Small world yields the most effective information spreading

Linyuan Lü[1,2], Duan-Bing Chen[1], and Tao Zhou[1*]
[1] *Web Sciences Center, University of Electronic Science and Technology of China, 610054 Chengdu, People's Republic of China*
[2] *Department of Physics, University of Fribourg, Chemin du Musée 3, CH-1700 Fribourg, Switzerland*

Spreading dynamics of information and diseases are usually analyzed by using a unified framework and analogous models. In this paper, we propose a model to emphasize the essential difference between information spreading and epidemic spreading, where the memory effects, the social reinforcement and the non-redundancy of contacts are taken into account. Under certain conditions, the information spreads faster and broader in regular networks than in random networks, which to some extent supports the recent experimental observation of spreading in online society [D. Centola, Science **329**, 1194 (2010)]. At the same time, simulation result indicates that the random networks tend to be favorable for effective spreading when the network size increases. This challenges the validity of the above-mentioned experiment for large-scale systems. More significantly, we show that the spreading effectiveness can be sharply enhanced by introducing a little randomness into the regular structure, namely the small-world networks yield the most effective information spreading. Our work provides insights to the understanding of the role of local clustering in information spreading.

PACS numbers: 89.75.Hc, 89.75.Fb, 05.70.Ln, 02.50.-r

I. INTRODUCTION

Understanding the dynamics of epidemic spreading is a long-term challenge, and has attracted increasing attention recently. Firstly, the fast development of data base technology and computational power makes more data available and analysable to scientific community. Secondly, many new objects of study come into the horizon of epidemiologists, such as computer virus, opinions, rumors, behaviors, innovations, fads, and so on. Lastly, in addition to the compartment model and population dynamics [1], novel models and tools appeared recently inspired by the empirical discoveries about network topology [2, 3], temporal regularities of human activities [4–6] and scaling laws in human mobility [7, 8].

In the simplest way, we can roughly divide the human-activated spreading dynamics into two classes according to the disseminules: one is the spreading of infectious diseases requiring physical contacts, and the other is the spreading of information including opinions, rumors and so on (Here we mainly consider the information whose value and authenticity need judge and verification by individuals, different from the information about jobs, discounts, etc.). In the early stage, scientists tried to describe these two classes by using a unified framework and analogous models (see, e.g., Ref. [9, 10]), emphasizing their homology yet overlooking their essential differences. Very recently, scientists started to take serious consideration about the specific features of information spreading [11, 12], as well as the different mechanisms across different kinds of information [13]. Dodds and Watts [14] studied the effects of limited memory on contagion, yet did not consider the social reinforcement. Some recent works indicate that the social reinforcement plays

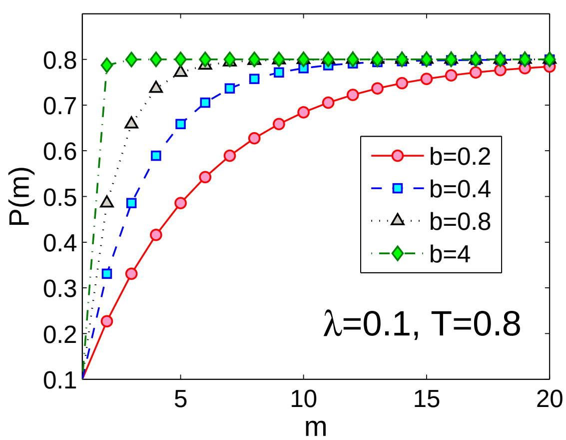

FIG. 1: (Color online) The approving probability as a function of m.

important role in the propagation of opinions, news, innovations and fads [15–19].

