Physica A 463 (2016) 485–492

Contents lists available at ScienceDirect

Physica A

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

Reconstruction of social group networks from friendship networks using a tag-based model

Yuan-Pan Guan, Zhi-Qiang You, Xiao-Pu Han *

Alibaba Research Center for Complexity Sciences, Hangzhou Normal University, Hangzhou 311121, China

HIGHLIGHTS

- A tag-based model is proposed to explain the mechanism on the growth of social groups.
- In the model, social groups expand on a friendship network based on users' tags of interest.
- Users' activity in joining group is related to their degree of friendship network.
- Various distributions of the simulated group network are in agreement with empirical findings.
- Our model throws light on the reconstruction of institute-based relationships.

ARTICLE INFO

Article history:
Received 29 October 2015
Received in revised form 27 June 2016
Available online 22 July 2016

Keywords:
Social groups
Social group networks
Friendship networks
Tag of interest
Network reconstruction

ABSTRACT

Social group is a type of mesoscopic structure that connects human individuals in microscopic level and the global structure of society. In this paper, we propose a tag-based model considering that social groups expand along the edge that connects two neighbors with a similar tag of interest. The model runs on a real-world friendship network, and its simulation results show that various properties of simulated group network can well fit the empirical analysis on real-world social groups, indicating that the model catches the major mechanism driving the evolution of social groups and successfully reconstructs the social group network from a friendship network and throws light on digging of relationships between social functional organizations.

1. Introduction

Gathering groups is widespread in societies covering different cultures and different historical periods [1]. Human individuals form associations, organizations and institutions, and there are usually relatively stable members, clear social tags and boundaries in each of these social institutes. In this paper, this type of social institutes is termed as "*social groups*". Generally, different from the widely-discussed implicit structure termed as "community" [2,3], social group is a type of explicit mesoscopic structure of society because of their clear tags and boundaries. Since these tags usually come from the real-world social functional institutes/organizations, social groups will have strong coincidence with these social institutes/organizations. For example, a user usually would like to join a group whose major members are his/her colleagues or teammates. This correlation actually provides a possible way to investigate the relationships and effective organizations between real-world functional institutes from the group information of online societies. And also, online groups usually are widely used to be a place for information releasing and public discussing, and they therefore play an important role in the spreading of online information and the formation of public opinions.

* Corresponding author.
E-mail address: xp@hznu.edu.cn (X.-P. Han).

http://dx.doi.org/10.1016/j.physa.2016.07.020
0378-4371/

Previous researches mainly focus on the implicit communities in social networks [2,3]. In this issue, topology properties of networks have aroused concerns [4–7], and serials of methods to identify and define communities in networks, as well as some confinement [8], have been proposed [9–14]. Community extraction arouses the greatest interest in the domain of social network analysis as well. In Ref. [15], the studies of community structures are allowed to promote further to in a general setting encompassing networks that evolve over time, have multiple types of links (multiplexity), and have multiple scales. Furthermore, finding communities among the system users opens the new possibilities and can be utilized in such disciplines as sociology, biology, and computer science [2] and others [12]. And also, some works explored the evolution of communities from social networks [16–20].

Nevertheless, the research on social groups is still rare. It is partially because of the lack of real-world datasets relating to social groups. Recently, You et al. [21] empirically analyzed the real-world online social groups using Tencent QQ dataset and reported many anomalous properties, including sudden growth of group size, wide-spread scaling anomalies on group size, degree and weight of connection network between groups, and special age effects and gender differences. Similar findings recently were also reported by the analysis for the group information on Tencent Wechat [22]. These findings indicate that there will be rich anomalous patterns hidden in real-world social groups, and raise the questions that what mechanism drives the emergence of these patterns and what impacts of social groups on social dynamics. Furthermore, these empirical findings will uncover some features of the organization of real-world social functional institutes/organizations because of the strong coincidence of social groups. In this sense, the mechanism studies on the evolution of social groups will be helpful for the understanding for the structure of real-world social institutes/organizations.

In this paper, we propose a model based on tag-driven expanding on social networks to mimic the expansion of social groups and reconstruct their connections from friendship networks. Numerical simulations of the model generate rich properties that can generally cover the empirical observations, indicating that the model successfully explains the origin of these anomalous properties of real-world social groups.

2. The model

Since the relevant empirical studies are mainly aimed at the online social groups, the rules of our model also are in the light of the online cases. Generally, a social group has a clear tag, and users who are interested in the tag or belong to the institution attaching to the one tag will join the group. For example, a tag naming "Complex networks" will attract many researchers studying complex networks. After a group is created by a user, its information will spread in his/her circles, and the group will spring from the creator to his/her friends and then extend to their further neighbors if the tag of the group can well fit their interests. This expansion is similar to the spreading of meme in society [23,24]. Actually, in a sense, social groups also can be considered to be a type of meme.

