

Research Article

The Effects of Decision-Making Processes and Population Turnover on the Formation of Social Networks

Zhijing Xu, Zhenghu Zu, Tao Zheng, Wendou Zhang, and Qing Xu

Beijing Institute of Biotechnology, 20 East Avenue, Fengtai District, Beijing 100071, China

Correspondence should be addressed to Tao Zheng; ztl19721@hotmail.com

Received 13 March 2013; Accepted 28 April 2013

Academic Editor: Guanghui Wen

Copyright © 2013 Zhijing Xu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The researches on the structure and formation mechanism of social networks lead to several models with differences in the attachment patterns of new links (edges). In fact, the driving factor behind the addition of new links is just as important as the attachment patterns, while very little attention has been devoted so far to this exploration. We present an agent-based model which could successfully reproduce large-scale social networks. We find that the structure of social networks is a consequence of continuous individuals' decision-making processes based on self-evaluations and the turnover of the population. The individuals' self-evaluation processes are key motivating factors for the addition of new links, while the attachment patterns and the turnover of the population should be responsible for the topology of social networks. The resulting networks of our model display dynamics between order and randomness, which is greatly consistent with current observations and research achievements of social networks. We also find that some plausible properties of empirical data are actually artifacts due to the boundedness of sampling. Our research has revealed the driving factors behind the evolution of social networks as well as the underlying evolving patterns. These findings will lead to a better understanding of social structures.

1. Introduction

Complex networks have gained increasing enthusiasm in various fields ranging from natural science to social science in the last few years. The achievements of data acquisition make it possible to calibrate some hypotheses once supposed to be reasonable in people's mind, one of which is networks of complex topology described with the random graph theory of Erdős-Rényi (ER) [1], whose degree obeys a Gaussian distribution. Barabási and Albert (BA) have developed a growing network model to interpret the emergence of scaling in networks, known as BA model [2], which leads to a degree distribution $P(k) \sim k^{-3}$. However, it is obviously inappropriate for applying it mechanically to social networks for two reasons: first, the population is relatively constant, which does not conform with the hypothesis of unceasing increasing of nodes; second, individuals could not be present in all the lifetime of social networks, instead, the living individuals will die after a few years' survival.

Researchers had realized these problems not long after BA model was proposed [3–7]. Furthermore, the analysis of empirical data [8, 9] shows that some social networks exhibit

single-scale properties for the degree distribution rather than power law regime. So González et al. [5] and Singer et al. [7] introduce constraints named "aging effect" to limit the addition of new links and the results indicate that these new observations are well fitted. However, there are still some problems with these models: (i) the empirical networks used in these models come from an in-school questionnaire among junior high school students from the USA (for acquaintance network, it comes from a questionnaire among 43 Utah Mormons), so are these social networks from specific groups universal and representative? Furthermore, could these models be used to explain the formation of large-scale social networks? (ii) Constraints introduced by these researches are more passive responses to the dilemma of continuous addition of new links, which could not lead to a deep understanding of the driving factors behind the addition of new links. We have developed a new agent-based model for large-scale friendship networks, which successfully reproduces acceptable properties of social networks such as small-world phenomenon and community structures [10, 11]. We also find that the resulting degree distribution of social networks displays obvious scale-free regime, indicating that the observed