In this paper, we propose a variant of the susceptible-infected-recovered (SIR) model for information spreading, which takes into account three different spreading rules from the standard SIR model: (i) memory effects, (ii) social reinforcement, and (iii) non-redundancy of contacts. The main contributions are twofold. Firstly, we show that when the spreading rate λ is smaller than a certain value λ^*, the information spreads more effectively in regular networks than in random networks, which to some extent supports the experiment reported by Centola [20]: behavior spreads faster and can infects more people in a regular online social network than in a random one (with no more than 200 people in the experiment). We further show that as the increasing of the network size, the value of λ^* will decrease, which challenges the validity

*Electronic address: zhutou@ustc.edu

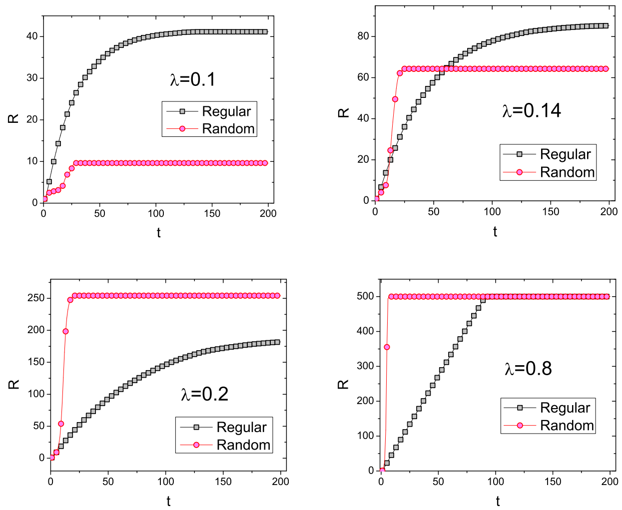

FIG. 2: (Color online) The number of approved nodes as a function of time on regular network (black squares) and random network (red circles). The parameters are $N = 500$, $k = 6$, $b = 0.8$ and $T = 1$. The results are obtained by averaging over 500 independent realizations.

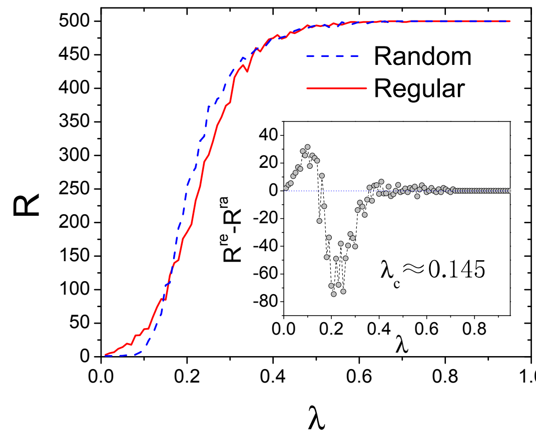

FIG. 3: (Color online) The dependence of the number of approved nodes at the final state on the parameter λ for regular (red solid line) and random (blue dash line) networks. The parameters are $N = 500$, $k = 6$, $b = 0.8$ and $T = 1$, as the same as those for Fig. 2. Inset shows the number of final approved nodes on regular network $R^{\rm re}$ minus that on random networks $R^{\rm ra}$, against λ. The results are obtained by averaging over 500 independent realizations.

of Centola's experiment [20] for very large-scale networks. Secondly, the effectiveness of information spreading can be remarkably enhanced by introducing a little randomness into the regular structure, namely the small-world networks [21] yield the most effective information spreading. This result is complementary to the traditional understanding of epidemic spreading on networks where the infectious diseases spread faster in random networks than in small-world networks.

II. MODEL

Given a network with N nodes and E links representing the individuals and their interactions, respectively. Hereinafter, for convenience, we use the language of news spreading, but our model can be applied to the spreading of many kinds of information like rumors and opinions, not limited to news. At each time step, each individual adopts one of four states: (i) *Unknown*–the individual has not yet heard the news, analogous to the susceptible state of the SIR model. (ii) *Known*–the individual is aware of the news but not willing to transmit it, because she is suspicious of the authenticity of the news. (iii) *Approved*–the individual approves the news and then transmits it to all her neighbors. (iv) *Exhausted*–after transmitting the news, the individual will lose interest and never transmit this news again, analogous to the recovered state in the SIR model.