Actually, besides the groups that can well fit the real-world social organizations or institutes and mainly contain strong social contacts, for example, the members in a club or a research team often build groups in social societies, from the daily life experience by the using of online social group, there is another type of group that usually has a tag of common interests (e.g. music, pop stars, etc.) and contains many weak ties. Similar two types of groups have been observed in other group-like online societies, like Tencent Wechat [22]. Empirical studies are difficult to distinguish the two types of groups and we do not know the real-world proportions of the two types directly. Nevertheless, Ref. [21] reported that real-world groups globally have the property of sudden growth, indicating that users usually would like to join a group that they have already intended to join or been interested in before the creation of the group, and thus the strong-tie-driven groups are the vast majority. We therefore construct our model from the tag-driven contacts between users and their neighboring groups.

Moreover, even though our datasets neither support the comparative study in individual level and nor provide direct evidence, it is natural in turn to suppose that users with more social contacts have more possibilities to join groups.

From the perceptions above, our model is therefore based on real-world social networks with the tag-driven growing process and various activities for group joining. For the coincidence with the empirical results of social groups of Tencent QQ, the friendship network of Tencent QQ users is used to be as a background network. The detailed description and sampling method of the dataset of both social groups and friendship networks can be found in the Appendix. With the friendship network, we run the model with the following rules:

(i) Each user (node) in the friendship network is granted an N dimensional binary tag vector $\mathbf{H} = (h_1, h_2, \ldots, h_N)^T$, $h_i = 0$ or 1, $i = 1, 2, \ldots, N$. Here, each h_i represents a tag in the node's interests, and $h_i = 1$ if the user has the corresponding tag and $h_i = 0$ if the user has not the corresponding tag, and N represents the total types of tags. Since each user is assumed to own at least one tag of interest, one randomly chosen component of its tag vector $\mathbf{H}$ is set as nonzero for each node at first. As for each of the rest components of $\mathbf{H}$, it is randomly to be nonzero with a probability ω (see Fig. 1).

(ii) The total number of groups that will be created in the model is $M_G = 3{,}432{,}642$, which is the total number of groups averaged by 10 independently-sampled group sets (see Appendix). The creating method is as follows: each node (user) in the friendship network firstly creates a group, and here we have M_U groups ($M_U = 1{,}052{,}199$ is the total number of nodes in the friendship network). For each of the remaining $M_G - M_U$ groups, with probability ρ, we randomly pick an edge from the friendship network and randomly choose a node on the edge to be the creator of the group; or with probability $1 - \rho$, we randomly pick a node to be the creator.

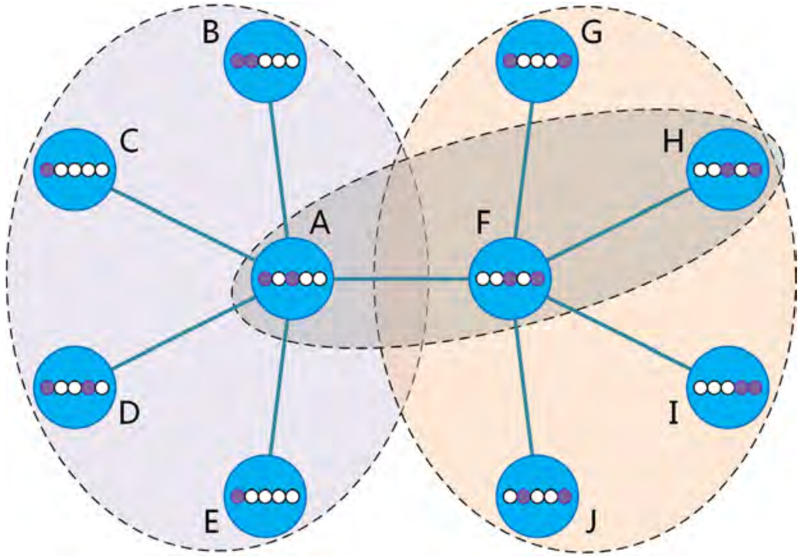

Fig. 1. (color online) Illustration of the possible expansion range of groups on friendship network. For the concision of the illustration, we suppose each user (denoted by solid blue circles) has 5 tags, i.e. $N = 5$. The violet dots and the white dots in the graph corresponding to $h_i = 1$ and $h_i = 0$, respectively. Here, the 3 dashed circles symbolize that there are 3 potential groups to be created by user A and F, which are, respectively, based on the first tag of interest of user A and the third and fifth tag of user F. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(iii) We randomly pick a nonzero tag from the creator's tag vector to be the only nonzero element for the tag vector of the group. For example, for a node's **H**, if $h_i = 1$, it will get possibilities, to create a group with tag h_i.