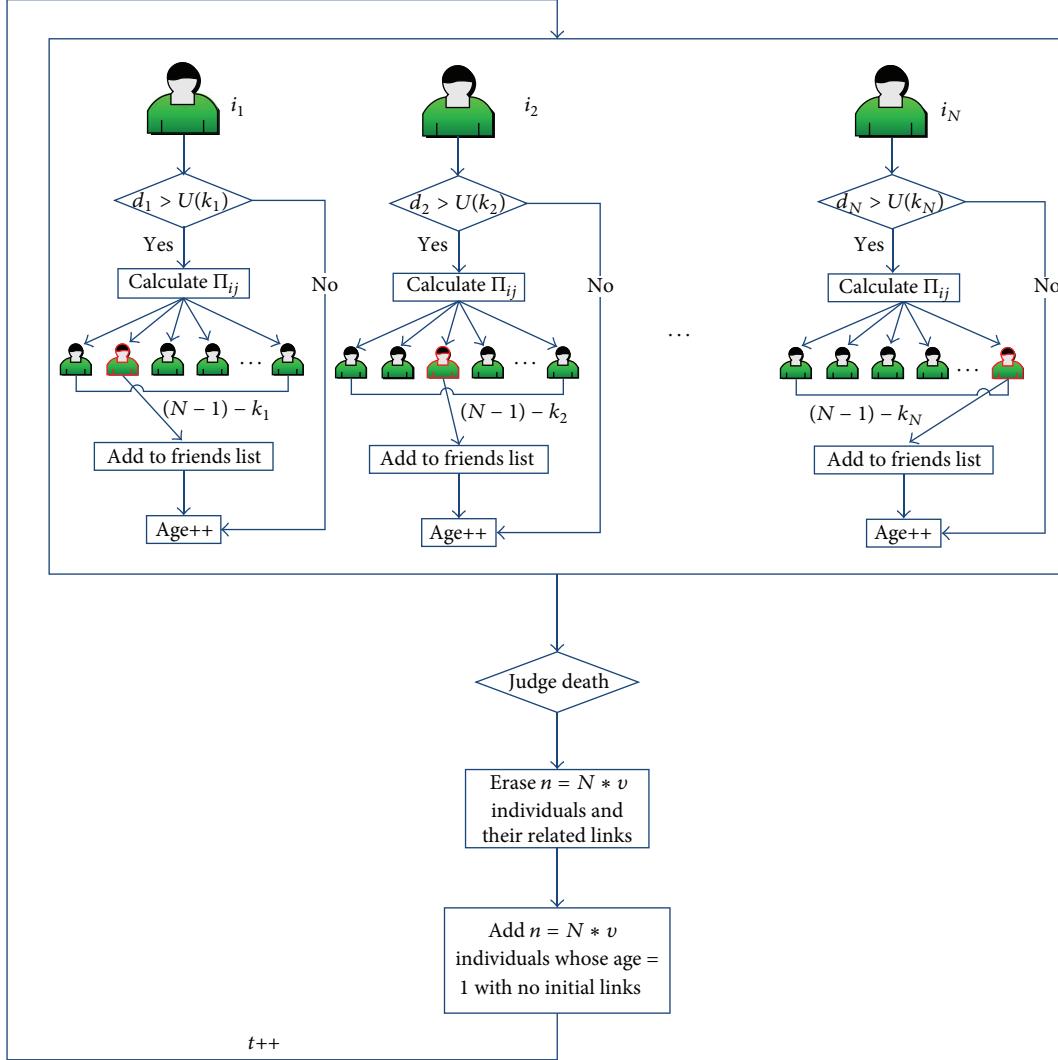


FIGURE 1: The model framework. At each time step t , every individual i will evaluate his satisfaction based on current friend circles. Here, we simply generate a random number d ranging from 0 to 1. If $d > U(k)$, individual i is not satisfied with current situation, and so he chooses a new friend j from those used to be strangers (the red individual) and adds j into his friends list. Meanwhile, j also adds i into his friends list. After all the individuals complete the decision-making processes independently, the network updates with $n = N * v$ aged individuals along with their links removed from the network and the same number of newborns with no initial links added into the network. Then time step t increases by 1.

single-scale properties of empirical networks are artifacts resulting from the boundedness of sampling. These findings give a clear answer to the questions above and confirm the drawbacks in the previous models.

2. Materials and Methods

There are two key features of real social networks neglected by foregoing researches: (i) the continuous turnover of the population—in real worlds, the aged die at a certain rate while new individuals are born at an appropriate rate resulting in an almost unchanged number of individuals present at the networks over time; (ii) on-going individuals' decision-making processes—the living individuals carry out social activities based on their current neighbors in the network including families, schoolmates, colleagues, and sometimes

strangers and in the case of dissatisfaction with current circles of friends, they will try to build new relationships with some strangers. These ingredients are incorporated into our model framework. We also introduce an index named social activeness index, which is a comprehensive measurement about individual's activeness in social networks, denoted by λ . An active individual will be likely to make more friends. Integrated with the definition of natural attributes, we finally get an agent-base model in which individuals are characterized by a set of natural and social attributes, including identity number, age, residence, and social activeness index, along with behavioral rules.

