

Growing scale-free small-world networks with tunable assortative coefficient

Qiang Guo^{a,*}, Tao Zhou^b, Jian-Guo Liu^c, Wen-Jie Bai^d,
Bing-Hong Wang^b, Ming Zhao^b

^a*School of Science, DaLian Nationalities University, Dalian 116600, PR China*

^b*Department of Modern Physics, University of Science and Technology of China, Hefei Anhui 230026, PR China*

^c*Institute of System Engineering, Dalian University of Technology, Dalian 116023, PR China*

^d*Department of Chemistry, University of Science and Technology of China, Hefei Anhui 230026, PR China*

Received 24 January 2006; received in revised form 3 March 2006

Available online 2 May 2006

Abstract

In this paper, we propose a simple rule that generates scale-free small-world networks with tunable assortative coefficient. These networks are constructed by two-stage adding process for each new node. The model can reproduce scale-free degree distributions and small-world effect. The simulation results are consistent with the theoretical predictions approximately. Interestingly, we obtain the nontrivial clustering coefficient C and tunable degree assortativity r by adjusting the parameter: the preferential exponent β . The model can unify the characterization of both assortative and disassortative networks.

© 2006 Elsevier B.V. All rights reserved.

Keywords: Complex networks; Scale-free networks; Small-world networks; Assortative coefficient

1. Introduction

In the past few years, no issue in the area of network researching attracted more scientists than the one related to the real networks, such as the Internet, the World-Wide Web, the social networks, the scientific collaboration and so on [1–6]. Recent works on the complex networks have been driven by the empirical properties of real-world networks and the studies on network dynamics [7–29]. Many empirical evidences indicate that the networks in various fields have some common topological characteristics. They have a small average distance like random graphs, a large clustering coefficient and power-law degree distribution [1,2], which are called the small-world and scale-free characteristics. The other characteristic is that the social networks are assortative while almost all biological and technological networks are opposite. The networks with high clustering coefficient and small averaged distance can be generated with an evolution by the small-world model of Watts and Strogatz [1], while the networks with power-law degree distribution can be

*Corresponding author. Tel.: +86 13050537943.

E-mail address: liujg004@yahoo.com.cn (Q. Guo).

generated with an evolution by the scale-free model of Barabási and Albert [2]. The BA model is a pioneering work in the studies on networks, which suggests that the growth and preferential attachment are two main self-organization mechanisms. Although BA model can generate the power-law degree distribution, its assortative coefficient r equals to zero in the limit of large size thus fail to reproduce the disassortative property that extensively exists in the real-world networks. Recently, some models that can generate either assortative or disassortative networks have been reported [30–35]. Wang et al. presented a mutual attraction model for both assortative and disassortative weighted networks. The model found that the initial attraction A of the newly added nodes may contribute to the difference of the assortative and disassortative networks [34]. Liu et al. [36] proposed a self-learning mutual selection model for weighted networks, which demonstrated that the self-learning probability p may be the reason why the social networks are assortative and the technological networks are disassortative. However, one should not expect the existence of a omnipotent model that can completely illuminate the underlying mechanisms for the emergence of disassortative property in various network systems. In this paper, beside the previous studies, we exhibit an alternative model that can generate scale-free small-world networks with tunable assortative coefficient, which may shed some light in finding the possible explanations to the different evolution mechanisms between assortative and disassortative networks.

Dorogovtsev et al. [37] proposed a simple model of scale-free growing networks. In this model, a new node is added to the network at each time step, which connects to both ends of a randomly chosen link undirected. The model can be equally described by the process that the newly added node connects to node i preferentially, then select a neighbor node of the node i randomly. Holme and Kim [38] proposed a model to generate growing scale-free networks with tunable clustering. The model introduced an additional step to get the trial information and demonstrated that the average number of trial information controls the clustering coefficient of the network. It should be noticed that the newly added node connects to the first node i preferentially, while connects to the neighbor node of the first node i randomly. In this paper, we will propose a growing scale-free network model with tunable assortative coefficient. Inspired by the above two models, the new node is added into the network by two steps. In the first step, the newly added node connects to the existing node i preferentially. In the second step, this node selects a neighbor node s of the node i with probability $k_s^\beta / \sum_{j \in \Gamma_i} k_j^\beta$, where β is the parameter named preferential exponent and Γ_i is the neighbor node set of node i . This model will be equal to the Holme–Kim model [37] when $\beta = 0$, and the MRGN model [35] when $\beta = 1$. Specifically, the model can generate a nontrivial clustering property and tunable assortativity coefficient. Therefore, one may find explanations to various real-world networks by our microscopic mechanisms.