(iv) With the definition that a user–group pair is created once a node joins a group, we set the total number of user–group pairs is $N_G = 3{,}709{,}712$, which is the number of real-world user–group pairs averaged by 10 independently-sampled group sets. M_G user–group pairs have been created in the procedures above, and the remaining $N_G - M_G$ pairs is created by the method similar to the above creator-picking: with probability ρ, we randomly choose a node from a randomly-picking edge of the friendship network; or with probability $1 - \rho$ we just randomly pick a node; then we connect the node to a group, which is randomly picked from the groups that the nodes' neighbors has joined and have a same tag of interest with the node, to create the user–group pair. If none of neighbors' group satisfies the same-tag condition, we repeat this selection till another pair is created.

There is not any dataset of social groups including the information of both social relationships and group institutes released so far. Due to the absence of correspondences between the user information in the two datasets, we have to compare the model running on a sampled friendship network with the empirical findings for sampled social groups. To keep comparability, the total of joined users in each of sampled social group is kept to be equal to the user number of the friendship network. More detailed sampled method is introduced in Appendix. Actually, this model describes a method that reconstructs relationships between social groups or social institutes from a friendship network.

3. Modeling results

Due to the limitation of datasets' information, it has some difficulties in the discussion at individual level. What we concerned here is the topological properties between different social groups. A natural connection between different social groups is by their common members. We therefore construct a network to describe the relationship between different groups using the following method that was used in the empirical analysis of social group reported by Ref. [21]: each group is a node of the group network, and each pair of groups is connected by an edge if they have at least one common member, and the number of common member is the weight, denoted as w, of the edge. In the following discussions, this network is called "group network" and symbolized by **G**.

Empirical findings indicate that the real-world group network is scale-free and has heterogeneous weighted degree distribution [21], therefore the following distributions of the modeling group network are mainly compared: (i) the group size distribution $p(S)$; (ii) the distribution $p(k_H)$ of the number of joined groups by individual users; (iii) the degree distribution $p(k_G)$ of group network **G**; (iv) the distribution $p(k_{GW})$ for the weighted degree of group network **G**, here for each node (node j, say), its weight degree k_{GW} is defined as the total of edge weight of all its connected edges: $k_{GW} = \Sigma_{k_G} w_G$.

During the simulation of our model, we have found that many similar statistical characteristics between the empirical findings and our modeling results come up. As shown in Fig. 2, with suitable parameter settings, the major region in the modeling distributions generally can be well parallel to the empirical curves, indicating that our model indeed catches the

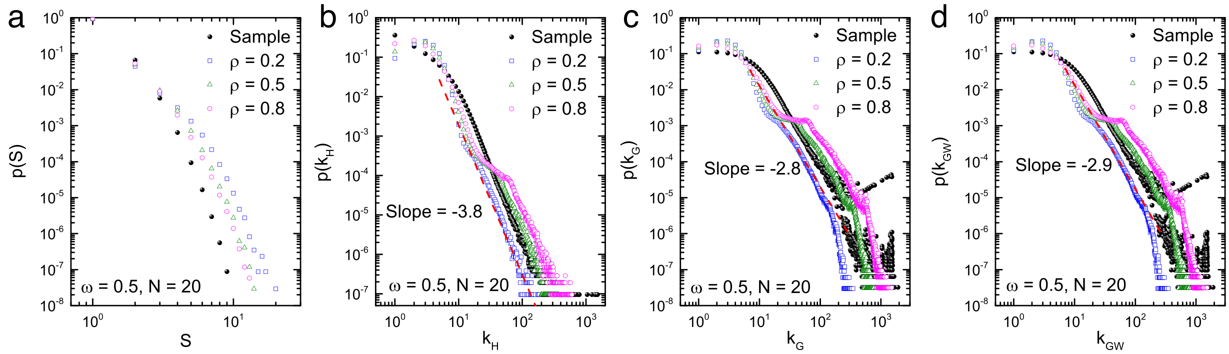

Fig. 2. (color online) Statistical patterns of the simulated group network for different ρ settings. Panels (a), (b), (c) and (d), respectively, are the group size distribution $p(S)$, the distribution $p(k_H)$ of the number of joined groups by individual users, the degree distribution $p(k_G)$ and the weighted degree distribution $p(k_{GW})$ of group network **G**. The black spheres show the empirical results, and other data points are the simulation results, and parameters $\omega = 0.5, N = 20$. The red dashed lines in panels (b), (c) and (d), respectively, show the power laws with exponent -3.8, -2.8 and -2.9. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

underlying mechanism driving the growth of social groups. Since our model do not consider the impact of limitation rules of Tencent QQ on group size, the tails of the modeling group size distribution are higher than the one of empirical results, as shown in Fig. 2(a).