At the beginning of the modeling, there are N individuals (agents) with no initial connections, who are randomly placed into an area of $x * y$ grids. This is close to the reality that the city is usually composed of a number of communities

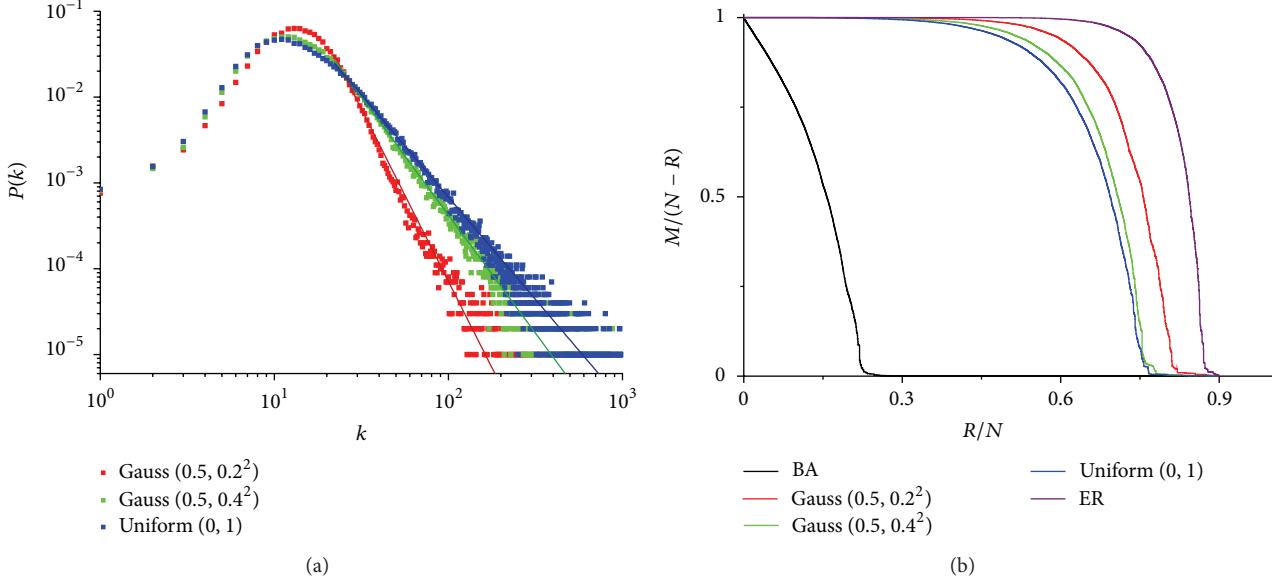


FIGURE 2: The resulting networks at $t = 2000$ for $N = 100000$, $v = 0.0005$, $\alpha = 0.2$, $\beta = 4$, $\gamma = 4$, and $p = 0.0001$ within an area of $10 * 10$ grids. (a) The distribution functions of connectivity: for red and green squares, $\rho(\lambda)$ is of truncated normal distribution $\text{Gauss}(\mu, \sigma^2)$; for blue squares, $\rho(\lambda)$ is of uniform distribution $\text{Uniform}(0, 1)$. The red, green, and blue lines have respectively slopes 4.02, 2.75, and 2.38. (b) The robustness of obtained networks: for red and green curves, $\rho(\lambda)$ is of truncated normal distribution $\text{Gauss}(\mu, \sigma^2)$. The red curve has $\zeta = 0.82$ while the green one has $\zeta = 0.76$; for blue curve, $\rho(\lambda)$ is of uniform distribution $\text{Uniform}(0, 1)$ and $\zeta = 0.74$; purple curve for ER, $\zeta = 0.92$, and black curve for BA, $\zeta = 0.16$.

and different individuals share the same community. For simplicity, we assume that the identity number of individuals increases from 0, so an elder individual will get a smaller identity number. Each individual's age is an integer selected randomly from 1 to max_age ($\text{max_age} = 1/v$, where v is the death rate of the population). Social activeness index (λ) is an important parameter indicating individual's social activeness, which is assigned a value ranging from 0 to 1 generated from a specific distribution $\rho(\lambda)$ when an individual is born. Figure 1 illustrates our model framework. At each time step, every individual will evaluate the situation of his social contacts based on the current number of neighbors in the network (degree). Here, we introduce the utility function [12] which is widely used to quantify consumer's total satisfaction from consuming a good or service in economics because of the observed striking similarities among many collective human activities [13–17]. We employ an exponential utility function $U(k) = 1 - e^{-\lambda k}$, where k is individual's degree and λ is the social activeness index. So $U(k)$ decreases with the decreasing of λ , that is, the individual would be more dissatisfied with the current situation and more likely to make a new contact. Under the circumstance of dissatisfaction, individual i will make a new friend with j based on three criteria: age difference (a_{ij}), grid distance (d_{ij}), and the number of shared friends (m_{ij}). These criteria have been observed in many social networks and employed by some foregoing studies [3, 5, 11, 18], while never integrated into a single model. So the probability i making a new contact with j is given by