2. Construction of the model

Our model is defined as follows:

- (1) Initial condition: the model starts with m_0 connected nodes.
- (2) Growth: at each time step, one new node v with m edges is added at every time step. Time t is identified as the number of time steps.
- (3) The first step: each edge of v is then attached to an existing node with the probability proportional to its degree, i.e., the probability for a node i to be attached to v is

$$\Pi_i = \frac{k_i}{\sum_j k_j}. \quad (1)$$

- (4) The second step: if an edge between v and w was added in the first step, then add one more edge from v to a randomly chosen neighbor s of w with probability P_t according to the following probability:

$$P_s = \frac{k_s^\beta}{\sum_{v \in \Gamma_i} k_v^\beta}. \quad (2)$$

If there remains no pair to connect, i.e., if all neighbors of w were always connected to v , go to the first step instead.

3. Characteristics of the model

3.1. Degree distribution

The degree distribution is one of the most important statistical characteristics of networks. Since some real-world networks are scale-free, whether the network is of the power-law degree distribution is a criterion to judge the validity of the model. By adopting the mean-field theory, the degree evolution of individual node can be described as

$$\frac{\partial k_i}{\partial t} = P(i) + \sum_{j \in \Gamma_i} P(i|j)P(j), \quad (3)$$

where $P(i)$ denotes the probability that the node i with degree k_i is selected at the first step, $P(i|j)$ denotes the conditional probability that node i is a neighbor of node j with degree k_j which has been selected at the first step.

According to the preferential attachment mechanism of the first step, one has

$$P(i) = \frac{k_i}{\sum_j k_j}. \quad (4)$$

The conditional probability $P(i|j)$ can be calculated by

$$P(i|j) = \frac{k_i}{\sum_{l \in \Gamma_j} k_l}. \quad (5)$$

According to the second step, one has that

$$\frac{\partial k_i}{\partial t} = \frac{k_i}{\sum_l k_l} + \sum_{j \in \Gamma_i} \left(\frac{k_i^\beta}{\sum_{s \in \Gamma_j} k_s^\beta} \frac{k_j}{\sum_l k_l} \right). \quad (6)$$

If $\beta = 0$, we get that

$$\frac{\partial k_i}{\partial t} = \frac{k_i}{\sum_l k_l} + \sum_{j \in \Gamma_i} \left(\frac{1}{k_j} \frac{k_j}{\sum_l k_l} \right) = \frac{2k_i}{\sum_l k_l}. \quad (7)$$

Then we can get that $P(k) \sim k^{-3}$, which has been proved by Holme and Kim [38]. If $\beta = 1$, the following formula can be obtained under the assumption that the present network is nonassortative:

$$\frac{\partial k_i}{\partial t} = \frac{k_i}{\sum_j k_j} + \sum_{j \in \Gamma_i} \frac{k_i k_l}{\langle k \rangle k_l \sum_j k_j} = \frac{2k_i}{\sum_j k_j}. \quad (8)$$

We can obtain that the degree distribution $p(k) \sim k^{-r}$ obeys the power-law and the exponent $\gamma = 3$. The numerical results are demonstrated in Fig. 1. From Fig. 1, we can get that the exponents γ of the degree distribution are around -3 when $\beta = -2, -1, 0, 1, 2$. When $\beta > 0$, the exponent γ would increase slightly as the β increases.

3.2. Average distance

The average distance is also one of the most important parameters to measure the efficiency of communication networks, which is defined as the mean distance over all pairs of nodes. The average distance plays a significant role in measuring the transmission delay. Firstly, we give the following lemma [39].

