Impacts of Local Events on Communities and Diseases

Xin-Jian Xu[1], Li-Jie Zhang[1], Guo-Hong Yang[1], and Xun Zhang[2]

[1] College of Science, Shanghai University, 99 Shangda Road, Shanghai 200444, China
xinjxu@shu.edu.cn
[2] Centre for Computational Science and Engineering,
National University of Singapore, Singapore 117542

Abstract. The study of community networks has attracted considerable attention recently. In this paper, we propose an evolving community network model based on local events, the addition of new nodes intra-community and new links intra- or inter-community. Employing growth and preferential attachment mechanisms, we generate the network with a generalized power-law distribution of nodes' degrees. Furthermore, we study epidemic spreading in the resulting network by the simple SIS model to understand the influence of the network structure on the dynamics. We find that the existence of communities in networks causes the critical behavior of the spreading dynamics and keeps epidemics endemic.

Keywords: complex networks, community networks, SIS model.

1 Introduction

Complex networks, evolved from the Erdös-Rényi random graph [1], are powerful models for describing many complex systems in biology, sociology, and technology [2]. In the past decade, the explosion of the general interest in the structure and the evolution of most real-world networks is mainly reflected in two striking characteristics. One is the small-world property [3], which suggests that a network has a highly degree of clustering like regular networks and a small average distance among any two nodes similar to random networks. The small-world phenomenon has been successfully described by network models with some degree of randomness [3,4]. The other is the scale-free behavior [5], which means a power-law distribution of connectivity, $P(k) \sim k^{-\gamma}$, where $P(k)$ is the probability that a node in the network has k connections to other nodes and γ is a positive real number determined by the given network. The origin of the scale-free behavior has been traced back to two mechanisms that are observed in many systems, growing and preferential attachment [5,6].

Recently, with the progress of research in networks, many other statistical characteristics of networks appeared on the stage. Of particular renown is the so-called "community"(or "modularity"). That is to say, a network is composed of many clusters of nodes, where the nodes in the same cluster are highly connected, while there are few links among the nodes belonging to different clusters. For

J. Zhou (Ed.): Complex 2009, Part I, LNICST 4, pp. 339–350, 2009.

instance, groups are formed in scientific collaboration networks [7]. Also, it has been found that dynamical processes on networks are affected by community structures, such as tendencies spread well within communities [8] and diffusion between different communities is slow [9].

In the study of community networks, most research has been directed in two distinct directions. On the one hand, attention has been paid to designing algorithms for detecting community structures in real networks. A pioneering method was made by Girvan and Newman [7], who introduced a quantitative measure for the quality of a partition of a network into communities. Later, a number of algorithms have been proposed in order to find a good optimization with the least computational cost. The fastest available procedures use greedy techniques [10] and extremal optimization [11], which are capable of detecting communities in large networks. On the other hand, research has focused on modeling of networks with community structures. In Ref. [12], a static social network was introduced where individuals belong to groups that in turn belong to groups of groups and so on. In Ref. [13], a networked seceder model was suggested to illustrate group formation in social networks. In Ref. [14], a growing bipartite network for social communities with group structures was proposed. Each of those models is constructed based on one aspect of reality.

In this paper, we introduce a network model with communities that gives a realistic description of local events [15,16,17]. The model incorporates three processes, the addition of new nodes intra-community and new links intra- or inter-community. Using growing and preferential attachment mechanisms, we generate the community network with a good right-skewed distribution of nodes' degrees, which has been observed in many social systems. Then, we investigate the standard SIS model on the generated network. We notice a great influence of the community structure on the epidemic dynamics over such complex networks.

2 Network Model

The Barabási-Albert network [5] only describes a particular type of evolving networks, the addition of new nodes preferential connecting to the nodes already present in the network. Systems in the real world, however, are much richer. For example, in friendship networks, a person usually makes friends with people belonging to different communities besides the community he belongs to. To give a realistic description of the network construction like that, we introduce a growing model of community networks based on local events, the addition of new nodes intra-community and new links intra- or inter-community. The proposed model is defined as follows.

