Eur. Phys. J. B **56**, 167–171 (2007)
DOI: 10.1140/epjb/e2007-00101-0

THE EUROPEAN
PHYSICAL JOURNAL B

Rank-based model for weighted network with hierarchical organization and disassortative mixing

L. Tian[a] and D.-N. Shi[b]

College of Science, Nanjing University of Aeronautics and Astronautics, Nanjing, 210016, P.R. China

Received 25 September 2006 / Received in final form 7 February 2007
Published online 12 April 2007 –

Abstract. In this paper, we study a rank-based model for weighted network. The evolution rule of the network is based on the ranking of node strength, which couples the topological growth and the weight dynamics. Analytically and by simulations, we demonstrate that the generated networks recover the scale-free distributions of degree and strength in the whole region of the growth dynamics parameter ($\alpha > 0$). Moreover, this network evolution mechanism can also produce scale-free property of weight, which adds deeper comprehension of the networks growth in the presence of incomplete information. We also characterize the clustering and correlation properties of this class of networks. It is showed that at $\alpha = 1$ a structural phase transition occurs, and for $\alpha > 1$ the generated network simultaneously exhibits hierarchical organization and disassortative degree correlation, which is consistent with a wide range of biological networks.

PACS. 89.75.-k Complex systems – 89.75.Hc Networks and genealogical trees

1 Introduction

A major source of the recent surge of interest in complex networks has been the discovery that a large number of real-world networks have power-law degree distributions, so called scale-free networks [1–4]. Due to the peculiar structural features and the critical dynamical processes taking place on them [5–8], there have been a tremendous number of works modeling networks with scale-free properties. The previous models of complex networks always incorporate the preferential attachment [4], which may results in scale-free properties. That is, a newly added node is connected to preexisting one with a probability *exactly* proportional to the degree or strength of the target node. In reality, however, this *absolute* quantity information of an agent is often unknown, while it is quite common to have a clear idea about the *relative* values of two agents. In this perspective, Fortunato et al. recently introduced a criterion of network growth that explicitly relies on the ranking of the nodes according to the prestige measure [9]. This rank-based model can well mimic the reality in many real cases that the *relative* values of agents are easier to access than their *absolute* values.

All of the researches mentioned above focused on the topological aspect of graphs, that is, unweighted networks. Recently, the availability of more complete empirical data has allowed scientists to consider the variation of the weights of links that reflect the physical characteristics of many real networks. It is well-known that networks are not only specified by their topology but also by the dynamics of weight taking place along the links. Take the world-wide airport networks (WAN) [10–12] for example: each given link weight w_{ij} is the number of available seats on direct flight connections between the airports i and j. In the scientific collaboration networks (SCN) [13–15], the nodes are identified with authors and the weight depends on the number of coauthored papers. In the light of this need, Barrat, Barthélemy, and Vespignani (BBV) proposed a model that integrated the topology and weight dynamical evolution to study the growth of weighted networks [16–18]. Their model yields scale-free properties of the degree, weight, and strength distributions, controlled by an introduced parameter δ. Recently, Wang et al. have studied the creation and reinforcement of internal connections in weighted network evolution [19–21], which is not considered in BBV model. On the other hand, Wu et al. integrated the deactivation mechanism in the evolution of weighted networks [22].

A weighted network is often denoted by a weighted adjacency matrix with element w_{ij} representing the weight of the link connecting node i and j. In the case of undirected graphs, weights are symmetric $w_{ij} = w_{ji}$, as we will focus on. A natural generalization of connectivity in the case of weighted networks is the node strength defined as $s_i = \sum_{j\in\mathcal{V}(i)} w_{ij}$, where the sum runs over the set $\mathcal{V}(i)$ (neighbors of node i). This quantity is a natural measure of the importance or centrality of a node in the network. As confirmed by measurement, real networks not only exhibit scale-free degree distribution $P(k) \sim k^{-\gamma}$ with $2 \leq \gamma \leq 3$ [10,11], but also the power-law strength

[a] e-mail: tianliang_1985@hotmail.com
[b] e-mail: shi@nuaa.edu.cn

distribution $P(s) \sim s^{-\eta}$ [11] and weight distribution $P(w) \sim w^{-\theta}$ [12]. Moreover, the strength is highly correlated with the degree, and usually displays a scale-free property $s \sim k^{\beta}$ [23,24].