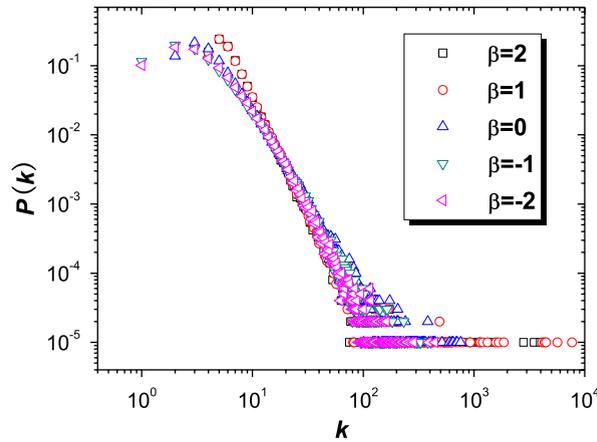


Fig. 1. Degree distribution of the present network with $N = 100\,000$ nodes when $m = 3$ and $P_t = 0.3$. In this figure, $p(k)$ denotes the probability of nodes with degree k in the network. The power-law degree distribution exponents γ of the four probability density function are $\gamma_{\beta=2} = 3.11 \pm 0.05$, $\gamma_{\beta=1} = 3.11 \pm 0.05$, $\gamma_{\beta=0} = 2.93 \pm 0.04$, $\gamma_{\beta=-1} = 2.96 \pm 0.05$ and $\gamma_{\beta=-2} = 2.95 \pm 0.04$.

Lemma 1. For any two nodes i and j , each shortest path from i to j does not pass through any nodes k satisfying that $k > \max\{i, j\}$.

Proof. Since the number i represent the node which is added into the network at time step i , we denote the node set of the shortest path from the node i to j of length $n + 1$ by $i \rightarrow x^1 \rightarrow x^2 \rightarrow \dots \rightarrow x^n \rightarrow j(SP_{ij})$, where $n > 0$. Suppose that $x^k = \max\{x^1, x^2, \dots, x^n\}$, if $x^k \leq \max\{i, j\}$, then the conclusion is true. If $x^k > \max\{i, j\}$, denote the youngest node of SP_{ij} by k . Denote the subpath passing through node k by $l \rightarrow k \rightarrow r$, where the node l and r are the neighbors of node k , then we can prove that node l and r are connected. The shortest path SP_{ij} passes from the node l to r directly, which is in conflict with the hypothesis.

Let $d(i, j)$ represent the distance between node i and j and $\sigma(N)$ as the total distance, i.e., $\sigma(N) = \sum_{1 \leq i < j \leq N} d(i, j)$. The average distance of the present model with order N , denoted by $L(N)$, is defined as follows:

$$L(N) = \frac{2\sigma(N)}{N(N-1)}. \tag{9}$$

According to Lemma 1, the newly added node will not affect the distance between the existing ones. Hence we have

$$\sigma(N+1) = \sigma(N) + \sum_{i=1}^N d(i, N+1). \tag{10}$$

Assume that the $(N+1)$ th node is added to the edge E_{y_1, y_2} , then Eq. (10) can be rewritten as

$$\sigma(N+1) = \sigma(N) + N + \sum_{i=1}^N D(i, y), \tag{11}$$

where $D(i, y) = \min\{d(i, y_1), d(i, y_2)\}$. Denote y as the edge connected the node y_1 and y_2 continuously, then we have the following equation:

$$\sigma(N+1) = \sigma(N) + N + \sum_{i \in A} d(i, y), \tag{12}$$

where the node set $A = \{1, 2, \dots, N\} - \{y_1, y_2\}$ has $(N-2)$ members. The sum $\sum_{i \in A} d(i, y)$ can be considered as the distance from each node of the network to node y in the present model with order $N-2$.