We start with M (≥ 2) isolated communities and each community consists of a small number n of isolated nodes. At each time step, we perform one of the following three operations.

(i) With probability p we add a new node in a randomly chosen community. Here the randomly chosen means that the community is selected according to the uniform distribution. The new node is only connected to one node that already presented in the selected community. We denote it as the uth commnuity. The probability that node i in community u will be selected is proportional to its intra-community degree

$$\prod(k_i^{\text{intra}}) = \frac{k_{u,i}^{\text{intra}} + 1}{\sum_j (k_{u,j}^{\text{intra}} + 1)}, \tag{1}$$

where the sum runs over nodes in community u and $k_{u,i}^{\text{intra}}$ is the intra-community degree of node i in community u.

(ii) With probability q we add a new link in a randomly chosen community. For this we randomly select a node in a randomly chosen community u as the starting point of the new link. The other end of the link is selected in the same community with the probability given by Eq. (1).

(iii) With probability r $(= 1 - p - q)$ we add a new link between two communities. For this we randomly select a node in a randomly chosen community u as the starting point of the new link. The other end i of the link selected in the other community v is proportional to its inter-community degree

$$\prod(k_i^{\text{inter}}) = \frac{k_{v,i}^{\text{inter}} + 1}{\sum_{v \neq u; j} (k_{v,j}^{\text{inter}} + 1)}, \tag{2}$$

where the sum runs over nodes in all communities except for community u and $k_{v,i}^{\text{inter}}$ is the inter-community degree of node i in community v.

After t time steps, this scheme generates a network of $Mn + pt$ nodes and t links. The parameters p, q, and r control the network structure. In the case of small r, the generated network will have a strong community structure. Notice that whatever process is chosen at each time step, only one link is added to the system (duplicate and self-connected edges are forbidden), however, this is not essential. We choose link probabilities $\prod(k_i^{\text{intra}})$ and $\prod(k_i^{\text{inter}})$ to be proportional to $k_i^{\text{intra}} + 1$ and $k_i^{\text{inter}} + 1$, respectively, such that there is a nonzero possibility of isolated nodes acquiring new links.

3 Degree Distribution

In our community network, the degree of a node consists of two parts, the intra-community degree and the inter-community degree. Increase in the node's connectivity can be divided into two processes, the increases of the intra-community degree and the inter-community degree. In each process, we assume that k_i^{intra} and k_i^{inter} change continuously, and the probabilities $\prod(k_i^{\text{intra}})$ and $\prod(k_i^{\text{inter}})$ can be interpreted as the rates at which k_i^{intra} and k_i^{inter} change, respectively. Thus, the operations (i)-(iii) all contribute to k_i, each being incorporated in the continuum theory as follows.

(i) Addition of a new node in a randomly chosen community with probability p:

$$\frac{\partial k_{u,i}^{\text{intra}}}{\partial t} = p \frac{1}{M} \frac{k_{u,i}^{\text{intra}} + 1}{\sum_j (k_{u,j}^{\text{intra}} + 1)}. \tag{3}$$

(ii) Addition of a new link in a randomly chosen community with probability q:

$$\frac{\partial k_{u,i}^{\text{intra}}}{\partial t} = q[\frac{1}{N} + \frac{1}{M}\frac{k_{u,i}^{\text{intra}}+1}{\sum_j (k_{u,j}^{\text{intra}}+1)}], \tag{4}$$

where N is the number of total nodes. The first term on the right-hand side (rhs) corresponds to the random selection of one end of the new link, while the second term on the rhs reflects the preferential attachment (Eq. (1)) used to select the other end of the link.

(iii) Addition of a new links between two communities with probability r :

$$\frac{\partial k_{v,i}^{\text{inter}}}{\partial t} = r[\frac{1}{N} + (1-\frac{1}{M})\frac{k_{v,i}^{\text{inter}}+1}{\sum_{v\neq u;j} (k_{v,j}^{\text{inter}}+1)}]. \tag{5}$$

The first term on the rhs represents the random selection of one end of the new link, while the second term on the rhs considers the preferential attachment (Eq. (2)) used to select the other end of the link in the other community.