In this paper, we propose a model for weighted network evolution with only ranking information available. We demonstrate that the generated networks recover scale-free distributions of degree and strength. Interestingly, this network evolution mechanism can also produce scale-free property of weight, which is obtained analytically and by simulations. This feature adds deeper comprehension of the networks growth in the presence of incomplete information. The clustering and correlation properties of this class of networks are also investigated, and we found the generated network simultaneously exhibits hierarchical organization and disassortative degree correlation.

2 The model

In the present model, the prestige ranking criterion is strength. The definition of the model is based on two coupled mechanisms: the topological growth and the weights' dynamics. The model dynamics starts from an initial seed of N_0 nodes connected by links with assigned weight w_0.

(1) *Topological growth.* At each time step, a new node n is created and m new links, with an assigned weight w_0 to each, are set between node n and pre-existing nodes. The previous nodes are ranked according to their strength, and the linking probability that the new node be connected to node i depends on the rank R_i of i:

$$\Pi_{n\to i} = \frac{R_i^{-\alpha}}{\sum_\nu R_\nu^{-\alpha}}, \tag{1}$$

where $\alpha > 0$ is a real-valued parameter. Note that the larger the rank of the node is, the more difficult for it to gain new links, which is reasonable in real life.

(2) *Weights' dynamics.* Analogous to the step in the model proposed by Barrat et al. (BBV model) [16], the introduction of the new link on node i will trigger local rearrangements of weights on the existing neighbors $j \in \mathcal{V}(i)$, according to the rule

$$w_{ij} \to w_{ij} + \delta \frac{w_{ij}}{s_i}, \tag{2}$$

where δ is the total induced increase in strength of node i.

3 Probability distributions

We firstly investigate the probability distribution of the generated network. Since the strength-based ranking of a node can change over time, it is hard to analyze the model directly by the ranking of node strength. However, for a growing weighted network, there is a strong correlation between the age of node and its strength, as the older nodes have more chances to receive links. For these considerations, we check this correlation by numerical simulations. Figure 1 shows the node strength as a function of its age. It can be found that the function is monotone increasing with certain fluctuations. Therefore, in the following analytical approach, an approximation is made that we use the ranking by age instead of strength.

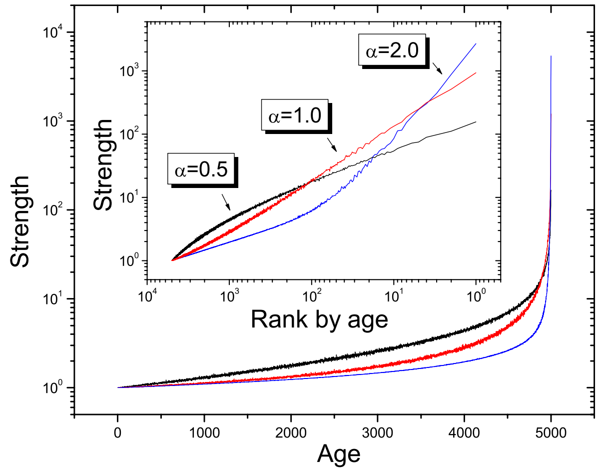

Fig. 1. (Color online) Node strength versus node age with $\alpha = 0.5$, $\alpha = 1.0$, and $\alpha = 2.0$ from the top to the bottom. Inset: Log-Log plot of the relation between node strength and node rank by age. All the data are averaged over 100 independent runs of network size $N = 5000$.

The network growth starts from an initial seed of N_0 nodes, and continues with the addition of one node per unit time, until a size N is reached. Hence, each node is labeled with respect to the time step of its generation, and the natural time scale of the model dynamics is the network size N. If the nodes are sorted by age, from the oldest to the newest, the label of each node coincides with its rank, i.e., $R_i = i \forall i$. Therefore, the node strength s_i is updated according to this evolution equation:

$$\begin{aligned} \frac{ds_i}{dt} &= m\frac{R_i^{-\alpha}}{\sum_j R_j^{-\alpha}}(1+\delta) + \sum_{j\in\mathcal{V}(i)} m\frac{R_j^{-\alpha}}{\sum_l R_l^{-\alpha}}\delta\frac{w_{ij}}{s_j} \\ &= m\frac{i^{-\alpha}}{\sum_j j^{-\alpha}}(1+\delta) + \sum_{j\in\mathcal{V}(i)} m\frac{j^{-\alpha}}{\sum_l l^{-\alpha}}\delta\frac{w_{ij}}{s_j}. \end{aligned} \tag{3}$$