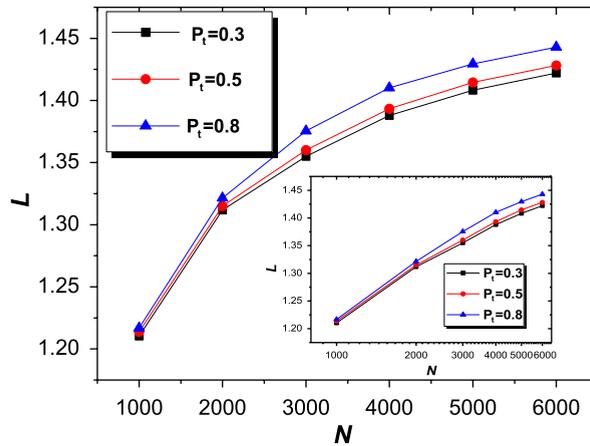


Fig. 2. The dependence between the average distance L and the order N of the present network, when $\beta = 1$ and $m = 4$. One can see that L increases very slowly as $\ln N$ increases. The inset exhibits the curve where L is considered as a function of N . All the data are obtained by 10 independent simulations.

Approximately, the sum $\sum_{i=A} d(i, y)$ is equal to $L(N - 2)$. Hence we have

$$\sum_{i=A} d(i, y) \approx (N - 2)L(N - 2). \tag{13}$$

Because the average distance $L(N)$ increases monotonously with N , this yields

$$(N - 2)L(N - 2) = (N - 2) \frac{2\sigma(N - 2)}{(N - 2)(N - 3)} < \frac{2\sigma(N)}{N - 1}. \tag{14}$$

Then we can obtain the inequality

$$\sigma(N + 1) < \sigma(N) + N + \frac{2\sigma(N)}{N - 1}. \tag{15}$$

Enlarge $\sigma(N)$, then the upper bound of the increasing tendency of $\sigma(N)$ reads

$$\frac{d\sigma(N)}{dN} = N + \frac{2\sigma(N)}{N - 1}. \tag{16}$$

This leads to the following solution:

$$\sigma(N) = (N - 1)^2 \log(N - 1) + C_1(N - 1)^2 - (N - 1). \tag{17}$$

The numerical results are demonstrated in Fig. 2.

3.3. Clustering property

The small-world characteristic consists of two properties: large clustering coefficient and small average distance. The clustering coefficient, denoted by C , is defined as $C = \sum_{i=1}^N C_i / N$, where C_i is the local clustering coefficient for node i . C_i is

$$C_i = \frac{2E(i)}{k_i(k_i - 1)}, \tag{18}$$

where $E(i)$ is the number of edges in the neighbor set of the node i , and k_i is the degree of node i . When the node i is added to the network, it is of degree $m + mP_i$ and $E(i) = mP_i$ (see Fig. 3). If a new node is added to be a neighbor of i at some time step, $E(i)$ will increase by mP_i since the newly added node will connect with one of the neighbors of the node i with probability P_i . Therefore, in terms of k_i , the expression of $E(i)$

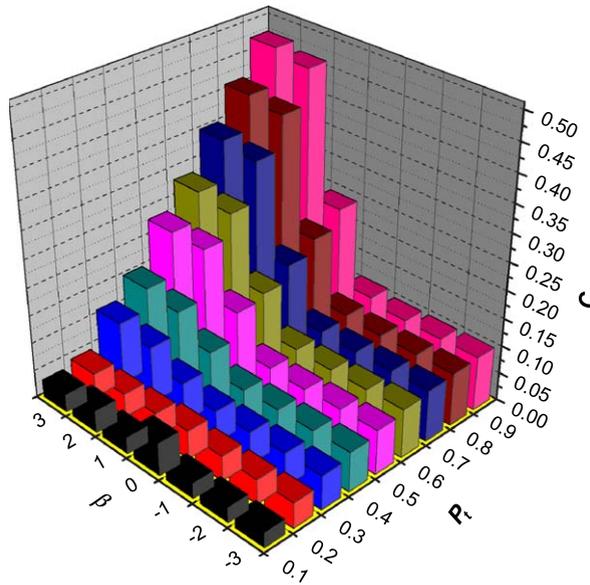


Fig. 3. The scale of C with various β and P_t when $m = 3$. The data are averaged over 10 independent runs of network size of $N = 6000$.