At the beginning, one node is randomly chosen as the "seed" and all others are in the unknown state. This seed node will transmit the news to all her neighbors, and then become exhausted. Once an individual (either in unknown or known state) receives a news, she will judge whether it is true depending on the number of times he has heard it—a news or a rumor is more likely to be approved if being heard many times (a very recent model allows the infectivity and/or susceptibility of hosts to be dependent on the number of infected neighbors [22]). The present rules imply two features of information spreading, namely the memory effects and social reinforcement, which are usually neglected in the standard SIR model and its variants for rumor propagation.

In our model, we assume that for a given individual if she receives the news at least once at the tth time step, and she has received $m(t)$ times of this news until time t ($m(t)$ is a cumulative number), the probability she will approve it at time t is $P(m) = (\lambda - T)e^{-b(m-1)} + T$, where $\lambda = P(1)$ is the approving probability for the first receival. $T \in (0, 1]$ is the upper bound of the probability indicating the maximal approving probability. Here we do not consider the interest decay, and we assume that the time scale of news spreading is much faster than our memory decay. After approval, she will transmit the news to all her neighbors in the next time step and then turn to be exhausted. If an individual, either in unknown or known states, does not receive any news in the tth time step, nothing will happen no matter how many times this individual has received the news. The memory effects are embodied by $m(t)$ which is a cumulative number instead of the independent spreading rates for different contacts in the standard SIR model. With the increasing of m,

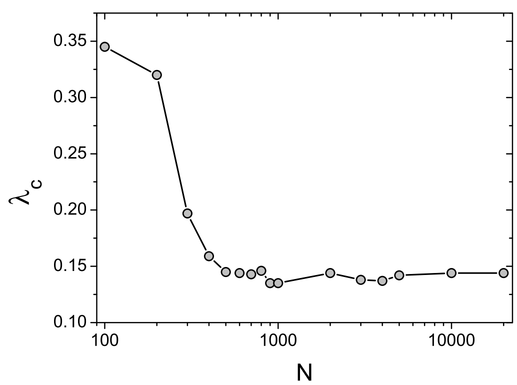

FIG. 4: The dependence of λ_c on the network size N. The parameters are $k = 6$, $b = 0.8$ and $T = 1$. The results are obtained by averaging over 500 independent realizations.

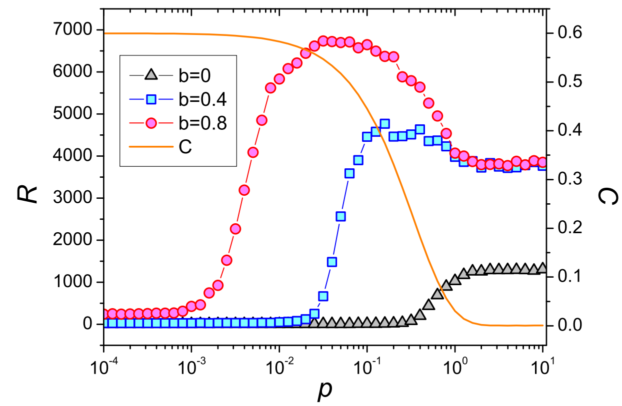

FIG. 5: (Color online) The number of final approved nodes against the randomness p given $b = 0$ (triangles), $b = 0.4$ (squares) and $b = 0.8$ (circles). Other parameters are $N = 10^4$, $k = 6$, $\lambda = 0.2$ and $T = 1$. The results are obtained by averaging over 10000 independent realizations. The clustering coefficient C, as a monotonic function of p, is also displayed.

$P(m)$ will infinitely approach to T and the speed is determined by the parameter $b > 0$ which reflects the social reinforcement effect. Figure 1 shows the approving probability as a function of m, given different b. Larger b indicates a stronger social reinforcement. For example, $P(2) = 0.227$ when $b = 0.2$, and it equals 0.486 for $b = 0.8$. Since an individual who has transmitted the news will immediately become exhausted, our model ensures that each link is used at most once without any redundancy of contacts. The spreading process comes to the end when no new individual approves the news and spreads it.