Combing the contribution of above processes, we have

$$\frac{\partial k_{u,i}^{\text{intra}}}{\partial t} = \frac{p+q}{M}\frac{k_{u,i}^{\text{intra}}+1}{\sum_j (k_{u,j}^{\text{intra}}+1)} + \frac{q}{N}, \tag{6}$$

$$\frac{\partial k_{v,i}^{\text{inter}}}{\partial t} = \frac{r}{N} + r\frac{M-1}{M}\frac{k_{v,i}^{\text{inter}}+1}{\sum_{v\neq u;j} (k_{v,j}^{\text{inter}}+1)}, \tag{7}$$

with

$$\begin{aligned}
\sum_j (k_{u,j}^{\text{intra}}+1) &= \sum_j k_{u,j}^{\text{intra}} + \frac{N}{M} \\
&= 2t(p\frac{1}{M} + q\frac{1}{M}) + \frac{Mn+pt}{M} \\
&= \frac{3p+2q}{M}t + n, \\
\sum_{v\neq u;j} (k_{v,j}^{\text{inter}}+1) &= \sum_{v\neq u;j} k_{v,j}^{\text{inter}} + N(1-\frac{1}{M}) \\
&= 2tr\frac{M-1}{M} + (Mn+pt)\frac{M-1}{M} \\
&= \frac{(2-p-2q)(M-1)}{M}t + (M-1)n.
\end{aligned}$$

We can simplify Eqs. (6) and (7) for large t, and get

$$\frac{\partial k_{u,i}^{\text{intra}}}{\partial t} \approx \frac{p+q}{3p+2q}\frac{(k_{u,i}^{\text{intra}}+1)}{t} + \frac{q}{pt}, \tag{8}$$

$$\frac{\partial k_{v,i}^{\text{inter}}}{\partial t} \approx \frac{1-p-q}{2-p-2q}\frac{(k_{v,i}^{\text{inter}}+1)}{t} + \frac{1-p-q}{pt}. \tag{9}$$

The boundary conditions of the intra-community degree and the inter-community degree at initial time t_s can be estimated in the sense of mathematical expectations, $k_{u,i}^{\text{intra}}(t_s) = p + q$ and $k_{v,i}^{\text{inter}}(t_s) = r$, respectively. So we write the solutions of Eqs. (8) and (9)

$$\begin{aligned} k_{u,i}^{\text{intra}}(t) &= \frac{p^3 + p^2 + 2p^2q + 4pq + pq^2 + 2q^2}{p(p+q)}\left(\frac{t}{t_s}\right)^{\frac{p+q}{3p+2q}} \\ &\quad - \frac{p^2 + 4pq + 2q^2}{p(p+q)}, \end{aligned} \tag{10}$$

$$\begin{aligned} k_{v,i}^{\text{inter}}(t) &= \frac{2 - 2q - pq + p - p^2}{p}\left(\frac{t}{t_s}\right)^{\frac{1-p-q}{2-p-2q}} \\ &\quad - \frac{2 - 2q}{p}. \end{aligned} \tag{11}$$

In random networks, the degree distribution can be calculated by

$$P(k) = \frac{1}{t}\sum_{i=1}^{t} \delta(k_i(t) - k), \tag{12}$$

which gives

$$\begin{aligned} P(k^{\text{intra}}) &= \frac{3p^2 + 2pq}{p^2 + 2q^2 + 4pq + 2p^2q + pq^2 + p^3} \\ &\quad \times \left[\frac{p^2 + 4pq + 2q^2 + (p^2 + pq)k^{\text{intra}}}{p^2 + 2q^2 + 4pq + 2p^2q + pq^2 + p^3}\right]^{-(3 + \frac{p}{p+q})}, \end{aligned} \tag{13}$$

$$\begin{aligned} P(k^{\text{inter}}) &= \frac{2p - 2pq - p^2}{2 - p - 4q - 2p^2 + 2q^2 + 2p^2q + pq^2 + p^3} \\ &\quad \times \left[\frac{2 - 2q + pk^{\text{inter}}}{2 + p - 2q - pq - p^2}\right]^{-(3 + \frac{p}{1-p-q})}. \end{aligned} \tag{14}$$

