

Emergence of Community Structure in the Adaptive Social Networks

Long Guo* and Xu Cai

Complexity Science Center and Institute of Particle Physics, Huazhong (Central China) Normal University, Wuhan, 430079, China.

Received 18 October 2009; Accepted (in revised version) 16 December 2009

Communicated by Dietrich Stauffer

Available online 17 May 2010

Abstract. In this paper, we propose a simple model of opinion dynamics to construct social networks, based on the algorithm of link rewiring of local attachment (RLA) and global attachment (RGA). Generality, the system does reach a steady state where all individuals' opinion and the complex network structure are fixed. The RGA enhances the ability of consensus of opinion formation. Furthermore, by tuning a model parameter p , which governs the proportion of RLA and RGA, we find the formation of hierarchical structure in the social networks for $p > p_c$. Here, p_c is related to the complex network size N and the minimal coordination number $2K$. The model also reproduces many features of large social networks, including the "weak links" property.

PACS: 87.23.Ge, 89.75.Fb

Key words: Opinion dynamics, social network, community structure, weak links property.

1 Introduction

Networks have in recent years undergone a remarkable development and have emerged as an invaluable tool for describing and quantifying complex systems in many branches of science [1]. It has been realized that many real complex networks, including social networks such as peer-to-peer social networks [2] and acquaintance networks [3], the technology networks such as the power grids [3,4], and biological networks such as the food webs [5] and metabolic networks [6], all share some distinctive characteristic properties. One such property is the "small-world effect", which means that the average shortest path length between vertices in network is short, usually scaling logarithmically with the size N of network. Another is the hierarchical structure, which means that vertices

*Corresponding author. *Email addresses:* longkuo0314@gmail.com (L. Guo), xcai@mail.ccnu.edu.cn (X. Cai)

are divided into groups that further subdivided into groups of groups, and so forth over multiple scales, in network [1]. Examples include the ecological niches in food webs, the modules in biochemical networks and communities in social networks [1, 7–9].

The community structure, the gathering of vertices into groups such that there is a higher density of edges within groups than between them [9], is a crucial property of social network and draws many scientists' attention to research how to detect and identify the communities within networks. Recently, a number of approaches, such as the divisive algorithm based on the edge betweenness [8, 10], the modularity algorithm [11], the self-contained algorithm [12] and a physical approach based on notions of voltage drops across networks [13], have been proposed. Another important problem is the mechanism of emergence of the hierarchical structure, especially the communities in social networks. Clauset et al. proposed a hierarchical random graph to study the hierarchical structure and predict the missing links in networks [1]. González et al. proposed a model of mobile agents to construct social networks and found the emergence of a giant cluster in the universality class of two-dimensional percolation above a critical collision rate [14]. And Kumpula et al. proposed a weight-topology dynamics model to generate a weighted networks with communities [15]. Nevertheless, the understanding of the hierarchical structure property in complex networks remains a challenge.

On the other hand, recent years have witnessed an attempt by physicists to research the collective phenomena emerging from the interactions of individuals as elementary units in social structure [16]. Castellano et al. review a wide list of topics of collective phenomena ranging from opinion and cultural and language dynamics to crowd behavior, hierarchy formation, human dynamics, and social spreading [16]. Many previous works have studied on static substrates: the interaction pattern is fixed and only opinion, not connections, are allowed to change [16]. This is the case of the dynamics on networks. The opposite case is the dynamics of networks: the links between vertices are formed or removed according to such fixed vertex properties, such as the network formation depending on the present degree [17] and weights [18] of the existing vertices. In fact, real social systems are mostly in between these two extreme cases: both intrinsic property of vertices (like opinions) and connections among them vary in time over comparable temporal scales. The interaction between those two evolutions is then a natural issue to be researched. Holme and Newman [19] proposed a simple model for the coevolution of opinions and social networks in a situation in which both adapt to the other with a single parameter ϕ . They found that the model undergoes a continuous phase transition as the parameter ϕ is varied. Stauffer et al. proposed the model of the coevolution of individual economic characteristics and socioeconomic networks, where links between agents with similar characteristics are more stable than those between agents with vastly different characteristics [20]. They found that a simple scaling law describes the number of distinct surviving characteristic realizations as a function of the number of agents and the number of possible distinct characteristics realizations. Allahverdyan and Petrosyan studied a model for a statistical network formed by interactions between its nodes and links, where each node can be in one of two states and the node-link interaction facilitates