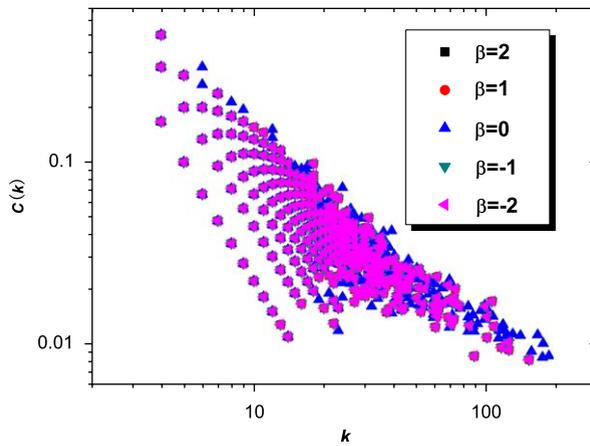


Fig. 4. $C(k)$ vs. k to various β and when $m = 3$ and $P_t = 0.3$. The data are averaged over 10 independent runs of network size of $N = 6000$.

can be written as

$$E(i) = mP_t + P_t[k_i - (m + mP_t)]. \tag{19}$$

Hence, we have that

$$C_i = \frac{2[mP_t + P_t(k_i - m - mP_t)]}{k_i(k_i - 1)} = 2 \left(\frac{mP_t^2}{k_i} + \frac{P_t - mP_t^2}{k_i - 1} \right). \tag{20}$$

This expression indicates that the local clustering scales as $C(k) \sim k^{-1}$, where $C(k)$ denotes the average clustering coefficient value of nodes with degree k . It is interesting that a similar scaling has been observed in many artificial models [38–41] and several real-world networks [42]. The degree-dependent average clustering

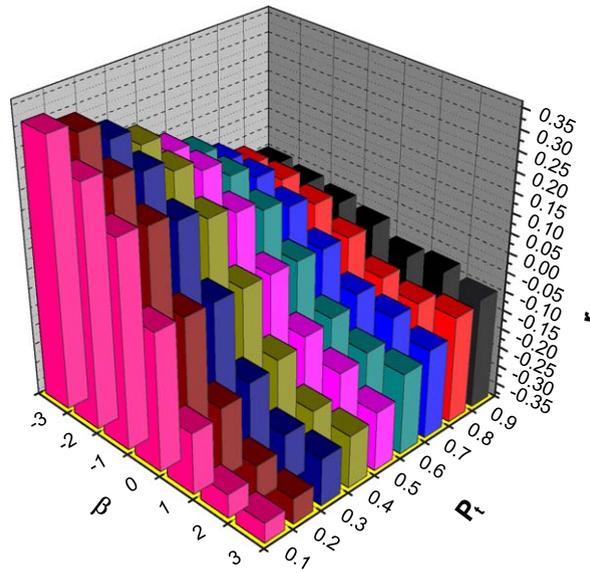


Fig. 5. The scale of r with various β and P_t when $m = 1$. The data are averaged over 10 independent runs of network size of $N = 6000$.

coefficient $C(k)$ has been demonstrated in Fig. 4. Consequently, we have

$$C = \frac{2}{N} \sum_{i=1}^N \left(\frac{mP_t^2}{k_i} + \frac{P_t - mP_t^2}{k_i - 1} \right). \tag{21}$$

Since the degree distribution is $P(k) = ak^{-3}$, where $k = k_{\min}, \dots, k_{\max}$. The constant a satisfies the normalization equation

$$\sum_{k_{\min}}^{k_{\max}} ak^{-3} = 1, \tag{22}$$

one can get that $a = 2k_{\min}^2$. The average clustering coefficient C can be rewritten as

$$C = \frac{2}{N} \sum_{k_{\min}}^{k_{\max}} \left(\frac{NP(k)mP_t^2}{k} + \frac{NP(k)(P_t - mP_t^2)}{k - 1} \right) \approx 2 \sum_{k_{\min}}^{k_{\max}} (ak^{-4}mP_t^2 + ak^{-3}(P_t - mP_t^2)/(k - 1)). \tag{23}$$

The numerical results are demonstrated in Fig. 3. From Fig. 3, we can get that if $\beta \leq 0$, the numerical results are consistent with the theoretical predictions approximately, while if $\beta > 0$, the fluctuations emerges. The departure from analysis results is observed, which may attribute to the fluctuations of the power-law exponent of degree distribution. It is also helpful to compare the present method with previous analysis approaches on clustering coefficient for Holme–Kim model [43,44].