Thus, the degree distribution of our network obeys a generalized power-law form

$$P(k) \sim [A(p,q)k + B(p,q)]^{-\gamma(p,q)}. \tag{15}$$

In Fig. 1 we present numerical results of distributions of the intra-community degree, the inter-community degree, and the total degree of nodes in log-log scale. The experimental network is generated by the proposed scheme with $N = 10^5$, $M = 10$, $n = 5$, $p = 0.4$, and $q = 0.4$, respectively. The distributions of the intra-community degree and the inter-community degree, shown in Figs. 1(a) and 1(b), agree with analytical results of Eqs. (13) and (14), respectively. The small deviations between computer simulations and analytical solutions at both ends of the distributions appears to be the mathematical approximation of the boundary conditions and the finite size effect due to the relatively small network size used in the simulations. According to the evolving rule of our network, nodes with larger intra- (or inter-) degree have higher probabilities to gain new

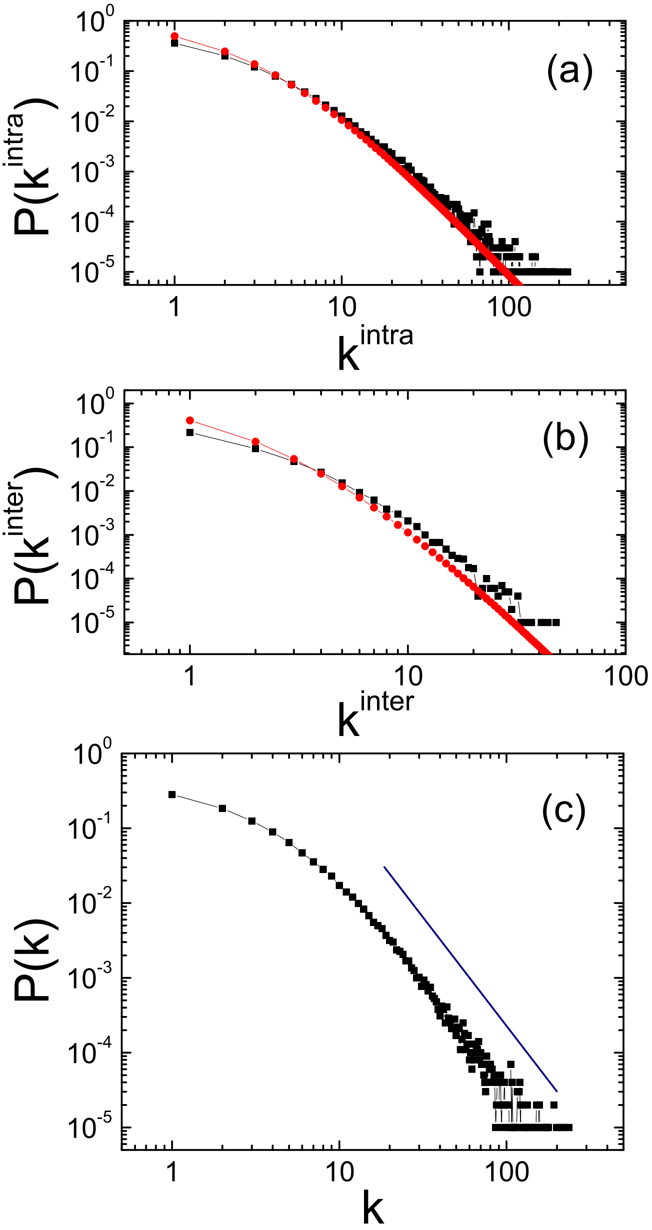

Fig. 1. (Color online) Log-log representation of distributions of intra-community degree (a), inter-community degree (b), and total degree (c) of nodes. All the simulation results (squares) display good right-skewed distributions. The circles in (a) and (b) denote analytical results predicted by Eqs. (13) and (14), respectively. The solid line in (c) is guide to the eye with power-law decay exponent $\gamma = 3$. The experiment network has a total number of nodes $N = 10^5$ with parameters $M = 10$, $n = 5$, $p = 0.4$, and $q = 0.4$, respectively.