$$\prod_{ij} = \frac{R_{ij}}{\sum_{l \neq i} R_{il}}, \quad (1)$$

where

$$R_{ij} = \begin{cases} p \cdot e^{-(\beta a_{ij} + \gamma d_{ij})} & k_i = 0, \\ \left(\left(\frac{m_{ij}}{k_i} \right)^\alpha (1-p) + p \right) \cdot e^{-(\beta a_{ij} + \gamma d_{ij})} & k_i \neq 0, \end{cases} \quad (2)$$

where k_i is the degree of i , α, β, γ are scale factors, and $p \ll 1$. Obviously, the effects of age difference and grid distance on social networks increase with the increasing of α and β , while γ does the opposite to m_{ij}/k_i . After every individual makes his decision independently, the network updates. We assume that the friendship is reciprocal, when i adds j into his friends list and so does j . Then a small number (n) of individuals are removed from the network along with the links with them and the same amount of individuals with no initial links is added, respectively, corresponding to individual's death and birth in real world. The number is determined by the death rate v with $n = N * v$ (noticing that we assume that the population is relatively constant, so v is also the birth rate).

3. Results and Discussion

In Figure 2(a), we show the degree distributions of truncated normal distribution $\rho(\lambda)$ with different μ and σ (the uniform distribution could be regarded as a limit form of truncated normal distribution with a large σ). Note that λ is the individual's social activeness index, so a normal distribution implies that most of individuals have similar social activeness, while only few of them are very active or inactive. With the increasing of σ , λ will be much more diverse. Our modeling suggests that the heavy tail will be more straightforward with