The model is driven by a hybrid mechanism including various activities on group-joining and tag-based expanding and has three major parameters: ρ, ω and N. For larger ρ, users with more social contacts have more certain probability to join groups. Fig. 2(a) shows that the modeling group size distributions $p(S)$ trend to close to the empirical results under the setting with larger ρ. As shown in Fig. 2(b), the modeling distribution $p(k_H)$ decays in the same tail slope as the empirical curve. The tail slope is about -3.8 and is insensitive to the value of ρ, however, the bump on the head of $p(k_H)$ is stronger with a larger ρ setting, since $p(k_H)$ is mainly dependent on the degree distribution $p(k_F)$ of the friendship network and the bump on $p(k_F)$ (see Fig. A.5 in Appendix) will also impact on $p(k_H)$ when ρ is large. Similar properties are also observed in the degree distribution $p(k_G)$ and the weighted degree distribution $p(k_{GW})$ of group network **G** (Fig. 2(c) and (d)).

Parameters ω and N determine the number of common tags of interests between two neighboring nodes. Our model limits the expanding of groups along the edges with common tag of interests. Larger ω means each pair of neighboring users have more possibility to have a common tag of interest, and in this case groups generally easily expand in social circles. However, simulations find out that all the distributions are insensitive to ω (Fig. 3), which partially results from that only a little part of potential user–group pairs can be created with the limitation M_G on the total of user–group pairs, as well as the selection rule for the creating of user–group pairs in the model. We therefore compare the simulation results under different limitation on the total of user–group pairs. Obviously, a fat tail on the group size distributions $p(S)$ emerges under a larger limitation $2M_G$ (see Fig. 3(a)). As shown in Fig. 3(b), since the dependence on the degree distribution of friendship network, $p(k_H)$ keeps slope for different limitations. Nevertheless, the slope of $p(k_G)$ and $p(k_{GW})$ obviously reduces when the limitation grows, along with the relaxation on the bump on curves (Fig. 3(c) and (d)), indicating that the consistency between the simulated $p(k_G)$ and $p(k_{GW})$ and empirical curves is based on the condition of the same total of user–group pairs. In addition, it implies that the origin of the bump on $p(k_G)$ and $p(k_{GW})$ is not only because of the impact of the bimodal-like degree distribution of friendship network, but also the insufficient group-joining of users.

Moreover, setting the limitation on the total of user–group pairs at $2M_G$, the effect of ω is more obvious: small ω causes more heterogeneous form on the three structural distributions (Fig. 3(b), (c) and (d)), because in this case the hubs of friendship network still keep a higher possibility in successfully joining in a group under a situation with more frequent reselection at individual level, but the higher failure probability is mainly taken by the nodes in periphery under the matching condition of tag. The reselection probability λ for different ω can be found in the inset of Fig. 3(d). It is important to notice the condition where the comparison between the modeling results and empirical findings is M_G, the limitation of total user–group pairs, even though in the above discussions we also show the case with a different limitation to discuss the detailed dynamics in the model. At last, keeping other parameters' setting constant, the total type N of tags does not show significant impact in the model (Fig. 4).

4. Discussions

The core of the proposed model is the hypothesis that, active users will have more possibilities to join groups, and every user in the friendship network has some tags of interest more or less, and they will create or join groups within their tags of interest. As a minimum model, neglecting several realistic cases that users possibly withdraw from groups, or groups can be dismissed is inevitably only an approximation for the process of real emergence. Nonetheless, to do this would introduce more parameters and cloud the basic result: the tendency of an individual to join a group is not only influenced by its tags of interest, but also crucially related to its current external state in the social environment. The basic hypothesis of the model