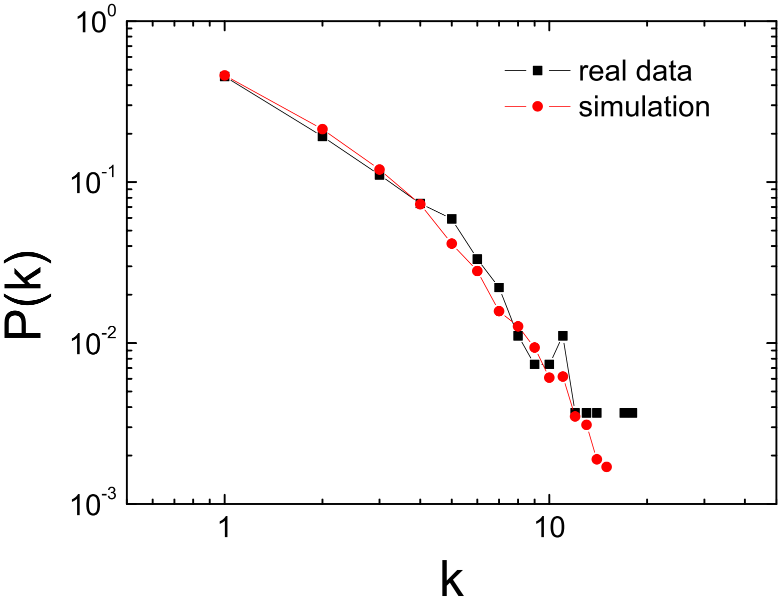

Fig. 2. (Color online) The degree distribution of econophysicists (squares) of an econophysics scientific collaboration network [19]. The circles correspond to computer simulations of our model with parameters $M = 10$, $n = 2$, $p = 0.4$, and $q = 0.4$, respectively.

links, then the usual degree preferential attachment is reasonably kept. This means that the right-skewed character of the network, such as the node's total degree, will retain. As shown in Fig. 1(c), the total degree distribution of nodes is well expected showing a good right-skewed character, which is reasonably in agreement with the condition of many realistic systems [18].

To illustrate the predictive power, we also compare the numerical result of our network with the statistics of an econophysics collaboration network. In the econophysics collaboration network, each node represents one scientist. If two scientists have collaborated one or more papers, they would be connected by an edge. Zhang *et al.* took the largest connected component of this network, which includes 271 nodes and 371 edges, and provided the best division, i.e., $M = 10$ [19]. In Fig. 2 we plot the degree distribution of econophysicists of the econophysics collaboration network which is fitted by computer simulations of our network starting with 10 communities. To gain p and q, we fit the connectivity distribution $P(k)$ obtained from this collaboration network with Eq. (15), obtaining a good overlap for $p = 0.75$ and $q = 0.15$ (Fig. 2).

4 Epidemic Spreading

In the consequent study of complex networks, an important topic is to inspect the effect of their topologies on dynamical behaviors and evolutionary processes. Of particular importance is the spread of infectious diseases, which has attracted increasing attention from scientific communities in both theoretical and experimental investigations [20]. It is easy to foresee that a good understanding and accurate characterization of epidemic dynamics over complex networks can provide immediate benefits to a large number of practical problems such as computer virus propagation and prevention, cascading failures of power grids, spreading of rumors and public opinions, etc. Motivated by this observation, some fundamental works have been done regarding an unified mathematical theory of disease spreading over complex networks [21,22,23], which have triggered a large number of following works to study epidemic models on different networks [24,25,26,27] as well as corresponding immunization strategies [28,29,30,31].