linking between the like nodes [21]. They found that herding and collaboration are not efficient either because linking is too costly or noises are too strong.

In the present paper, we propose a simple model of opinion formation to construct social networks with community property. Our present work is based somewhat on ideas that presented in [19–21], such as two nodes connected via a third one tend to link directly and nodes with the same property (for example the same opinions) tend to connect, etc. However, our present model shows the intrinsic interactionism between the property of vertices (like opinions) and the connections among them, which is different from that in [19] where the opinions of vertices and the connections among them vary in time with probability $(1 - \phi)$ and ϕ respectively.

The rest of the paper is structured as follows. In the following section, we present a simple model of the opinion dynamics in adaptive social network; then, we analyze the crucial role of the rewiring of local attachment (RLA) and the rewiring of global attachment (RGA) in the opinion formation and the emergence of the community structure in social network(s); finally, we discuss the results obtained.

2 Model

As well known, many complex networks arising in real complex systems play the role of underlying frameworks where dynamics occur. Examples include the opinion (rumors, disease) spreads in our society, the computer virus spreads in Internet and the electric current flows through the power grid. The obvious questions are how the dynamics and its underlying network structure interact with each other and how the hierarchical structure of complex network emerges. To see this, we propose a simple model of continuous opinion dynamics to construct social network with hierarchical structure, based on the algorithm of link rewiring of local attachment (RLA) and global attachment (RGA). RLA refers to forming ties with one's network neighbors — "friends of friends" [15]. RGA, in contrast, refers to forming ties independently of the geodesic distance and is attributed to forming social ties through sharing the same activities, [15] take the opinion in our present work for example. As for the opinion dynamics, we choose the celebrated Defiant model (D model) of continuous opinion dynamics [22–25]. Each individual (i.e., vertex or node in network) has a continuous opinion varying from zero to one. Each individual selects randomly one of its neighbors and checks first if an exchange of opinions makes sense. If the two opinion differ by less than ϵ ($0 < \epsilon < 1$), each opinion moves partly in the direction of the other, by amount $\mu \Delta s$, where Δs is the two opinions difference and μ the convergence parameter ($0 < \mu \leq 0.5$); otherwise, the two refuse to discuss seriously and no opinion is changed. The parameter ϵ is called confidence bound or confidence parameter.

For convenience, our model starts from a lattice consisting of N vertices arranged in a ring. Each vertex is connected to all of its neighbors up to some fixed range K to make the network with the minimal coordination number $2K$, which is the minimal links of each