We perform our model on three kinds of networks with identical node degree k. (i) *Regular networks.*–A regular network is a one dimensional ordered network with periodic boundary conditions, where each node is connected to its k nearest neighbors, namely to the $k/2$ nearest neighbors clockwise and counterclockwise [21]. Notice that, in the literature of graph theory, the term "regular networks" usually stands for networks whose nodes are of the same degree, and thus the following homogeneous small-world networks are also regular. In this article, we follow the literature of complex networks and use the term "regular networks" to represent networks with ordered structure. (ii) *Homogeneous small-world networks.*–The homogenous small-world network is constructed by randomly reshuffling links of a regular network, while keeping the degree of each node unchanged [23]. According to the link exchanging method [24], at each time step, we randomly select a pair of edges A-B and C-D. These two edges are then rewired to be A-D and B-C. To prevent the multiple edges connecting the same pair of nodes, if A-D or B-C already exists in the network this step is aborted and a new pair of edges is randomly selected. We implement pE steps, where p indicates the randomness of the network. (iii) *Random networks.*–Repeating the above rewiring operations many times leads to a homogenous random network. Theoretically speaking, a homogenous random network is obtained only for $p \to \infty$, we here consider $p \in [0, 10]$ and when $p > 1$, the topological statistics are very close to the ones of random networks. In all simulations, the node degree is set to be $k = 6$, and we have carefully checked that the results are not sensitive to the node degree unless k is very large or very small.

III. RESULTS

Denote by R the number of approved nodes of the news. Larger R at the final state indicates a broader spreading. We firstly compare the spreading processes on regular and random networks. Figure 2 reports four typical examples with different λ and fixed $b = 0.8$. Surprisingly, for small λ (e.g., Fig. 2(a)), the spreading on regular networks is faster and broader than on random networks. These results are in accordance with the online social experiment of Centola [20], yet against the traditional understanding of network spreading [28]. With the increasing of λ, the random networks will be favorable for faster and broader spreading. Figure 3 shows the dependence of the number of approved nodes at the final state on the parameter λ. There is a crossing point at about $\lambda_c \approx 0.145$, after which R of random networks exceeds that of regular networks. The inset shows the difference between numbers of final approved nodes on regular and random networks, namely $R^{\mathrm{re}} - R^{\mathrm{ra}}$ against λ. With very large λ, almost every node will run into the approved state, and thus R is not sensitive to the

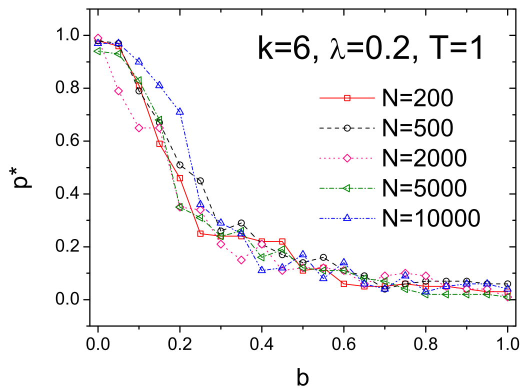

FIG. 6: (Color online) The dependence of optimal randomness p^* on the strength of social reinforcement b given different N. The results are obtained by averaging over 500 independent realizations.