There are several classical epidemiological models taking into account different characteristics of disease transmission. A typical one is the SIS model. This model is defined on a network, where nodes represent individuals or groups of individuals and edges represent social contacts or relations among them. In the SIS model, an individual is described by a single dynamical variable having one of the two states: susceptible and infected. A susceptible individual at time t will be changed to the infected state at time $t+1$, with the transmission rate $\nu > 0$, if it is connected to an infected individuals. An infected individual at time t will be changed back to the susceptible state again at time $t+1$, with the transmission rate $\delta > 0$, due to immunity or medical treatment. Here, an effective spreading rate is defined to be $\lambda = \nu/\delta$ and, without loss of generality, one may assume $\delta = 1$ since it only affects the definition of the time scale of epidemic propagation.

In present work, we have performed Monte-Carlo (MC) simulations of the SIS model with synchronously updating after the network has been constructed. The two courses, the evolution of the network and the epidemic dynamics, are independent. Initially, the number of randomly infected nodes is 1 percent of the size of the network. After appropriate relaxation times, the systems stabilize in a steady state. Simulations were averaged over 10 different starting configurations, performed on 10 different realizations of the network. Given a network, an important observable is the prevalence ρ, which is the time average of the fraction of infected individuals in steady states.

Figures 3(a) and 3(b) depict the steady density of infected nodes ρ vs the spreading rate λ and its reciprocal, respectively. Four epidemic plots are presented in Fig. 3(a). Here we choose $p = q$ in simulations, just as an illustration of the influence of the local events on the dynamics. One can easily notice a drop of the epidemic prevalence as p and q become larger. In addition, the critical behavior of the dynamics is implied by Fig. 3(b), namely, there is a transition from the epidemic state to the endemic one as the network gets more modular. This is different from the results obtained in Balabási-Albert scale-free networks, although the connectivity of our community network displays a similar power-law decay. Only in cases that the addition of links among groups are equal to or larger than that inside groups, i.e., the networks are reduced to typical growing scale-free networks, can the disease show the feature of the absence of the threshold. To describe the critical behavior in detail, we plot the epidemic threshold