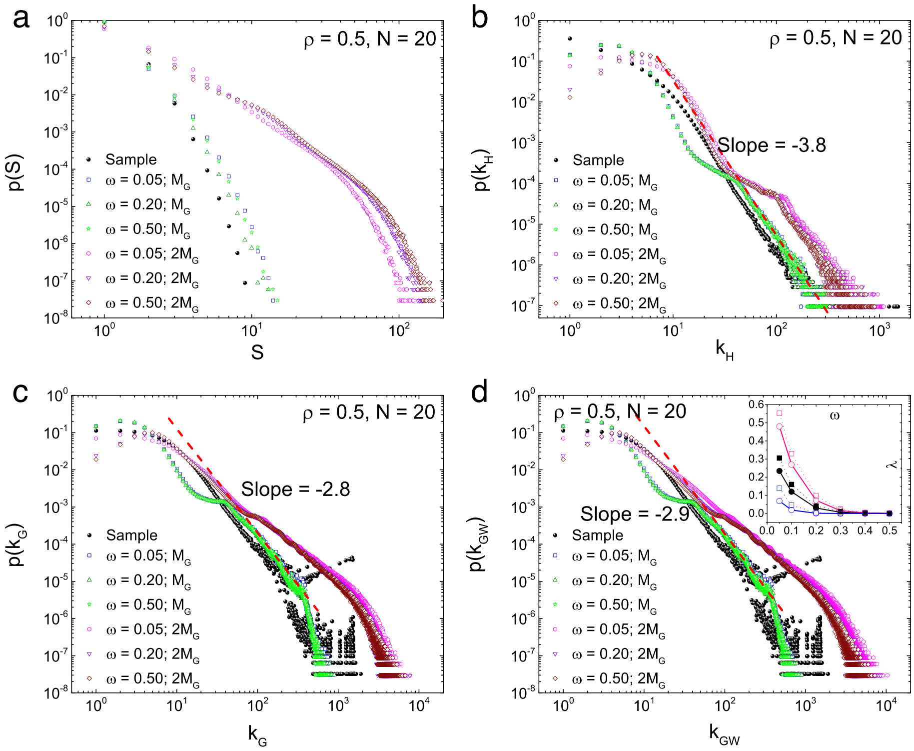

Fig. 3. (color online) Statistical patterns of the simulated group network for different ω settings and different settings on total of user–group pairs. Panels (a), (b), (c) and (d), respectively, are the group size distribution $p(S)$, the distribution $p(k_H)$ of the number of joined groups by individual users, the degree distribution $p(k_G)$ and the weighted degree distribution $p(k_{GW})$ of group network **G**. The black spheres show the empirical results, and other data points are the simulation results, where ω settings, respectively, are 0.05, 0.20 and 0.50, and the limitations on the total of user–group pairs, respectively, are M_G and $2M_G$, and parameters $\rho = 0.5, N = 20$. The red dashed lines in panels (b), (c) and (d), respectively, show the power laws with exponent −3.8, −2.8 and −2.9. The inset in panel (d) shows the probability λ of reselection of a node in group joining vs. the value of ω, where the spheres and squares, respectively, show the global averaged λ with different limitation M_g and $2M_g$, and the blue and pink circles/blocks, respectively, show the averaged λ for large degree nodes ($k_F > 10$) and small degree nodes ($k_F \le 10$) with limitation M_g (circles) and $2M_g$ (block) on the total of user–group pairs. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

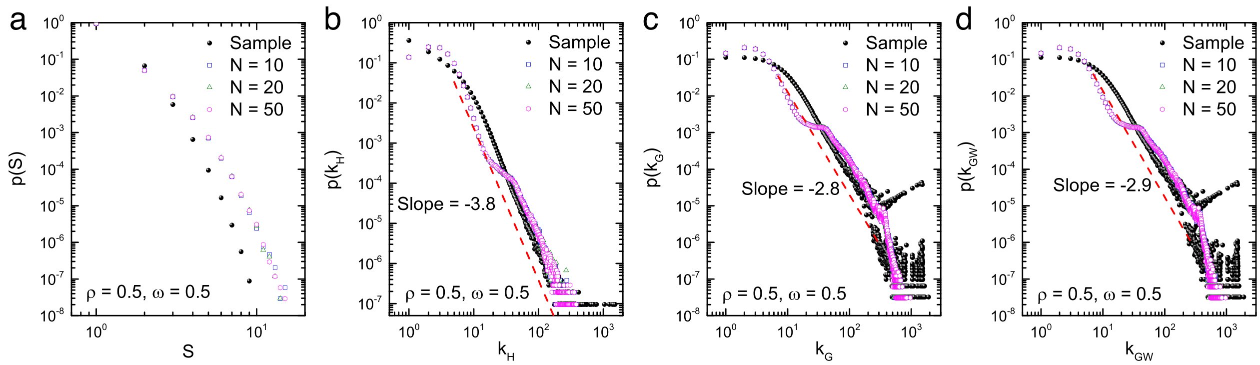

Fig. 4. (color online) Statistical patterns of the simulated group network for different N settings. Panels (a), (b), (c) and (d), respectively, are the group size distribution $p(S)$, the distribution $p(k_H)$ of the number of joined groups by individual users, the degree distribution $p(k_G)$ and the weighted degree distribution $p(k_{GW})$ of group network **G**. The black spheres show the empirical results, and other data points are the simulation results, and parameters $\rho = 0.5, \omega = 0.5$. The red dashed lines in panels (b), (c) and (d), respectively, show the power laws with exponent −3.8, −2.8 and −2.9. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

is verified by the high consistency between the simulation results and the empirical findings. This mechanism provides a heuristic insight to the understanding for both the emergence and the growth of social groups and the reconstruction of social group network from friendship network.

Moreover, with another perspective, because online social groups usually correspond to offline functional social institutes/organizations, our model actually provides a method to reconstruct the relationships between these social institutes/organizations from friendship networks. Obviously, these institute-based relationships are important for social functional mechanism. However, partially because of large number of non-occupational social contacts in many online societies, it is often difficult to directly analyze these institute-based relationships. Our model thus throws light on this problem: we can firstly reconstruct a basic network using our model to represent the relationships between institutes from friendship network, and then amendment its detailed connections with the supporting information.

