L. Barabási, Bose–Einstein condensation in complex networks, Physical Review Letters 86 (2001) 5632–5635.
[20] M.E.J. Newman, D.J. Watts, S.H. Strogatz, Random graph models of social networks, Proceedings of the National Academy of Sciences of the United States of America 99 (2002) 2566–2572.
[21] H.J. Larson, Introduction to Probability, Addison-Wesley, Reading, MA, 1995.
[22] E.N. Bruce, Biomedical Signal Processing and Signal Modeling, Wiley, New York, NY, 2001.
[23] G.F. Davis, M. Yoo, W.E. Baker, The small world of the American corporate elite 1982–2001, Strategic Organization 1 (2003) 301–326.
[24] M.J. Conyon, M.R. Muldoon, The small world of corporate boards, Journal of Business Finance and Accounting 33 (2006) 1321–1343.