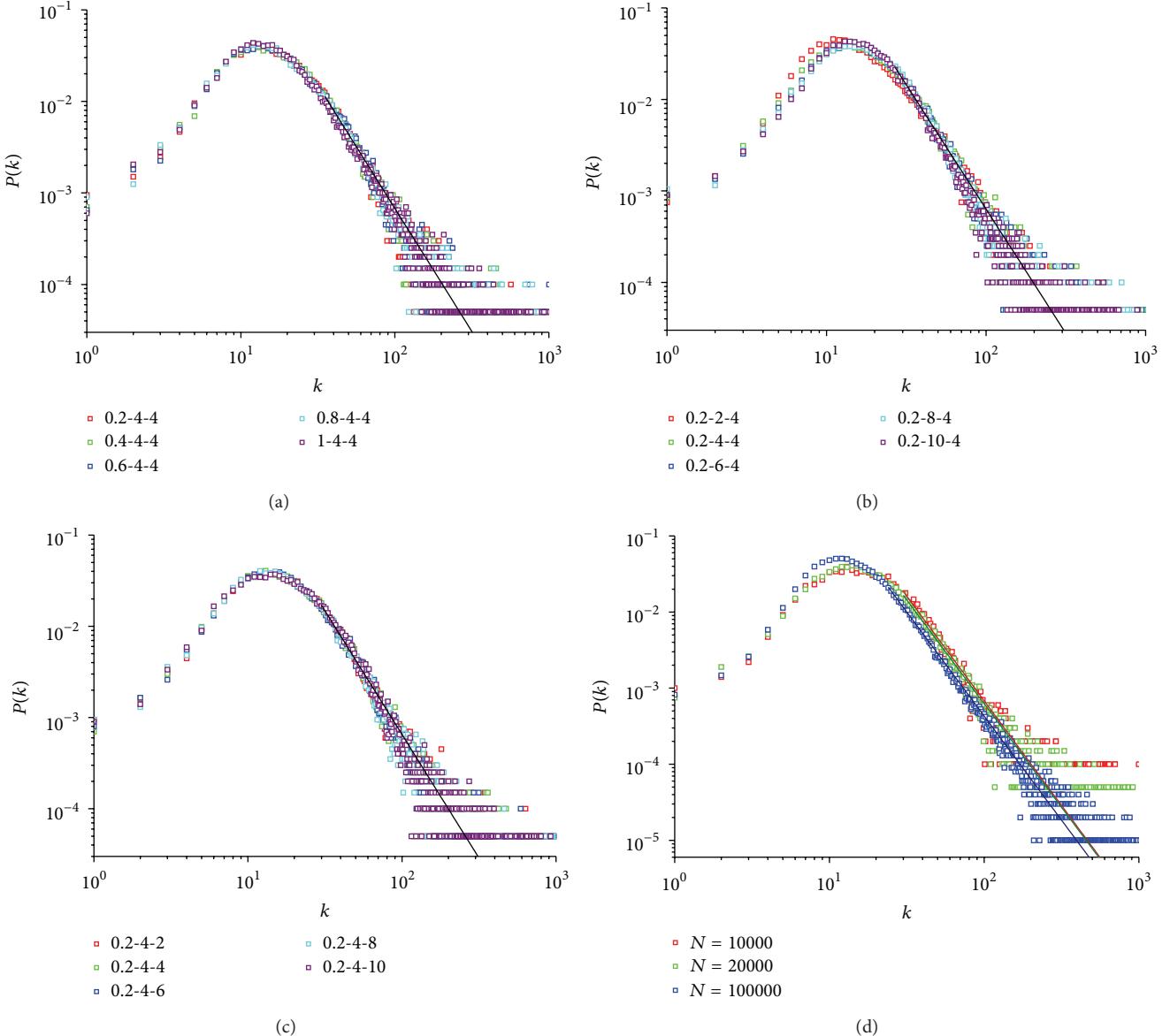


FIGURE 3: The sensitivity analysis. (a) $\rho(\lambda) = \text{Gauss}(0.5, 0.40^2)$ with different α . The black solid line has slope 2.69. (b) $\rho(\lambda) = \text{Gauss}(0.5, 0.40^2)$ with different β . The black solid line has slope 2.72. (c) $\rho(\lambda) = \text{Gauss}(0.5, 0.40^2)$ with different γ . The black solid line has slope 2.71. (d) $\rho(\lambda) = \text{Gauss}(0.5, 0.20^2)$. The black solid line has slope 4.01.

a decreasing of the scaling exponent. When coming to the uniform distribution, the scaling exponent is approximately equal to 2.38. Sensitivity analyses of N , α , β , and γ show that they do not significantly change the scaling exponent (see Figure 3).

When exploring various complex networks, we have a gut feeling that social networks are of strong robustness. It is easily accountable—individuals' death should not greatly shatter social networks, although they are hubs (people with a lot of links). So we compare the output of our model to some other networks, for example, random network from ER theory and scale-free network from BA model. As Figure 2(b) shows, the obtained network's robustness is between BA and ER. If we regard the BA network as a well-ordered network and

correspondingly the ER network as a random network, the obtained networks will be between order and randomness, just as Watts [18] described. In order to make the contrast of different networks more clearly, we define a shatter index ζ as

$$\zeta = \frac{1}{N} \sum_{R=1}^N \frac{M}{(N-R)}, \quad (3)$$

where R denotes the number of nodes removed, and M is the nodes' number of the maximal connected subnetwork in the remained network after R nodes have been removed. Noticing that $\zeta = 1$ for a full connected network, the geometric meaning of ζ is actually the area enclosed by curves in Figure 2(b) and the coordinate axes.