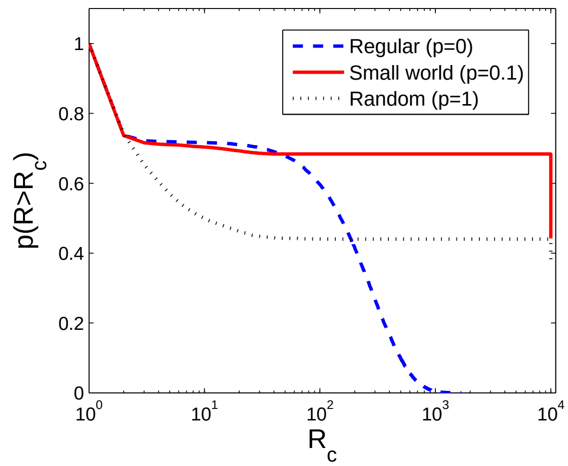

FIG. 7: (Color online) The cumulative probability that in a realization, the information reaches more than R_c individuals for regular, random and small-world networks. The parameters are $\lambda = 0.2$, $k = 6$, $T = 1$ and $N = 10000$. These distributions are obtained from 10000 realizations.

network structure, but the spread on random networks is still faster than on regular networks (see, for example, Fig. 2(d)).

Figure 4 displays the crossing point λ_c as a function of the network size N. When N is small, λ_c decreases sharply with the increasing of N, while when N gets larger λ_c becomes insensitive to N. As a whole, λ_c shows a non-increasing behavior versus N. Notice that, the phenomenon that spreading on regular networks is faster and broader than on random networks is more remarkable and easier to be observed if λ_c is large. Therefore, our result about $\lambda_c(N)$ indicates that for large-scale systems, Centola's experimental results may be not hold or will be weaken to some extent.

In previous study on SIR model, it was pointed out that the number of recovered nodes at the end of evolution increases with the increasing of randomness p in small-world networks [29]. In contrast, our simulations show that the number of approved nodes in the final state does not monotonously increase with the increasing of p, instead, an optimal randomness p^* exists subject to the highest R. Figure 5 shows the dependence of the number of final approved nodes on the randomness p given $b = 0$ (triangles), $b = 0.4$ (squares) and $b = 0.8$ (circles). With strong social reinforcement, even a very small randomness can bring a remarkable improvement of the number of final approved nodes, R. Take the case $b = 0.8$ for example, on the regular networks (i.e., $p = 0$), R is 205, while by introducing a tiny randomness $p = 0.02$, this number will suddenly increase to 6593, which is also higher than the random networks (i.e., $p = 1$, $R = 4049$). We also plot the clustering coefficient C as a function of p in figure 5. As expected, C decreases as the increasing of p. The results indicate that the local clustering can to some extent enhance the approving rate of information, which refine the completely negative valuation of clustering coefficient in epidemic spreading [25–27].

The dependence of optimal randomness p^* on the strength of social reinforcement b given different N are shown in Fig. 6, where one can observe that the stronger social reinforcement (i.e., larger b) results in a smaller p^*. In the presence of weak social reinforcement (i.e., small b), our result (p^* is close to 1) is analogous to the well-known one [28, 29] that the speed and range of spreading obey the relation "Random > Small-World > Regular". In contrast, the small-world networks yield the most effective spreading when the social reinforcement plays a considerable role (i.e., large b).

To further investigate the advantages of small-world networks for information spreading, we calculate the complementary cumulative distribution $p(R > R_c)$, namely the probability that in a realization the information has reached more than R_c individuals. As shown in figure 7, comparing with random networks, the advantages of small-world networks are twofold. On one hand, it has higher probability to spread out (see the region when R_c is small). For example, in small-world networks $p(R > 10) = 0.703$, while for random networks, this number is only 0.460. If the information can spread out, like an epidemic for a disease, in both two kinds of networks it can reach the majority of population. In contrast, comparing with regular networks, information in small-world networks can spread wide. According to figure 7, the maximum R in regular networks is only 1680 while in small-world networks it can reach 9900 individuals with probability 0.684.