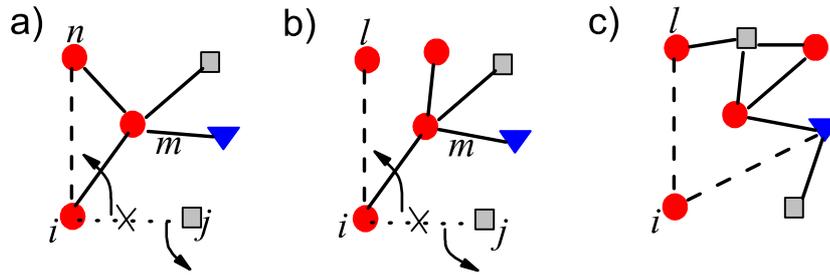


Figure 1: (color online) The model algorithm. (a): a degree local search starts from active vertex i to its confidence nearest neighbor m and then to its second-nearest neighbor n , which is the confidence neighbor of m . Then, vertex i rewires its rewiring link to the chosen vertex n . (b): the active vertex i rewires its rewiring link to one of vertices, which are not the neighbors of i , in the whole network randomly. (c): the active vertex i creates $(2K - k_i)$ links to random vertices while its current links $k_i < 2K$. The same signal of vertices are the confidence ones, for example, the pair of vertices i (circle) and m (circle) is the confidence one and the pair of vertices i and j (square) is not the confidence one in (a) and (b). The dot line with cross will be divided into two rewiring lines, and then each rewiring line rewires to another vertex according to the arrow. The dash line represents the new rewired one or the new added one.

vertex during the evolution process of our model. Initially, each vertex has a continuous opinion from zero to one chosen from a uniform distribution. At time step t , the vertex i 's opinion is labeled as $s_i(t)$. Then, at each time step, one vertex and one of its neighbors, says vertices i and j , are chosen as the active vertices randomly. If $|s_i(t) - s_j(t)| < \epsilon$, i.e., the two opinions differ by less than the confidence parameter ϵ , then each opinion moves partly in the direction of the other as follows:

$$\begin{cases} s_i(t+1) = s_i(t) + \mu[s_j(t) - s_i(t)]; \\ s_j(t+1) = s_j(t) + \mu[s_i(t) - s_j(t)]. \end{cases} \quad (2.1)$$

Otherwise, the link between vertices i and j is broken and divided into two links, which are called as the rewiring links of vertices i and j respectively. One end of each of those two links is attached to vertices i and j respectively, the other ends of those two links are rewired according to RLA with probability p and according to RGA with probability $(1 - p)$, which is different from the original rule in the D model, see (a) and (b) in Fig. 1 respectively. Here, p is a tunable parameter, which shows the competition between the RGA and the RLA during the evolution process of our model. For convenience, we call the two vertices the confidence pair while their opinions differ by less than ϵ in network. More specially, vertex i chooses one of its confidence neighbors, vertex m , with probability $k_m / \sum_{l \in \Gamma'_i} k_l$, where Γ'_i is the subgraph of vertex i 's confidence neighbors and k_m is the connectivity degree of vertex m . If the chosen vertex m has other confidence neighbors, which are not the neighbors of vertex i , apart from i , it chooses one of them, say n , with probability $k_n / (\sum_{l \in \Gamma'_m} k_l - k_i)$. Then, vertex i rewires its rewiring link to the vertex n by the introduction of their common confidence neighbor m . We call this rewiring attachment mechanism as the rewiring of local attachment (RLA), which is difference from that in [15] and enhances the local interaction in social network, see Fig. 1(a). Second, one

vertex, denoted by l , of vertices which are not connected to i by a link is chosen with probability $1/N'_i$, where N'_i is the number of vertices that are not connected to i by a link in the whole network. Hence, $1/N'_i$ varies as time elapses. Then, vertex i rewires its rewiring link to the chosen vertex l . The mechanism corresponds to establishing a new interaction in the whole network by its own action, and we call it the rewiring of global attachment (RGA), Fig. 1(b). Simultaneously, the vertex j also rewires its rewiring link to other vertices following the mentioned algorithm of vertex i above. Finally, if a vertex, says i , has links less than the minimal links $2K$, i.e., $k_i < 2K$, it creates $(2K - k_i)$ links to other vertices in network randomly and its opinion also is reset from the uniform distribution randomly in order to adapt the social environment, see Fig. 1(c). Therefore, the network size of the system and the minimal links of each vertex remain fixed at N and $2K$ respectively.

3 Results

We investigate by numerical simulations the coevolution of the complex network structure and the opinion dynamics. Generality, the system of opinion dynamics reaches a fragmentation state or consensus state, which is related to the confidence parameter ϵ that also been found in many previous works about D mode and is not shown here, as time t elapses. Here, the fragmentation state is defined as that individuals can be divided into two or more camps according to their opinions. Each camp has its opinion that different from others obviously. The consensus state is defined as that all the individuals share the same opinion, see the inset of Fig. 2.