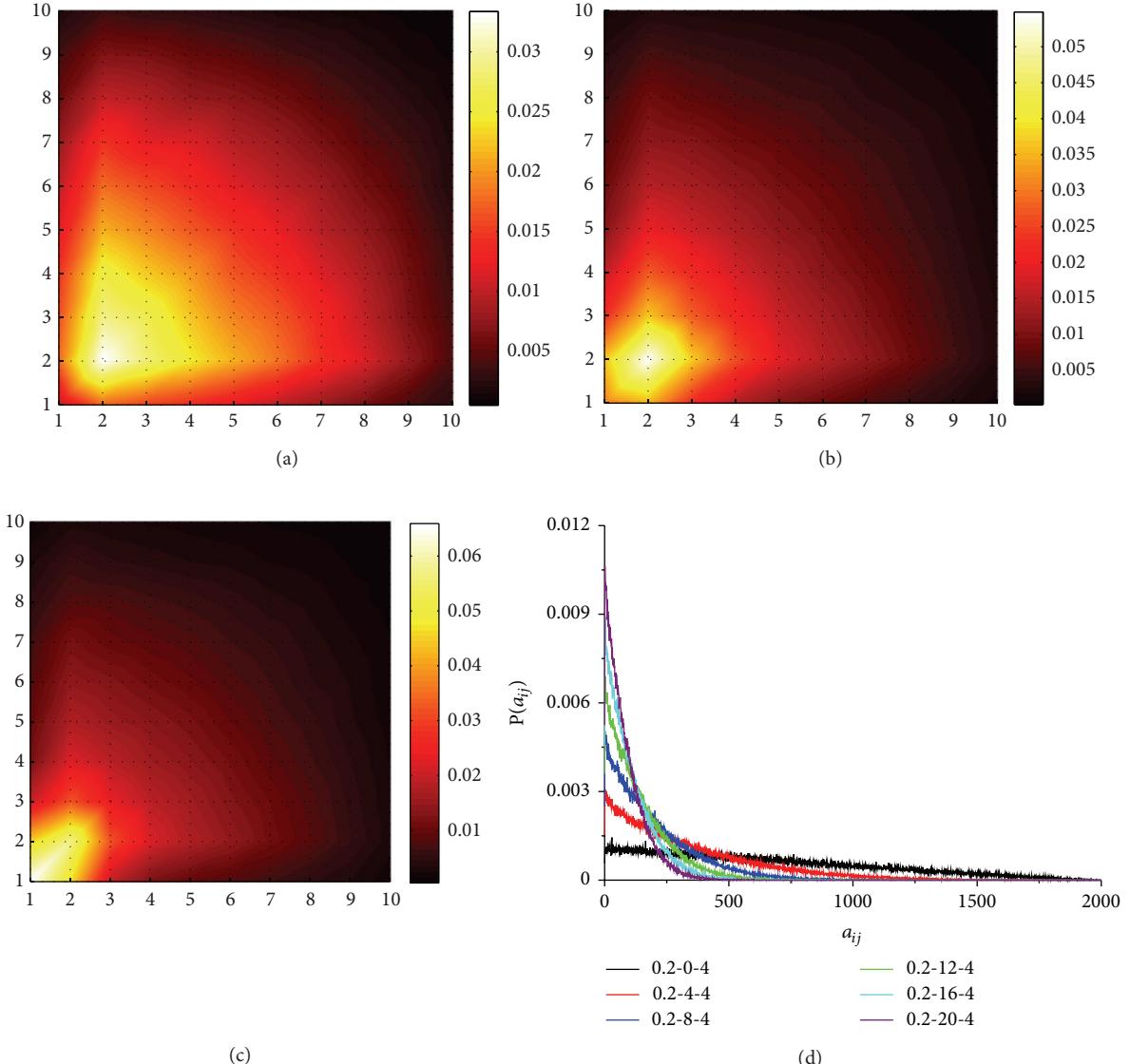


FIGURE 4: The grid distance and age difference between connected nodes at $t = 2000$ for $N = 20000$, $v = 0.0005$, and $p = 0.0001$ and $\rho(\lambda)$ is of truncated normal distribution with $\mu = 0.5$, $\sigma = 0.20$ within an area of $10 * 10$ grids. In (a), (b), and (c), the color denotes the proportion of node pairs with grid distance (x, y) in total nodes pairs. All of them use $\alpha = 0.2$ and $\beta = 4$, while $\gamma = 0$ for (a), 10 for (b), and 20 for (c). (d) The age difference between connected node pairs uses $\alpha = 0.2$, $\gamma = 4$, and different β (see the legend $\alpha\text{-}\beta\text{-}\gamma$). The horizontal axis denotes age difference, while the vertical axis denotes the proportion of node pairs with age difference a_{ij} in total node pairs.

We also check the grid distance and age difference between friends in the networks. As Figures 4(a), 4(b), and 4(c) show, the regional structure will be more conspicuous with the increasing of γ . Figure 4(d) gives the analysis result of age difference between node pairs in obtained networks with different β . A significant positive correlation between the age effect and β emerges from the comparison of different curves.

The shortage of empirical large-scale social networks makes a straightforward verification of the obtained networks from comparison impossible. However, researches in related areas have provided some strong circumstantial evidences. In

the past few years, the tight correlation between the spreading of infectious diseases and social networks has aroused widespread concerns. Eubank et al. [19] proposed a large-scale simulation framework based on realistic urban social networks. Though edges in their networks are defined as contacts between individuals, they also give a lot of valuable information about friendship networks because of the interplay between contact network and friendship network. Namely, close friends tend to make more contacts, while frequent contacts will lead to an intimate relationship. Their studies show that real networks are strongly connected small-world graphs with a well-defined scale for the degree