IV. CONCLUSION AND DISCUSSION

Thanks to the fast development of data base technology and computational power, the detailed analysis about information spreading in large-scale online systems become feasible nowadays. In our opinion, the similarity between information spreading and epidemic spreading are over emphasized in the previous studies (see, for example, the models summarized in the review article [30]), and currently we should turn to the other side of the matter: revealing the essential difference between them. The significant difference may include: (i) *Time decaying effects.*– An infectious disease can exist more than thousands of years in human society and still keep active, but no one is willing to spread a news one year ago. Actually, our attention on information decays very fast [31], and thus when we model the information spreading, especially if it involves multiple information competing for attention, we have to consider the time decaying effects. (ii) *Tie strength.*– It is well known that in social networks, ties with different strengths play different roles in maintaining the network connectivity [32], information filtering [33], information spreading [34], and so on. We guess the weak ties provide faster paths for information spreading while the strong ties provide trust paths (i.e., with high infectivity). However, this point is still not clear till far. (iii) *Information Content.*– Information with different contents may have far different spreading paths, and even with the same content, different expressions may lead to far different performances. Some of them are born with fashionable features while others are doomed to be kept from known. Whether these two kinds of information are only different quantitatively or they follow qualitatively different dynamic patterns are still under investigation [35]. (iv) *Role of spreaders.*– Recent analysis on Twitter show that different kinds of spreaders, such as media, celebrities, bloggers and formal organizations, play remarkably different roles in network construction and information spreading [36], which may result in different spreading pathes and outbreaking mechanisms from epidemic spreading. (v) *Memory effects.*– Previous contacts could impact the information spreading in current time [14]. Such memory effects can be direct since an agent may tend to be interested in our disgusted with objects heard many times, and/or indirect since previous contacts could change the tie strength that further impact the current interactions. (vi) *Social reinforcement.*– If more than one neighbor approved the information and transferred it to you, you are of high probability to approve it. Generally speaking, is an agent receives twice an information item recommended from her neighbors, the approval probability should be much larger than the twice of the approval probability with a single recommending. (vii) *Non-redundancy of contacts.*– People usually do not transfer an information item more than once to the same guy, which is far different from the sexually transmitted diseases. To name just a few.

In this paper, we propose a simple model for information spreading in social networks that considers the memory effects, the social reinforcement and the non-redundancy of contacts. Under certain conditions, the information spreads faster and broader in regular networks than in random networks, which to some extent supports the Centola's experiment [20]. At the same time, we show that the random networks tend to be favorable for effective spreading when the network size increases, which challenges the validity of the Centola's experiment for large-scale systems. Furthermore, simulation results suggest that by introducing a little randomness into regular structure, the small-world networks yield the most effective information spreading. Although this simple model cannot take into account all the above-mentioned features in information spreading, it largely refines our understanding about spreading dynamics. For example, traditional spreading models on complex networks show that the diseases spread faster and broader in random networks than small-world networks [28, 29], yet our results suggest that the small world may be the best structure for effective spreading under the consideration of social reinforcement. Indeed, information in small-world networks has much higher probability to spread out than in random networks, and can spread much broader than in regular networks. In addition, the local clustering is well-known to play a negative role in spreading [25–27], while our model indicates that local clustering are very helpful in facilitating the acceptance/approval of the information for individuals and thus can to some extent fasten the spreading.

Acknowledgments

We acknowledge Xiao-Ke Xu for valuable discussion. This work is supported by the National Natural Science Foundation of China under Grant No. 90924011, the Fundamental Research Funds for the Central Universities, and the Swiss National Science Foundation under Grant No. 200020-132253.

[1] R. L. May and R. M. Anderson, Infectious diseases of humans: dynamics and control (Oxfrod: Oxford University Press, 1991).
[2] R. Pastor-Satorras and A. Vespignani, Phys. Rev. Lett. **86**, 3200 (2001).
[3] A. Barrat, M. Barthelemy, and A. Vespignani, Dynamical Processes on Complex Networks (Cambridge: Cambridge University Press, 2008).
[4] A. Vazquez, B. Rácz, A. Lukács, and A.-L. Barabási, Phys. Rev. Lett. **98**, 158702 (2007).