And also, as a major platform for information publishing and public discussing, social groups and their background functional social institutes/organizations play an important role in the online information spreading and the public opinion. With the understanding of community topology properties and overlapping community structure of complex networks in nature and society [25], the issues about public opinion emergence, information spreading and the co-evolution with social structures or identifiable social tags [26–30] are becoming more prominent and more relaxable to analyze. Similarly, it is expected that the properties and the mechanism can supply some heuristic sight to these issues. Consequently, our model also shows a profitable significance for the research of these issues.

In summary, our study uncovers the underlying mechanism driving the emergence and growth of social groups and finds a way to statistically reconstruct social group networks from a friendship network, and bridges the gap between the action of microscopic individuals and the emergence of macroscopic structures of social networks by social groups in mesoscopic level.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grants Nos. 11305043, 61304150, and 61503110), the research startup fund of Hangzhou Normal University (Grant No. 2015QDL005), the European Commission FP7 Project GROWTHCOM (Grant No. 611272), and the CCF-Tencent Open Research Fund.

Appendix. Datasets description and sampling

Tencent QQ (the website of Tencent QQ: http://www.qq.com) is an instant communication tool developed by Tencent Holdings Limited. It now has over 700 million active users and has become the largest online society in China.

The real-world social group dataset of Tencent QQ was released from the online open database: http://qun.col.pw. It covers more than 58,523,079 groups and 274,335,183 users, of which 48,676,355 groups has the information with all ID, member list, and date. The detailed analysis of this dataset can be found in Ref. [21].

In this paper, the model runs on a background friendship network. To be comparable, the friendship network of Tencent QQ users is used here. The dataset of this friendship network, with 1,052,129 anonymous users and 8,022,535 friendship edges, was released from the CCF-Tencent Open Research project (http://ur.tencent.com) in 2014.

This friendship network was sampled from the global Tencent database using the following method: (i) 10,000 randomly picked users was assumed as seed nodes, which were within active users, as was defined as those who had logged at least one time within 30 days before this extraction time and whose registration time had been more than 1 year; (ii) add all the friends of those seed users into the dataset; (iii) supplement the edges having been in the global database among these nodes in the dataset yet not included in the dataset. And the network we are adopting is the largest connected component.

The average node clustering coefficient and the average distance of this sampled friendship network are 0.609 and 4.167, respectively. Its degree distribution is heterogeneous and bimodal-like, and the latter region can be well fitted by power law with slope −1.69, as shown in Fig. A.5.

It is noticed that the release of the dataset of friendship network is independent with the dataset of QQ group, i.e., the former is not the subset extracted from the latter. Therefore we are unable to know the corresponding relations between users' information in the two datasets. In our discussions, we use the simulated results to compare with the average distributions of 10 independent group samples with same user size.

The sampling method of real-world social groups includes the following steps:

(i) we randomly pick M_U users from the social group datasets as a sample of user set, and we assume these M_U users are "known" users, where $M_U = 1{,}052{,}199$ is the total number of users in the friendship network;
(ii) all the groups that these M_U users joined to construct a sample of real-world social groups;
(iii) remove all the group members that are not among the set of M_U users from the sample of groups.

In this way, the sampled group set contains all the groups that related to the sampled user set and has none of "unknown" users, which can be comparable with the modeling results that reconstructed from the friendship network with same size of user set. However, some global distributions of the sampled group set have differences with the patterns of total groups. Due to the sampled groups only keep the members in the set of M_U users, the average size of the sampled groups is largely reduced, and thus the group size distribution $p(S)$ of the sampled groups is much narrower than the one of population, and the deep bumps on the population distribution $p(S)$ caused by the limitation rule of Tencent QQ on group size are no longer obvious, as shown Fig. A.6(a). The distributions of $p(k_H)$ are well matched together (Fig. A.6(b)), and the middle region of $p(k_G)$ and $p(k_{GW})$ of the sampled group set keeps parallel to the one of total groups (see Fig. A.6(c) and (d)). The obvious branches in the tail of $p(k_G)$ and $p(k_{GW})$ of the sampled group set appear since few users in the sampled set join a large number of groups.