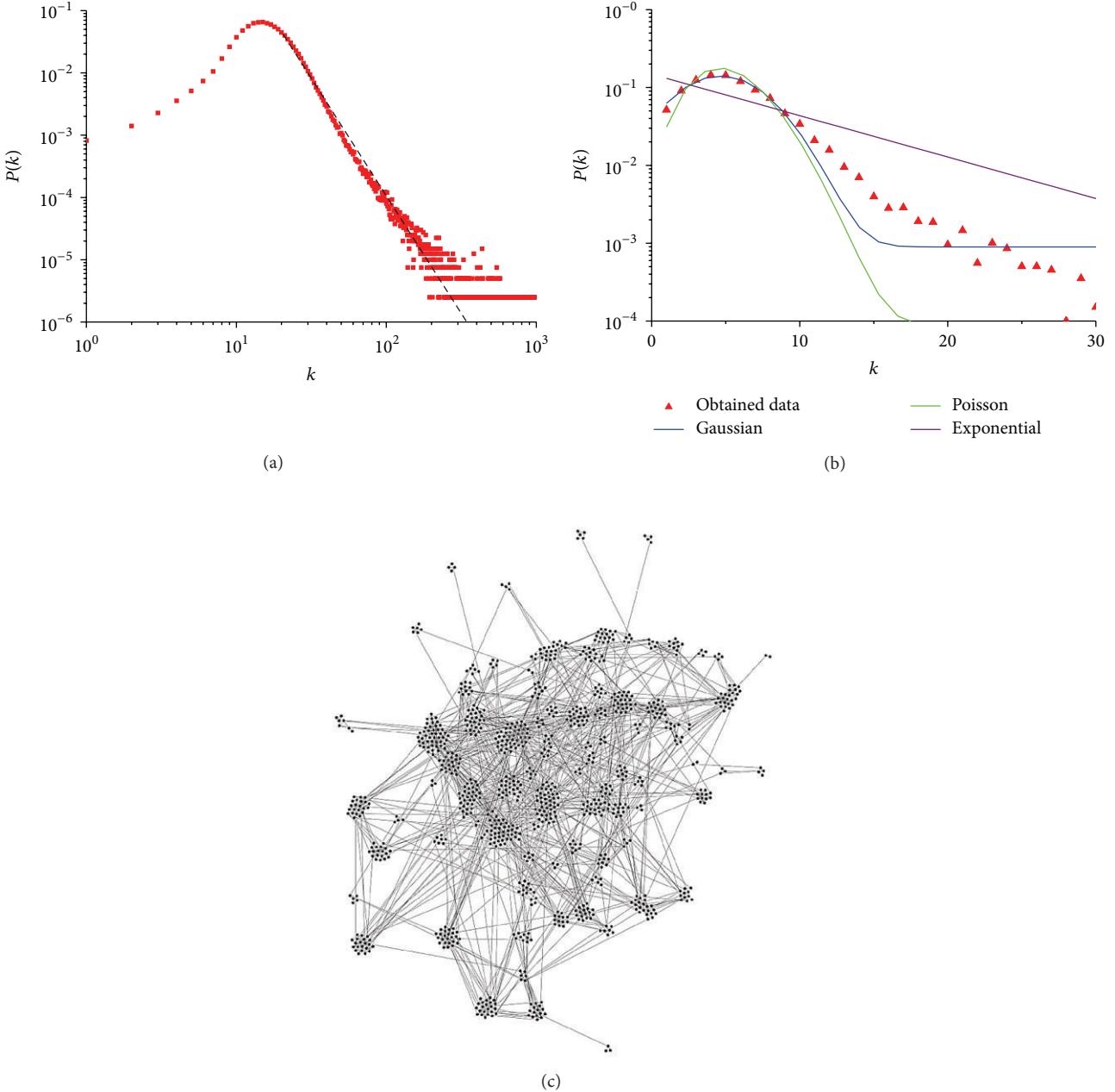


FIGURE 5: The in-school friendship networks at $t = 2000$ using $N = 20000$, $v = 0.0005$, $p = 0.0001$, $\mu = 0.5$, $\sigma = 0.20$, $\alpha = 0.8$, $\beta = 16$, and $\gamma = 2$ within an area of $10 * 10$ grids. The results have been averaged over 20 realizations. (a) The degree distribution of obtained full social networks. The dashed line has slope 4.02. (b) The degree distribution of sampled in-school friendship network. (c) A snapshot of the in-school friendship networks visualized with the Gephi software.

distribution and could not be shattered by removing a small number of high-degree vertices, which coincides with the obtained network of our model.

Next, we will show by sampling individuals from the obtained social networks and reconstructing their interconnections that our model successfully reproduces the observed single-scale properties in the in-school friendship networks. Owing to the obvious age effect in in-school friendship networks, we employ a large β , indicating that when individual makes a new friend, there is an obvious tendency to

individuals from the same generation. After evolution over few time steps, we pick up a group of individuals of similar age as well as their interconnections. In this way, we get a subnetwork corresponding to in-school friendship network, as shown in Figure 5(c). It is visually apparent that the obtained in-school friendship network is composed of a number of different size communities. Figure 5(b) shows the connectivity distribution of in-school friendship network, a striking similarity between them and the observations [5, 7, 20] could easily be detected; however, the degree

distribution of the full social network exhibits a visible heavy tail property (Figure 5(a)), indicating that the observed single-scale properties in empirical in-school friendship and acquaintance networks are more artifacts than universal phenomena, leading to an inability of foregoing models to describe social systems.