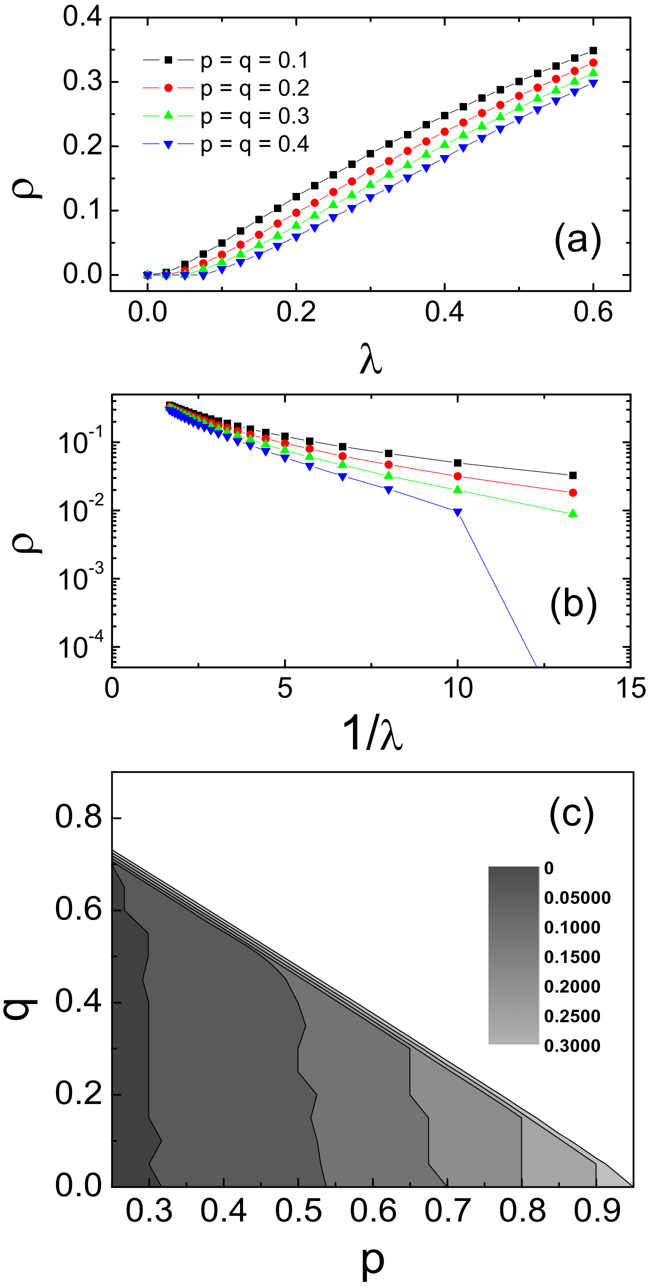

Fig. 3. (Color online) Density of infected nodes ρ vs λ (a) and $1/\lambda$ (b) for different local events. The trend of the network modularity enhances the critical value of the spreading rate. (c) The epidemic threshold λc in the p? q plane. Parameters of the underlying networks are the same as in Fig. 1.

λ_c in the $p - q$ plane in Fig. 3(c). We observe the absence of the threshold in regions about $p \leq 0.35$ and $q \leq 0.7$. While the addition of links intra-community is larger than that inter-community (in the regions of large p or q), we find that $\lambda_c > 0$. It can be easily understood. For community networks with large p or q, the infection will be confined within the groups where the initially infected seeds are chosen because the number of links among groups is much smaller than that inside the groups. Thus the existence of communities in networks will suppress spreading.

5 Conclusions

Networks with community structures underlie many natural and artificial systems. It is becoming essential to model and study this kind topological feature. We presented a simplified mechanism for networks organized in communities, which corresponds to local events during the system growth. The generated network is highly clustered and has a good right-skewed distribution of connectivity, which have been found very common in most realistic systems. Furthermore, we investigated the spreading process of infectious diseases in the community network by using the SIS model and observed the containment of modular structures to epidemic propagation. That is to say, the network is robust to diseases when connections inside communities has an overwhelming majority than that among communities. This might contribute to understanding realistic epidemics with critical behavior even the underline structures are scale free.

The present paper only suggests a simple way for generating community networks. The shape of the resulting network is deterministic in some extent. It is more interesting to model the evolution of communities, especially the self organization (or emergence) of communities in the natural world [32], which is left to future work.

Acknowledgments

The authors acknowledge financial support from NSFC under Grant No. 10805033 and STCSM under Grant No. 08ZR1408000. This work is sponsored by Shanghai Leading Academic Discipline Project.

References

1. Erdös, P., Rényi, A.: On the Evolution of Random Graphs. Publ. Math. Inst. Hung. Acda. Sci. 5, 17–65 (1960)
2. Albert, R., Barabási, A.-L.: Statistical Mechanics of Complex Networks. Rev. Mod. Phys. 74, 47–97 (2002)
3. Watts, D.J., Strogatz, S.H.: Collective Dynamics of “Small-world”Networks. Nature 393, 440–442 (1998)
4. Newman, M.E.J., Watts, D.J.: Scaling and Percolation in the Small-world Network Model. Phys. Rev. E 60, 7332–7342 (1999)