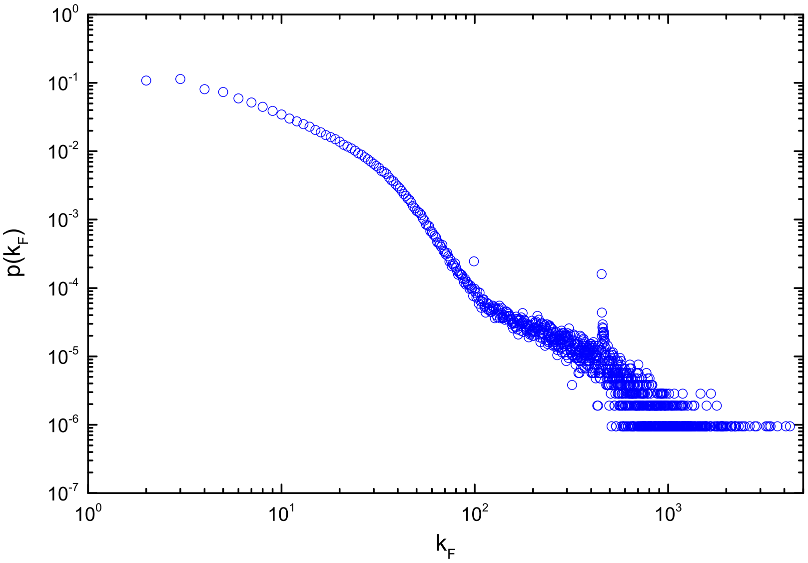

Fig. A.5. (Color online) The degree distribution $p(k_F)$ of the QQ friendship network.

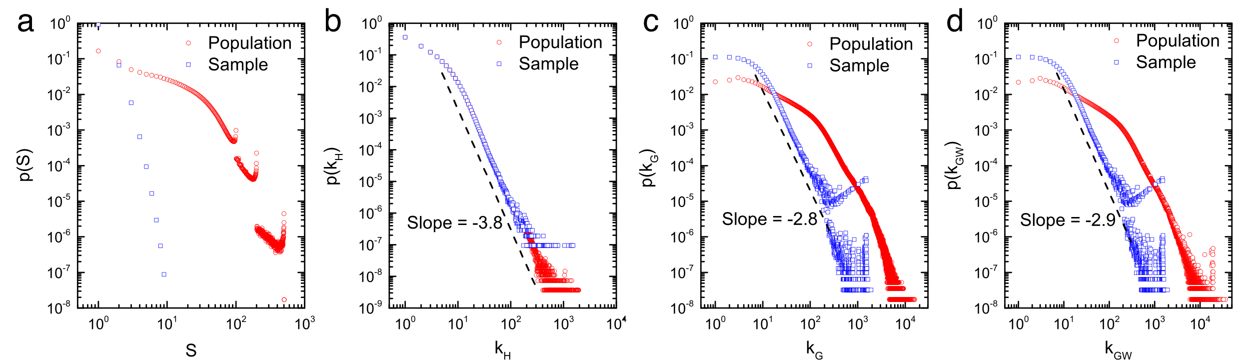

Fig. A.6. (color online) Comparison of four distributions of the population of QQ groups and group samples. Panels (a), (b), (c) and (d), respectively, are the group size distribution $p(S)$, the distribution $P(k_H)$ of the number of joined groups by individual users, the degree distribution $P(k_G)$ of group network, and the weighted degree distribution $P(k_{GW})$ of group network. The statistics of group samples are averaged by 10 independent samplings (blue data points). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

References

[1] J. Coleman, Foundations of Social Theory, Harvard, 1990.
[2] S. Fortunato, Community detection in graphs, Phys. Rep. 486 (35) (2010) 75–174. http://dx.doi.org/10.1016/j.physrep.2009.11.002.
[3] M.E.J. Newman, Detecting community structure in networks, Eur. Phys. J. B 38 (2) (2004) 321–330. http://dx.doi.org/10.1140/epjb/e2004-00124-y.
[4] M.E.J. Newman, M. Girvan, Finding and evaluating community structure in networks, Phys. Rev. E 69 (2) (2004) 026113. http://dx.doi.org/10.1103/PhysRevE.69.026113.
[5] M. Girvan, M.E.J. Newman, Community structure in social and biological networks, Proc. Natl. Acad. Sci. USA 99 (12) (2002) 7821–7826. http://dx.doi.org/10.1073/pnas.122653799.
[6] R.M. Shiffrin, K. Börner, Mapping knowledge domains, Proc. Natl. Acad. Sci. USA 101 (suppl 1) (2004) 5183–5185. http://dx.doi.org/10.1073/pnas.0307852100.
[7] A. Arenas, L. Danon, A. Daz-Guilera, P. Gleiser, R. Guimer, Community analysis in social networks, Eur. Phys. J. B 38 (2) (2004) 373–380. http://dx.doi.org/10.1140/epjb/e2004-00130-1.
[8] S. Fortunato, M. Barthélemy, Resolution limit in community detection, Proc. Natl. Acad. Sci. USA 104 (1) (2007) 36–41. http://dx.doi.org/10.1073/pnas.0605965104.
[9] B.W. Kernighan, S. Lin, An efficient heuristic procedure for partitioning graphs, Bell Syst. Tech. J. 49 (2) (1970) 291–307. http://dx.doi.org/10.1002/j.1538-7305.1970.tb01770.x.
[10] F. Radicchi, C. Castellano, F. Cecconi, V. Loreto, D. Parisi, Defining and identifying communities in networks, Proc. Natl. Acad. Sci. USA 101 (9) (2004) 2658–2663. http://dx.doi.org/10.1073/pnas.0400054101.
[11] M.E.J. Newman, Fast algorithm for detecting community structure in networks, Phys. Rev. E 69 (6) (2004) 066133. http://dx.doi.org/10.1103/PhysRevE.69.066133.
[12] P. Bródka, T. Filipowski, P. Kazienko, An introduction to community detection in multi-layered social network, in: Information Systems, E-Learning, and Knowledge Management Research, Springer, 2013, pp. 185–190. http://dx.doi.org/10.1007/978-3-642-35879-1_23.
[13] V.D. Blondel, J.-L. Guillaume, R. Lambiotte, E. Lefebvre, Fast unfolding of communities in large networks, J. Stat. Mech. Theory Exp. 2008 (10) (2008) P10008. http://dx.doi.org/10.1088/1742-5468/2008/10/P10008/.