Using the continuous approximation, we treat s, w, and time t as continuous variables and approximate the sums with integrals. Solving equation (3) yields the strength evolution equation:

$$s_i(t) \sim \left(\frac{t}{i}\right)^{\alpha}. \tag{4}$$

Consequently, we can easily obtain in the infinite size limit the probability distribution:

$$P(s) \sim s^{-(1+1/\alpha)}, \tag{5}$$

which shows that the strength distribution of the network follows a power law with exponent $\eta = 1 + 1/\alpha$ for any value of α. It should be noted that this result recovers the earlier one in reference [9].

Similarly to the previous quantities, it is possible to obtain analytical expressions for the evolution of weights and the relative statistical distribution. The rate equation of weight w_{ij} can be written as:

$$\begin{aligned} \frac{dw_{ij}}{dt} &= m\frac{R_i^{-\alpha}}{\sum_j R_j^{-\alpha}}\delta\frac{w_{ij}}{s_i} + m\frac{R_j^{-\alpha}}{\sum_j R_j^{-\alpha}}\delta\frac{w_{ij}}{s_j} \\ &= m\frac{i^{-\alpha}}{\sum_j j^{-\alpha}}\delta\frac{w_{ij}}{s_i} + m\frac{j^{-\alpha}}{\sum_j j^{-\alpha}}\delta\frac{w_{ij}}{s_j}. \end{aligned} \quad (6)$$

Incorporating with equation (4), the above equation can be solved that $w_{ij} \sim (t/t_{ij})^{2\delta(1-\alpha)}$, where $t_{ij} = \max(i,j)$ is the time at which the edge is established. Therefore, the probability distribution $P(w)$ is in this case also a power law $P(w) \sim w^{-\theta}$, where

$$\theta = 1 + \frac{1}{\alpha} + \frac{1}{2\alpha\delta}. \quad (7)$$

This scale-free property of weight is indeed new for networks growth in the presence of incomplete information. We think this result gives more evidence to the significance of the present network evolution mechanism.

In order to check the analytical predictions, we performed numerical simulations of networks generated by the present model, where the prestige ranking criterion is strength. In the upper pattern of Figure 2, we plot the cumulative strength distributions of the networks corresponding to various values of the exponent α. In the logarithmic scale of the plot, they exhibit power-law behaviors in agreement with theoretical results. The relation between α and the exponent η of the strength distribution is showed in the lower pattern of Figure 2, which confirms the validity of equation (5). Together, the power-law distribution of weight $P(w)$ is shown in Figure 3. The analytical predictions can be well confirmed by numerical simulations. Noting the weights' dynamics step in the definition of the model, the triggered increase δ is only arranged locally. Therefore, we expect the proportionality relation $s \sim k$, by which we easily obtain the scale-free distribution of degree $P(k) \sim k^{-\gamma}$ with $\gamma = \eta = 1 + 1/\alpha$. Since there exist no new properties, we do not show them again here.