5. Barabási, A.-L., Albert, R.: Emergence of Scaling in Random Networks. Science 286, 509–512 (1999)
6. Krapivsky, P.L., Redner, S., Leyvraz, F.: Connectivity of Growing Random Networks. Phys. Rev. Lett. 85, 4629–4632 (2000)
7. Girvan, M., Newman, M.E.J.: Community Structure in Social and Biological Networks. Proc. Natl. Acad. Sci. USA 99, 7821–7826 (2002)
8. Bettencourt, L.M.A.: Tipping the Balances of a Small World. arXiv:cond-mat/0304321
9. Eriksen, K.A., Simonsen, I., Maslov, S., Sneppen, K.: Modularity and Extreme Edges of the Internet. Phys. Rev. Lett. 90, 148701 (2003)
10. Newman, M.E.J.: Fast Algorithm for Detecting Community Structure in Networks. Phys. Rev. E 69, 066133 (2004)
11. Duch, J., Arenas, A.: Community Detection in Complex Networks Using Extremal Optimization. Phys. Rev. E 72, 027104 (2005)
12. Watts, D.J., Dodds, P.S., Newman, M.E.J.: Identity and Search in Social Networks. Science 296, 1302–1305 (2002)
13. Grönlund, A., Holme, P.: Networking the Seceder Model: Group Formation in Social and Economic Systems. Phys. Rev. E 70, 036108 (2004)
14. Noh, J.D., Jeong, H.-C., Ahn, Y.-Y., Jeong, H.: Growing Network Model for Community with Group Structure. Phys. Rev. E 71, 036131 (2005)
15. Albert, R., Barabási, A.-L.: Topology of Evolving Networks: Local Events and Universality. Phys. Rev. Lett. 85, 5234 (2000)
16. Dorogovtsev, S.N., Mendes, J.F.F.: Scaling Behaviour of Developing and Decaying Networks. Europhys. Lett. 52, 33–39 (2000)
17. Li, X., Chen, G.: A Local-world Evolving Network Model. Physica A 328, 274–286 (2003)
18. Amaral, L.A.N., Scala, A., Barthelemy, M., Stanley, H.E.: Classes of Small-world Networks. Proc. Natl. Acad. Sci. USA 97, 11149–11152 (2000)
19. Zhang, P., Li, M., Wu, J., Di, Z., Fan, Y.: The analysis and Dissimilarity Comparison of Community Structure. Physica A 367, 577–585 (2006)
20. Biley, N.T.J.: The Mathematical Theory of Infectious Diseases and its Applications, 2nd edn. Griffin, London (1975)
21. Pastor-Satorras, R., Vespignani, A.: Epidemic Spreading in Scale-Free Networks. Phys. Rev. Lett. 86, 3200–3203 (2001)
22. Newman, M.E.J.: Spread of Epidemic Disease on Networks. Phys. Rev. E 66, 016128 (2002)
23. Warren, C.P., Sander, L.M., Sokolov, I.M.: Epidemics, Disorder, and Percolation. Physica A 325, 1–8 (2003)
24. Eguíluz, V.M., Klemm, K.: Epidemic Threshold in Structured Scale-free Networks. Phys. Rev. Lett. 89, 108701 (2002)
25. Barthélemy, M., Barrat, A., Pastor-Satorras, R., Vespignani, A.: Dynamical Patterns of Epidemic Outbreaks in Complex Heterogeneous Networks. J. Theor. Biol. 235, 275–288 (2005)
26. Vazquez, A.: Polynomial Growth in Branching Processes with Diverging Reproductive Number. Phys. Rev. Lett. 96, 038702 (2006)
27. Volz, E.: SIR Dynamics in Random Networks with Heterogeneous Connectivity. J. Math. Biol. 56, 293–310 (2008)
28. Pastor-Satorras, R., Vespignani, A.: Immunization of Complex Networks. Phys. Rev. E 65, 036104 (2002)

29. Cohen, R., Havlin, S., ben-Avraham, D.: Efficient Immunization Strategies for Computer Networks and Populations. Phys. Rev. Lett. 91, 247901 (2003)
30. Hufnagel, L., Brockmann, D., Geisel, T.: Forecast and Control of Epidemics in a Globalized World. Proc. Natl. Acad. Sci. USA 101, 15124–15129 (2004)
31. Gross, T., Dommar D'Lima, C.J., Blasius, B.: Epidemic Dynamics on an Adaptive Network. Phys. Rev. Lett. 96, 208701 (2006)
32. Kumpula, J.M., Onnela, J.-P., Saramäki, J., Kaski, K., Kertész, J.: Emergence of Communities in Weighted Networks. Phys. Rev. Lett. 99, 228701 (2007)