[14] P. Expert, T.S. Evans, V.D. Blondel, R. Lambiotte, Uncovering space-independent communities in spatial networks, Proc. Natl. Acad. Sci. USA 108 (19) (2011) 7663–7668. http://dx.doi.org/10.1073/pnas.1018962108.

[15] P.J. Mucha, T. Richardson, K. Macon, M.A. Porter, J.-P. Onnela, Community structure in time-dependent, multiscale, and multiplex networks, Science 328 (5980) (2010) 876–878. http://dx.doi.org/10.1126/science.1184819.

[16] R. Toivonen, J.-P. Onnela, J. Saramäki, J. Hyvönen, K. Kaski, A model for social networks, Physica A 371 (2) (2006) 851–860. http://dx.doi.org/10.1016/j.physa.2006.03.050.

[17] G. Palla, A.-L. Barabási, T. Vicsek, Quantifying social group evolution, Nature 446 (7136) (2007) 664–667. http://dx.doi.org/10.1038/nature05670.

[18] Y. Wang, B. Wu, N. Du, Community evolution of social network: feature, algorithm and model. arXiv:0804.4356.

[19] L. Backstrom, D. Huttenlocher, J. Kleinberg, X. Lan, Group formation in large social networks: membership, growth, and evolution, in: ACM SIGKDD 12, ACM, 2006, pp. 44–54. http://dx.doi.org/10.1145/1150402.1150412.

[20] J.P. Bagrow, D. Brockmann, Natural emergence of clusters and bursts in network evolution, Phys. Rev. X 3 (2) (2013) 021016. http://dx.doi.org/10.1103/PhysRevX.3.021016.

[21] Z.-Q. You, X.-P. Han, L.-Y. Lü, C.H. Yeung, Empirical studies on the network of social groups: the case of tencent QQ, PLoS One 10 (7) (2015) e0130538. http://dx.doi.org/10.1371/journal.pone.0130538.

[22] J. Qiu, Y. Li, J. Tang, Z. Lu, H. Ye, B. Chen, Q. Yang, J. Hopcroft, The lifecycle and cascade of social messaging groups. arXiv:1512.07831.

[23] X. Wei, N. Valler, B. Prakash, I. Neamtiu, M. Faloutsos, C. Faloutsos, Competing memes propagation on networks: A network science perspective, IEEE J. Sel. Areas Commun. 31 (6) (2013) 1049–1060. http://dx.doi.org/10.1109/JSAC.2013.130607.

[24] J.P. Gleeson, J.A. Ward, K.P. O'Sullivan, W.T. Lee, Competition-induced criticality in a model of meme popularity, Phys. Rev. Lett. 112 (2014) 048701. http://dx.doi.org/10.1103/PhysRevLett.112.048701.

[25] G. Palla, I. Derényi, I. Farkas, T. Vicsek, Uncovering the overlapping community structure of complex networks in nature and society, Nature 435 (7043) (2005) 814–818. http://dx.doi.org/10.1038/nature03607.

[26] Z. Liu, B. Hu, Epidemic spreading in community networks, Europhys. Lett. 72 (2) (2005) 315. http://dx.doi.org/10.1209/epl/i2004-10550-5.

[27] J. Chen, H. Zhang, Z.-H. Guan, T. Li, Epidemic spreading on networks with overlapping community structure, Physica A 391 (4) (2012) 1848–1854. http://dx.doi.org/10.1016/j.physa.2011.10.011.

[28] X. Wu, Z. Liu, How community structure influences epidemic spread in social networks, Physica A 387 (23) (2008) 623–630. http://dx.doi.org/10.1016/j.physa.2007.09.039.

[29] L. Miaa, Q.-M. Zhang, D.-C. Nie, Shi-Min Cai, Whether information network supplements friendship network, Physica A 419 (2015) 301–306. http://dx.doi.org/10.1016/j.physa.2014.10.021.

[30] Y.-D. Zhao, S.-M. Cai, M. Tang, M.-S. Shang, A fast recommendation algorithm for social tagging systems: A delicious case. arXiv:1512.08325.