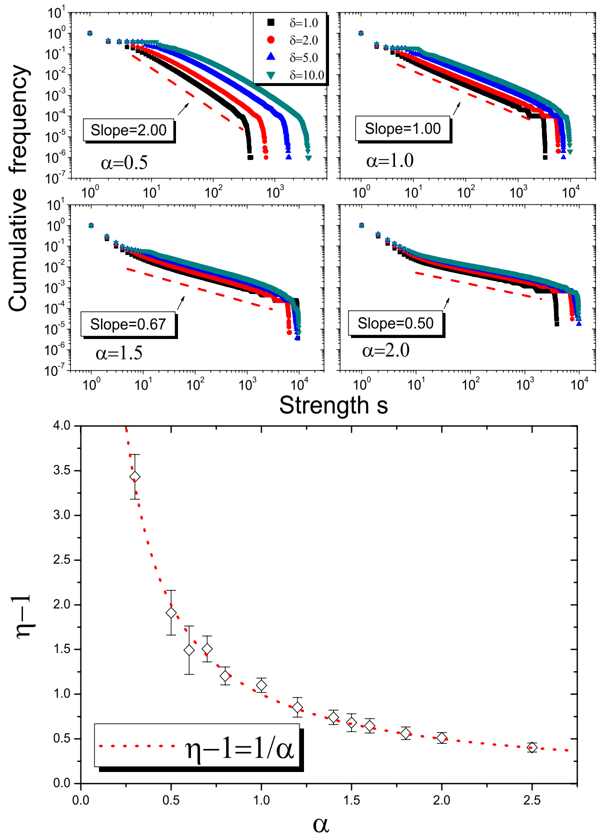

Fig. 2. (Color online) The upper pattern shows the cumulative strength distributions of networks generated by using our model with different parameters $\alpha = 0.5$, $\alpha = 1.0$, $\alpha = 1.5$, and $\alpha = 2.0$. The four dashed lines have slopes 2.00, 1.00, 0.67, and 0.50 separately for comparisons. The lower pattern shows the value of the strength distribution exponent η as a function of α obtained from numerical simulations. The dotted line is the prediction of equation (5). All the data are averaged over 100 independent runs of network size $N = 10^4$.

4 Clustering and correlation

Complex networks display an architecture imposed by the structural and administrative organization of these systems that is not fully characterized by the distributions $P(k)$ and $P(s)$. Indeed, the structural organization of complex networks is mathematically encoded in the various correlations existing among the properties of different vertices. For this reason, a set of topological quantities are customarily studied in order to uncover the network architecture. The first widely used quantity is clustering of nodes. The clustering of a node i is defined as

$$c_i = \frac{1}{k_i(k_i-1)}\sum_{j,h} a_{ij}a_{ih}a_{jh}, \quad (8)$$

where a_{ij} is the adjacency matrix element of the network. It measures the local cohesiveness of the network in the neighborhood of the node. Indeed, it yields the fraction of interconnected neighbors of a given node. The average over all nodes gives the network *clustering coefficient* C which describes the statistics of the density of connected triples.

Further information can be gathered by inspecting the average clustering coefficient $C(k)$ restricted to classes of nodes with degree k:

$$C(k) = \frac{1}{NP(k)}\sum_{i,k_i=k} c_i. \quad (9)$$

In many networks, the degree-dependent clustering coefficient $C(k)$ is a decreasing function of k. It shows that

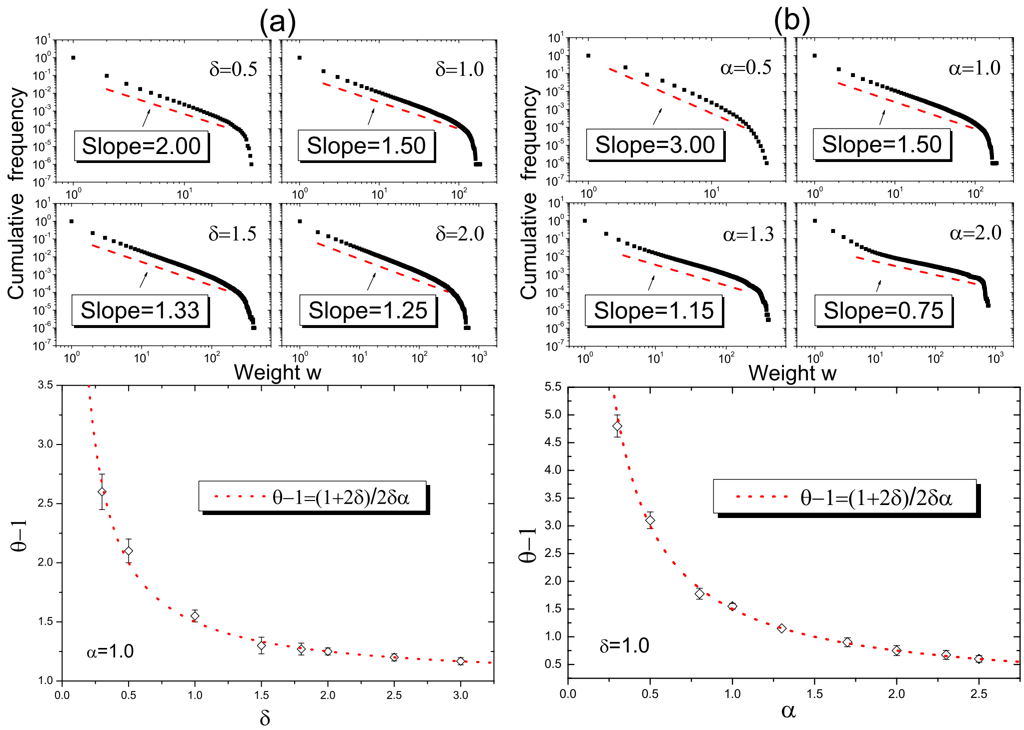

Fig. 3. (Color online) Upper patterns: cumulative weight distributions of networks built according to the present model for different value of (a) δ and (b) α. The lower patterns show the value of the weight distribution exponent θ as a function of δ and α obtained from numerical simulations. The dotted lines are the prediction of equation (7). All the data are averaged over 100 independent runs of network size $N = 10^4$.