3.4. Assortative coefficient

The assortative coefficient r can be calculated from

$$r = \frac{M^{-1} \sum_i j_i k_i - [M^{-1} \sum_i \frac{1}{2}(j_i + k_i)]^2}{M^{-1} \sum_i \frac{1}{2}(j_i^2 + k_i^2) - [M^{-1} \sum_i \frac{1}{2}(j_i + k_i)]^2}, \tag{24}$$

where j_i, k_i are the degrees of the vertices of the i th edge, for $i = 1, 2, \dots, M$ [45,46].

From Fig. 5, we can find that when $\beta > 0$ the assortative coefficient r increases with the probability P_t , while r decreases with the probability P_t when $\beta < 0$. As $\beta = 0$, r equals to zero approximately.

4. Conclusion and discussion

In this paper, we propose a simple rule that generates scale-free small-world networks with tunable assortative coefficient. The inspiration of this model is to introduce the parameter β to Holme–Kim model. The simulation results are consistent with the theoretical predictions approximately. Interestingly, we obtain the nontrivial clustering coefficient C and tunable degree assortativity r , depending on the parameters β . The model can unify the characterization of both assortative and disassortative networks. Specially, studying the degree-dependent average clustering coefficient $C(k)$ also provides us with a better description of the hierarchies and organizational architecture of weighted networks. Our model may be conducive to future understanding or characterizing real-world networks.

Acknowledgment

This work has been supported by the Chinese Natural Science Foundation of China under Grant nos. 70431001, 70271046 and 70471033.