In order to show the crucial role of the adaptive complex network structure in the dynamics of opinion formation, we study the evolution process of our model with the parameters $N = 1000$, $2K = 10$, $\mu = 0.1$ and $\epsilon = 0.3$, since the system reaches the consensus state that is independent of the p . We define the critical time t_c as the time when the system reaches the consensus state first, see the Inset of Fig. 2. In Fig. 2 we represent the evolution of the critical time t_c as a function of the tunable parameter p . We find that t_c increases with increasing p , i.e., the weaker is the local attachment during our model, the easier the system reaches the consensus state. Furthermore, comparing with the previous works about D model [22–25], we find that the RGA enhances the ability of consensus of opinion formation.

As mentioned above, the system will reach a steady state where the opinion dynamics and the complex network structure do not change again as time elapses. Hence, we focus on the complex network structure properties, such as the clustering coefficient, the characteristic path length and the community property, when the system reaches the steady state. There exists the two limit cases about the tunable parameter p . One limit case is $p = 0$, where the link which connects the pair of active vertices whose opinions differ by more than ϵ is always rewired according to the RGA. The characteristic path length $L(0)$ and clustering coefficient $C(0)$ are smaller than those in random network till

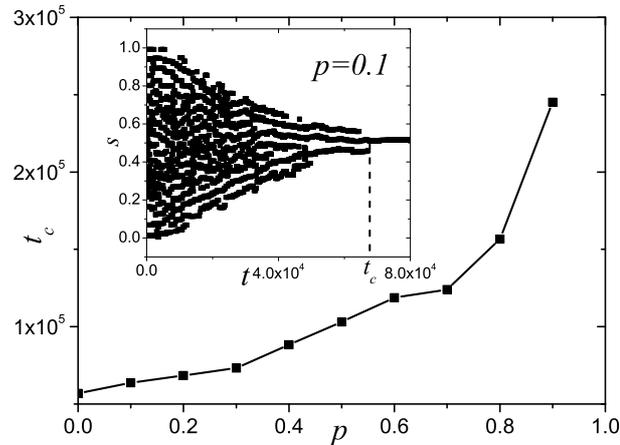


Figure 2: The critical time t_c when all the individuals share the same opinion as a function of the tunable parameter p . Here, each dot has been averaged over 10 realizations. Inset: a special case of opinion evolution as a function of time t for $p=0.1$. The parameters of the model are: $N=1000$, $2K=10$, $\epsilon=0.3$ and $\mu=0.1$.

the coevolution of the complex network structure and the opinion dynamics is dynamical steady. The other is $p=1$, where the link which connects the pair of active vertices whose opinion differ by more than ϵ is always rewired according to the RLA. We find that the clustering coefficient $C(1)=0.606$, which is close to that ($C=0.642$) in the original regular lattice. More interestingly, the characteristic path length $L(1)=9.85$ is much smaller than that ($L=62.86$) in the original regular lattice. Hence, the complex network constructed by the model with parameter $p=1$ has the small-world effect, which means that the average shortest path length between vertices in the network is short and the clustering coefficient of the network is larger than that of the random network. What's more, we also find that the degree distributions are Gaussian and exponential for $p=0$ and $p=1$, see the insets of Fig. 3(a) and (b) respectively.

On the other hand, we pay most of our attention to the tunable parameter p , which shows the role of competition between the algorithms of RLA and RGA in our present model. In Fig. 3, we represent the normalized characteristic path length $L(p)/L(1)$ and clustering coefficient $C(p)/C(1)$ as a function of parameter p . We find that the characteristic path length $L(p)$ and the clustering coefficient $C(p)$ increases with the tunable parameter p increasing, and $C(p)$ increases faster than $L(p)$ as p increases. From Fig. 3, we find that the larger is the tunable parameter p , the more and the stronger communities the network will be divided into. Interestingly, each community in the model is the confidence one, i.e., the difference of each pair of vertices' opinion is less than the confidence parameter ϵ in the same community. There also exists a few vertices (or links), which have a larger betweenness centrality and play the crucial role of bridge in connecting different communities, belonging to more than one community in the network, see the blue vertices and links in Fig. 4. Hence, the network has the "weak links" property. This formation process of the communities in the model is confirmed in Fig. 4, where we