4. Conclusions

In this study, we have uncovered the structure and formation of social networks by capturing some important features of real world. This has significant potential in interpreting lots of social phenomena related to human activities. Further study will focus on the effects of social networks on individual's migration, which may help understand heterogeneity of human geographical distribution.

Conflict of Interest

The authors declare that they have no conflict of interests.

Acknowledgments

This study was funded by the Ministry of Science and Technology of China Grant 2013ZX10004605 and the National Natural Science Foundation of China (NSFC) Grant 90924019.

References

- [1] P. Erdős and A. Rényi, *On the Evolution of Random Graphs*, Akad. Kiadó, 1960.
- [2] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *American Association for the Advancement of Science. Science*, vol. 286, no. 5439, pp. 509–512, 1999.
- [3] L. A. N. Amaral, A. Scala, M. Barthélémy, and H. E. Stanley, "Classes of small-world networks," *Proceedings of the National Academy of Sciences*, vol. 97, no. 21, pp. 11149–11152, 2000.
- [4] G. Caldarelli, A. Capocci, P. De Los Rios, and M. A. Muñoz, "Scale-free networks from varying vertex intrinsic fitness," *Physical Review Letters*, vol. 89, no. 25, Article ID 258702, 4 pages, 2002.
- [5] M. C. González, P. G. Lind, and H. J. Herrmann, "System of mobile agents to model social networks," *Physical Review Letters*, vol. 96, no. 8, Article ID 88702, 4 pages, 2006.
- [6] J. M. Kumpula, J. P. Onnela, J. Saramäki, K. Kaski, and J. Kertész, "Emergence of communities in weighted networks," *Physical Review Letters*, vol. 99, no. 22, Article ID 228701, 4 pages, 2007.
- [7] H. M. Singer, I. Singer, and H. J. Herrmann, "Agent-based model for friendship in social networks," *Physical Review E*, vol. 80, no. 2, Article ID 026113, 5 pages, 2009.
- [8] H. R. Bernard, P. D. Killworth, M. J. Evans, C. McCarty, and G. A. Shelley, "Studying social relations cross-culturally," *Ethnology*, vol. 27, no. 2, pp. 155–179, 1988.
- [9] T. J. Fararo and M. H. Sunshine, *A Study of a Biased Friendship Net*, Youth Development Center, Syracuse University, 1964.
- [10] G. Palla, I. Derényi, I. Farkas, and T. Vicsek, "Uncovering the overlapping community structure of complex networks in nature and society," *Nature*, vol. 435, no. 7043, pp. 814–818, 2005.
- [11] G. Palla, A. L. Barabasi, and T. Vicsek, "Quantifying social group evolution," *Nature*, vol. 446, no. 7136, pp. 664–667, 2007.
- [12] J. E. Ingersoll, *Theory of Financial Decision Making*, Rowman & Littlefield Pub Inc, 1987.
- [13] P. W. Anderson, "Nonlinear dynamics, mathematical biology, and social science," *Complexity*, vol. 3, no. 2, pp. 50–51, 1997.
- [14] T. Lux, M. Marchesi, and B. Universität, "Scaling and criticality in a stochastic multi-agent model of a financial market," *Nature*, vol. 397, no. 6719, pp. 498–500, 1999.
- [15] X. Gabaix, P. Gopikrishnan, V. Plerou, and H. E. Stanley, "A theory of power-law distributions in financial market fluctuations," *Nature*, vol. 423, no. 6937, pp. 267–270, 2003.
- [16] A. L. Barabasi, "The origin of bursts and heavy tails in human dynamics," *Nature*, vol. 435, no. 7039, pp. 207–211, 2005.
- [17] J. C. Bohorquez, S. Gourley, A. R. Dixon, M. Spagat, and N. F. Johnson, "Common ecology quantifies human insurgency," *Nature*, vol. 462, no. 7275, pp. 911–914, 2009.
- [18] D. J. Watts, *Small Worlds: The Dynamics of Networks between Order and Randomness*, Princeton University Press, 1999.
- [19] S. Eubank, H. Guclu, V. S. Kumar et al., "Modelling disease outbreaks in realistic urban social networks," *Nature*, vol. 429, no. 6988, pp. 180–184, 2004.
- [20] K. M. Harris and J. R. Udry, *National Longitudinal Study of Adolescent Health (Add Health), 1994–2008*, Inter-University Consortium for Political and Social Research (ICPSR), 2012.