References

- [1] D.J. Watts, S.H. Strogatz, *Nature* 393 (1998) 440.
- [2] A.-L. Barabási, R. Albert, *Science* 286 (1999) 509.
- [3] R. Albert, A.-L. Barabási, *Rev. Mod. Phys.* 74 (2002) 47.
- [4] S.N. Dorogovtsev, J.F.F. Mendes, *Adv. Phys.* 51 (2002) 1079.
- [5] M.E.J. Newman, *SIAM Rev.* 45 (2003) 167.
- [6] X.F. Wang, *Int. J. Bifurcat. Chaos* 12 (2002) 885.
- [7] W. Li, X. Cai, *Phys. Rev. E* 69 (2004) 046106.
- [8] R. Wang, X. Cai, *Chin. Phys. Lett.* 22 (2005) 2715.
- [9] P.-P. Zhang, K. Chen, Y. He, T. Zhou, B.-B. Su, Y.-D. Jin, H. Chang, Y.-P. Zhou, L.-C. Sun, B.-H. Wang, D.-R. He, *Physica A* 360 (2005) 599.
- [10] P.-P. Zhang, Y. He, T. Zhou, B.-B. Su, H. Chang, Y.-P. Zhou, B.-H. Wang, D.-R. He, *Acta Phys. Sin.* 55 (2006) 60.
- [11] M.H. Li, Y. Fan, J.-W. Chen, L. Gao, Z.-R. Di, J.-S. Wu, *Physica A* 350 (2005) 643.
- [12] J.Q. Fang, Y. Liang, *Chin. Phys. Lett.* 22 (2005) 2719.
- [13] F.C. Zhao, Y.J. Yang, B.-H. Wang, *Phys. Rev. E* 72 (2005) 046119.
- [14] H.J. Yang, F.C. Zhao, L.Y. Qi, B.L. Hu, *Phys. Rev. E* 69 (2004) 066104.
- [15] J.-G. Liu, Y.-Z. Dang, Z.-T. Wang, *Physica A* 366 (2006) 571.
- [16] A.E. Motter, Y.-C. Lai, *Phys. Rev. E* 66 (2002) 065102.
- [17] K.-I. Goh, D.-S. Lee, B. Kahng, D. Kim, *Phys. Rev. Lett.* 91 (2003) 148701.
- [18] T. Zhou, B.-H. Wang, *Chin. Phys. Lett.* 22 (2005) 1072.
- [19] T. Zhou, B.-H. Wang, P.-L. Zhou, C.-X. Yang, J. Liu, *Phys. Rev. E* 72 (2005) 046139.
- [20] M. Zhao, T. Zhou, B.-H. Wang, W.-X. Wang, *Phys. Rev. E* 72 (2005) 057102.
- [21] W.Q. Duan, Z. Chen, Z.R. Liu, W. Jin, *Phys. Rev. E* 72 (2005) 026133.
- [22] F. Jin, L. Xiang, F.W. Xiao, *Physica A* 355 (2005) 657.
- [23] B. Wang, H.W. Tang, C.H. Guo, Z.L. Xiu, T. Zhou, Preprint arXiv:cond-mat/0509711.
- [24] B. Wang, H.W. Tang, C.H. Guo, Z.L. Xiu, Preprint arXiv:cond-mat/0506725.
- [25] J.-G. Liu, Z.-T. Wang, Y.-Z. Dang, *Mod. Phys. Lett. B* 19 (2005) 785.
- [26] J.-G. Liu, Z.-T. Wang, Y.-Z. Dang, Preprint arXiv:cond-mat/0509290.
- [27] C.P. Zhu, S.J. Xiong, Y.J. Tian, N. Li, K.S. Jiang, *Phys. Rev. Lett.* 92 (2004) 218702.
- [28] F. Comellas, G. Fertin, A. Raspaud, *Phys. Rev. E* 69 (2004) 037104.
- [29] F. Comellas, M. Sampels, *Physica A* 309 (2002) 231.
- [30] A. Vázquez, *Phys. Rev. E* 67 (2003) 056104.
- [31] R. Xulvi-Brunet, I.M. Sokolov, *Phys. Rev. E* 70 (2004) 066102.
- [32] M. Catanzaro, G. Caldarelli, L. Pietronero, *Phys. Rev. E* 70 (2004) 037101.
- [33] A. Barrat, M. Barthélemy, R. Pastor-Satorras, A. Vespignani, *Proc. Natl. Acad. Sci. USA* 101 (2004) 3747.
- [34] W.-X. Wang, B. Hu, T. Zhou, B.-H. Wang, Y.-B. Xie, *Phys. Rev. E* 72 (2005) 046140.
- [35] W.-X. Wang, B.-H. Wang, B. Hu, G. Yan, Q. Ou, *Phys. Rev. Lett.* 94 (2005) 188702.
- [36] J.-G. Liu, Y.-Z. Dang, Z.-T. Wang, *Chin. Phys. Lett.* 23 (3) (2006) 746.
- [37] S.N. Dorogovtsev, J.F.F. Mendes, A.N. Samukhin, *Phys. Rev. E* 63 (2001) 062101.
- [38] P. Holme, B.J. Kim, *Phys. Rev. E* 65 (2002) 026107.
- [39] T. Zhou, G. Yan, B.-H. Wang, *Phys. Rev. E* 71 (2005) 046141.

- [40] J.S. Andrade, H.J. Herrmann, R.F.S. Andrade, L.R. da Silva, *Phys. Rev. Lett.* 94 (2005) 018702.
- [41] S.N. Dorogovtsev, A.D. Goltsev, J.F.F. Mendes, *Phys. Rev. E* 65 (2002) 066122.
- [42] E. Ravasz, A.-L. Barabási, *Phys. Rev. E* 67 (2003) 026112.
- [43] G. Szabó, M. Alava, J. Kertész, *Phys. Rev. E* 67 (2003) 056102.
- [44] X. Wu, B.-H. Wang, T. Zhou, W.-X. Wang, M. Zhao, H.-J. Yang, *Chin. Phys. Lett.* 23 (4) (2006) 1046.
- [45] M.E.J. Newman, *Phys. Rev. Lett.* 89 (2002) 208701.
- [46] M.E.J. Newman, *Phys. Rev. E* 67 (2003) 026126.