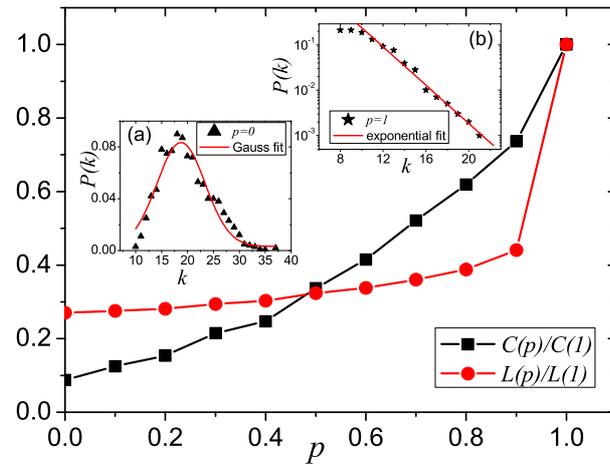


Figure 3: (color online). Normalized Characteristic path length $L(p)$ and Clustering coefficient $C(p)$ as a function of the tunable parameter p . The data are normalized by the values $L(1) = 9.85$ and $C(1) = 0.606$. Insets: the degree distribution of the adaptive social networks for (a), $p = 0$ in normal-normal representation and (b), $p = 1$ in log-normal representation. The parameters of the model are: $N = 1000$, $2K = 10$, $\epsilon = 0.3$ and $\mu = 0.1$.

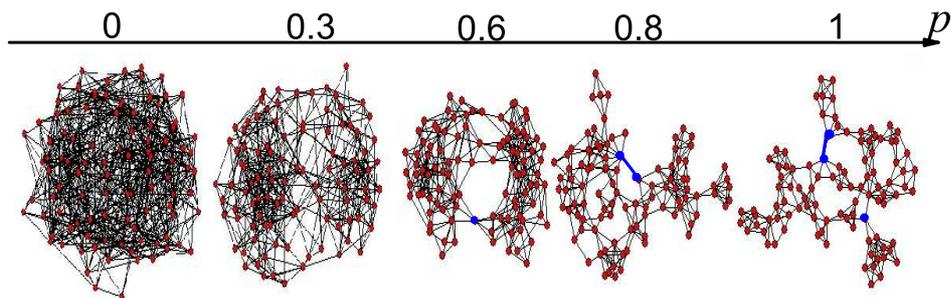


Figure 4: (color online). The complex network structure for several values of p in our model have been studied. Those networks are made up of 100 nodes, in order to have a vivid process picture of the emergence of communities in the model with the tunable parameter p increasing. The blue vertices and links have been emphasized, since each of them has a larger betweenness centrality and plays the crucial role of bridge in connecting different communities. The parameters of the complex network are: $N = 100$, $2K = 4$, $\epsilon = 0.3$ and $\mu = 0.1$.

represent the formation process of communities in our model with various p . Furthermore, we also analyze the degree distribution as a function of p . We find that the degree distribution is between the Gaussian distribution ($p = 0$) and the exponential distribution ($p = 1$), which not shown here.