low-degree nodes generically belong to well interconnected communities while high-degree sites are linked to many nodes that may belong to different groups which are not directly connected [25]. This is generally the signature of a nontrivial architecture in which the high degree nodes play a distinct role in the network.

Numerical simulations indicate that for $\alpha = 0.5$ the clustering coefficient C seems to converge to zero. This is seen by the accurate fits to algebraic decay forms in Figure 4a. Meanwhile, $C(k)$ is uncorrelated with k, denoting that the network does not possess hierarchical structure. For $\alpha = 2.0$, C approaches a stationary value of about 0.9 in the limit of large N, which is showed in Figure 4b. In this case, a simple scaling form of clustering coefficient, $C(k) \sim k^{-1}$, is obtained, which indicates that the network topology exhibits hierarchical manner.

Another commonly studied network property is the degree correlation (or the mixing pattern) of node i and its neighbor. The average nearest neighbor degree is proposed to measure these correlations

$$k_{nn,i} = \frac{1}{k_i} \sum_j a_{ij} k_j. \tag{10}$$

If degrees of the neighboring nodes are uncorrelated, $k_{nn,i}$ is a constant. When correlation are present, two main classes of possible correlations have been identified: *assortative* behavior if $k_{nn,i}$ increases with k, which indicates that large degree nodes are preferentially connected with other large degree nodes, and *disassortative* if $k_{nn,i}$ decreases with k, which denotes that links are more easily built between large degree nodes and small ones.

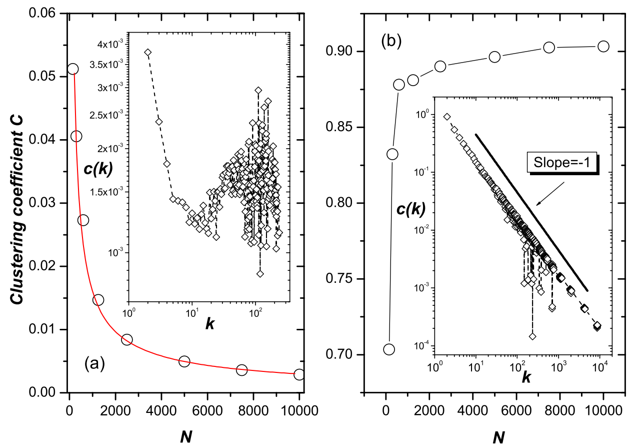

Fig. 4. (Color online) Illustration of the average clustering coefficient C as a function of networks size N for (a) $\alpha = 0.5$ and (b) $\alpha = 2.0$. The insets show the behavior of $C(k)$ depending on degree k. The curves in (a) is fit to algebraic decay form, $2.50 \times N^{-0.73}$. The solid line in the inset of (b) has slope -1 for comparison. All the data are averaged over 100 independent runs.

A simpler measure to quantify this structure is assortative mixing coefficient [26]:

$$r = \frac{L^{-1}\sum_i j_i k_i - [L^{-1}\sum_i \frac{1}{2}(j_i + k_i)]^2}{L^{-1}\sum_i \frac{1}{2}(j_i^2 + k_i^2) - [L^{-1}\sum_i \frac{1}{2}(j_i + k_i)]^2}, \tag{11}$$