[5] J. L. Iribarren and E. Moro, Phys. Rev. Lett. **103**, 038702 (2009).
[6] Z. Yang, A.-X. Cui, and T. Zhou, Physica A **390**, 4543 (2011).
[7] P. Wang, M. C. González, C. A. Hidalgo, and A.-L. Barabási, Science **324**, 1071 (2009).
[8] D. Balcan, V. Colizza, B. Gonçalves, H. Hu, J. J. Ramasco, and A. Vespignani, Proc. Natl. Acad. Sci. U.S.A. **106**, 21484 (2009).
[9] Z. Liu, Y.-C. Lai, and N. Ye, Phys. Rev. E **67**, 031911 (2003).
[10] Y. Moreno, M. Nekovee, and A. F. Pacheco, Phys. Rev. E **69**, 066130 (2004).
[11] A. L. Hill, D. G. Rand, M. A. Nowak, and N. A. Christakis, PLoS Comput. Biol. **6**, e1000968 (2010).
[12] T. House, J. R. Soc. Interface **8**, 909 (2011).
[13] D. M. Romero, B. Meeder, and J. Kleinberg, Proc. 20th Intl. Conf. WWW (New York: ACM Press, 2011).
[14] P. S. Dodds and D. J. Watts, Phys. Rev. Lett. **92**, 218701 (2004).
[15] D. Centola, V. M. Eguíluz, and M. W. Macy, Physica A **374**, 449 (2007).
[16] M. Medo, Y.-C. Zhang, and T. Zhou, EPL **88**, 38005 (2009).
[17] G. Cimini, M. Medo, T. Zhou, D. Wei, and Y.-C. Zhang, Eur. Phys. J. B **80**, 201 (2011).
[18] D. Wei, T. Zhou, G. Cimini, P. Wu, W. Liu, and Y.-C. Zhang, Physica A **390**, 2117 (2011).
[19] P. L. Prapivsky, S. Redner, and D. Volovik, arXiv: 1104.4107.
[20] D. Centola, Science **329**, 1194 (2010).
[21] D. J. Watts and S. H. Strogatz, Nature (London) **393**, 440 (1998).
[22] F. J. Pérez-Reche, J. J. Ludlam, S. N. Taraskin, and C. A. Gilligan, Phys. Rev. Lett. **106**, 218701 (2011).
[23] F. C. Santos, J. F. Rodrigues, and J. M. Pacheco, Phys. Rev. E **72**, 056128 (2005).
[24] S. Maslov and K. Sneppen, Science **296**, 910 (2002).
[25] V. M. Eguíluz and K. Klemm, Phys. Rev. Lett. **89**, 108701 (2002).
[26] T. Petermann and P. De Los Rios, Phys. Rev. E **69**, 066116 (2004).
[27] T. Zhou, G. Yan, and B.-H. Wang, Phys. Rev. E **71**, 046141 (2005).
[28] T. Zhou, Z.-Q. Fu, and B.-H. Wang, Prog. Nat. Sci. **16**, 452 (2006).
[29] D. H. Zanette, Phys. Rev. E **64**, 050901(R) (2001).
[30] C. Castellan, S. Fortunato, and V. Loreto, Rev. Mod. Phys. **81**, 591 (2010).
[31] F. Wu and B. A. Huberman, Proc. Natl. Acad. Sci. U.S.A. **104**, 17599 (2007).
[32] J.-P. Onnela, J.Saramäki, J. Hyvönen, G. Szabó, D. Lazer, K. Kaski, J. Kerész, and A.-L. Barabási, Proc. Natl. Acad. Sci. U.S.A. **104**, 7332 (2007).
[33] L. Lü and T. Zhou, EPL **89**, 18001 (2010).
[34] G. Miritello, E. Moro, and R. Lara, Phys. Rev. E **83**, 045102(R) (2011).
[35] R. Crane and D. Sornette, Proc. Natl. Acad. Sci. U.S.A. **105**, 15649 (2008).
[36] S. Wu, J. M. Hofman, W. A. Mason, and D. J. Watts, Proc. 20th Intl. Conf. WWW (New York: ACM Press, 2011).