The formation of communities in complex networks can be quantified by modularity Q [9, 10]:

$$Q = \sum_i (e_{ii} - a_i^2), \tag{3.1}$$

where e_{ij} is the fraction of all edges in the network that link vertices in community i to

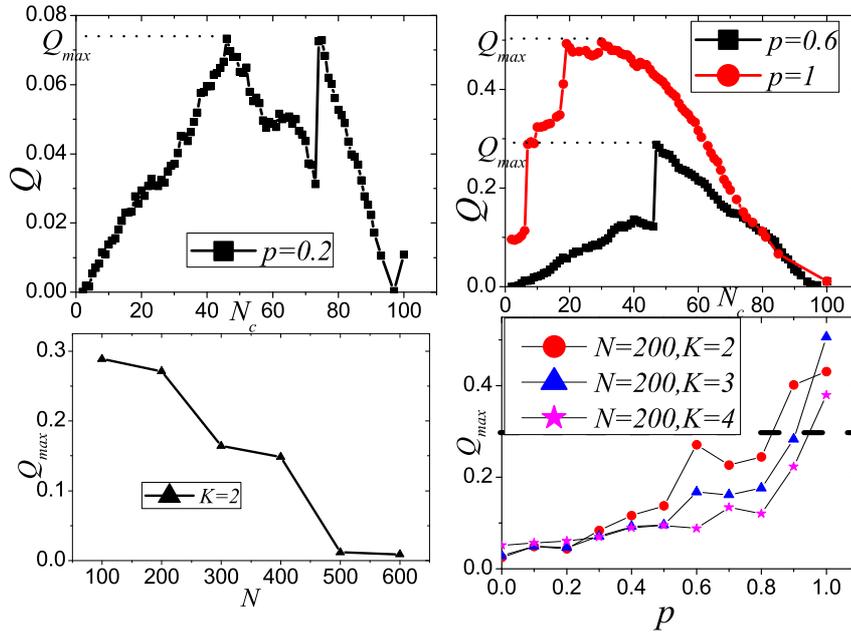


Figure 5: (color online) The modularity Q as a function of the number of communities N_c with various tunable parameters (a) $p=0.2$, (b) $p=0.6$ (square) and $p=1$ (circle) for the complex networks with the same $N=100$, $2K=4$. (c) The evolution of the maximum modularity Q_{max} as a function of the network size N with the same $p=0.6$ and $2K=4$. (d) The Q_{max} as a function of the tunable parameter p in complex networks with different coordination number $2K$. The dash line is $Q_{max}=0.3$. The parameters are $\epsilon=0.3$ and $\mu=0.1$.

vertices in community j and $a_i = \sum_j e_{ij}$ represents the fraction of edges that connect to vertices in community i . In practice, it is found that the modularity value above about 0.3 is a good indicator of significant community structure in a network [9].

Generally, we calculate Q for each split of a network into communities as we move down the dendrogram, and look for local peaks in its value, which indicate particularly satisfactory splits [10]. In Fig. 5, we represent the evolution of the modularity Q as a function of the number of communities N_c with various parameter (a) $p=0.2$, (b) $p=0.6$ (square) and $p=1$ (circle). We find that there are only one or two peaks, and the positions of those peaks correspond closely to the expected number N_c of communities that the original network is divided into. In order to analyze the comparison between RLA and RGA in the formation of complex system structure, we define Q_{max} , which plays the role of order parameter and indicates the best expected divisions of complex systems, as the maximum value among all peaks. In Fig. 5, Q_{max} are (a) 0.0734 for $p=0.2$, (b) $0.288 \approx 0.3$ for $p=0.6$ and 0.493 for $p=1$. On the other hand, we find that Q_{max} increases with p increasing, see Fig. 5(d). The larger is the parameter p , the larger the maximum modularity Q_{max} is. Furthermore, $Q_{max} > 0.3$ for $p > p_c$, where p_c is related to the complex system size N and the coordination parameter $2K$. The larger is the complex system size N , the smaller the maximum modularity Q_{max} is, see the Fig. 5(c). In Fig. 5(d), we find

that the maximum modularity Q_{\max} decreases with $2K$ increasing for $p > 0.5$. All those results show that the factor of RLA enhances the formation of community structure in the adaptive social networks.

4 Conclusions

In summary, we discuss the emergence of community structure in adaptive social network, which inspired by some previous celebrated model, such as the small-world network model [26], the scale-free network model [17] and the weighted complex network model [15]. Generality, the system reaches a steady state, i.e., the opinions and the social network structure do not change again for $t \rightarrow \infty$. The RGA enhances the ability of consensus of opinion formation. Most important of all, the larger is the tunable parameter p , the larger the modularity Q of the complex network is. The modularity $Q > 0.3$ for $p > p_c$, where p_c is related to the initial complex network parameters N and $2K$. Namely, the RLA enhances the formation of community structure in complex social system. Our present work provides a new perspective and tools to understand the formation of community structure and the evolution of opinion dynamics in our real society directly.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant Nos. 10647125, 10635020, 10975057 and 10975062) and the Programme of Introducing Talents of Discipline to Universities under Grant No. B08033.

References

- [1] A. Clauset, C. Moore, M. E. J. Newman, *Nature* 453 (2008) 98.
- [2] F. Wang, Y. Moreno, Y. Sun, *Phys. Rev. E* 73 (2006) 036123.
- [3] L. A. N. Amaral, A. Scala, M. Barthélemy, H. E. Stanley, *Proc. Natl. Acad. Sci. USA* 97 (2000) 11149.
- [4] R. Albert, I. Albert, G. L. Nakarado, *Phys. Rev. E* 69 (2004) 025103.
- [5] R. J. Williams, N. D. Martinez, *Nature*, 404 (2000) 180.
- [6] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, A.-L. Barabási, *Nature* 407 (2000) 651.
- [7] M. C. Lagomarsino, P. Jona, B. Bassetti, H. Isambert, *Proc. Natl. Acad. Sci. USA* 104 (2007) 5516.
- [8] M. Girvan, M. E. J. Newman, *Proc. Natl. Acad. Sci. USA* 99 (2002) 7821.
- [9] A. Clauset, M. E. J. Newman, C. Moore, *Phys. Rev. E* 70 (2004) 066111.
- [10] M. E. J. Newman, M. Girvan, *Phys. Rev. E* 69 (2004) 026113.
- [11] M. E. J. Newman, *Phys. Rev. E* 69 (2004) 066133.
- [12] F. Radicchi, C. Castellano, F. Cecconi, V. Loreto, D. Parisi, *Proc. Natl. Acad. Sci. USA* 101 (2004) 2658.
- [13] F. Wu, B. A. Huberman, *Eur. Phys. J. B* 38 (2004) 331.
- [14] M. González, P. G. Lind, H. J. Herrmann, *Phys. Rev. Lett.* 96 (2006) 088702.

- [15] J. M. Kumpula, J. P. Onnela, J. Saramáki, K. Kaski, J. Kertész, *Phys. Rev. Lett.* 99 (2007) 228701.
- [16] C. Castellano, S. Fortunato, V. Loreto, *Rev. Mod. Phys.* 81 (2009) 591.
- [17] A.-L. Barabási, R. Albert, *Science* 286 (1999) 509.
- [18] K.-I. Goh, B. Kahng, D. Kim, *Phys. Rev. Lett.* 87 (2001) 278701.
- [19] P. Holme, M. E. J. Newman, *Phys. Rev. E* 74 (2006) 056108.
- [20] D. Stauffer, M. Hohnisch, S. Pittnauer, *Physics A* 370 (2006) 734.
- [21] A. E. Allahverdyan, K. G. Petrosyan, *Europhys. Lett.* 75 (2006) 908.
- [22] G. Deffuant, D. Neau, F. Amblard, G. Weisbuch, *Adv. Complex Systems* 3 (2000) 87.
- [23] G. Weisbuch, *Eur. Phys. J. B* 38 (2004) 339.
- [24] G. Weisbuch, G. Deffuant, F. Amblard, *Physica A* 353 (2005) 555.
- [25] L. Guo, X. Cai, *Commun. Comput. Phys.* 5 (2009) 1045.
- [26] D. J. Watts, S. H. Strogatz, *Nature* 393 (1998) 440.