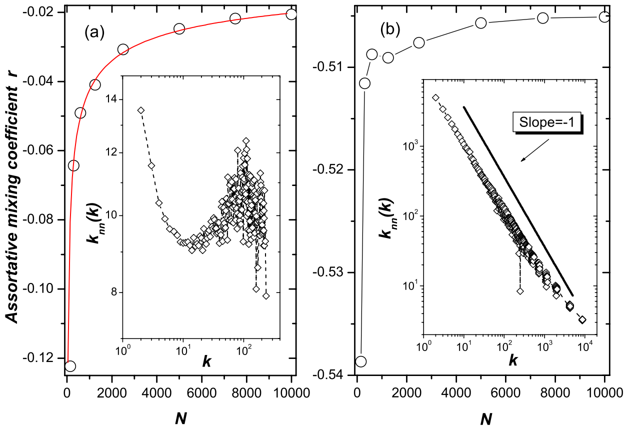

Fig. 5. (Color online) Illustration of the assortative mixing coefficient r as a function of networks size N for (a) $\alpha = 0.5$ and (b) $\alpha = 2.0$. The insets show the behavior of $k_{nn}(k)$ depending on degree k. The curves in (a) is fit to algebraic decay form, $-0.40 \times N^{-0.32}$. The solid line in the inset of (b) has slope -1 for comparison. All the data are averaged over 100 independent runs.

where j_i, k_i are the degrees of nodes at the ends of the ith edges, with $i = 1, ..., L$ (L is the total number of edges in the graph). This quantity takes values in the interval $[-1, 1]$, where positive values mean *assortative* and negative values mean *disassortative*.

The simulation results are shown in Figure 5. When $\alpha = 0.5$, the value of r converges algebraically to zero, and $k_{nn}(k)$ is unrelated with k, which denotes that correlations are absent. On the contrary, when $\alpha = 2.0$ the assortative mixing coefficient is almost independent of network size for large N. Meanwhile, $k_{nn}(k) \sim k^{-1}$, characterizing the *disassortative* degree correlation in the network.

It is indicated that a structural phase transition occurs when the growth dynamics parameter $\alpha = 1$ [27]. For $\alpha < 1$ ($\gamma > 2$), $C(k)$ and $k_{nn}(k)$ are observed as a horizontal line subject to fluctuations, and clustering coefficient C and assortative mixing coefficient r converge to zero in the large limit of network size N. For $\alpha > 1$ ($1 < \gamma < 2$), there emerge a few hub nodes in the network which are linked to almost every other site ,and the generated network exhibits hierarchical topology and *disassortative* degree correlation. Moreover, the clustering coefficient C is independent of network size N and approaches a high value.

5 Conclusion

To sum up, we studied a rank-based model for weighted network. The scale-free properties of probability distributions of degree, strength, and weight are obtained analytically and by simulation. Furthermore, we investigate the clustering and correlation of the network. Specially, in the region of $\alpha > 1$ ($1 < \gamma < 2$), the generated networks can well mimic the biological networks which always appear to be *disassortative* [26] and possess hierarchical organization [25,28–30]. We think that this class of network provides us with a new method to reconstruct the hierarchies and organizational architecture of biological networks, and it may be beneficial for future understanding or charac- terizing the biological networks.

We thank NCET (04-0510) for financial supports.

References

1. R. Albert, A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002)
2. S.N. Dorogrovtsev, J.F.F. Mendels, Adv. Phys. **51**, 1079 (2002)
3. M.E.J. Newman, SIAM Rev. **45**, 167 (2003)
4. A.-L. Barabási, R. Albert, Science **286**, 509 (1999)
5. R. Albert, H. Jeong, A.-L. Barabási, Nature **406**, 378 (2000)
6. R. Pastor-Satorras, A. Vespignani, Phys. Rev. Lett. **86**, 3200 (2001)
7. P. Holme, B.J. Kim, Phys. Rev. E **65**, 066109 (2002)
8. Y. Moreno, R. Pastor-Satorras, A. Vázquez, A. Vespignani, Europhys. Lett. **62**, 292 (2003)
9. S. Fortunato, A. Flammini, F. Menczer, Phys. Rev. Lett. **96**, 218701 (2006)
10. R. Guimera, S. Mossa, A. Turtschi, L.A.N. Amearal, Proc. Natl. Acad. Sci. USA **102**, 7794 (2005)
11. A. Barrat, M. Barthélemy, R. Pastor-Satorras, A. Vespignani, Proc. Natl. Acad. Sci. USA **101**, 3747 (2004)
12. W. Li, X. Cai, Phys. Rev. E **69**, 046106 (2004)
13. M.E.J. Newman, Phys. Rev. E **64**, 016132 (2001)
14. A.-L. Barabási, H. Jeong, Z. Néda, E. Ravasz, A. Schubert, T. Vicsek, Physica A **311A**, 590 (2002)
15. M. Li, J. Wu, D. Wang, T. Zhou, Z. Di, Y. Fan, e-print `arXiv:cond-mat/0501655`
16. A. Barrat, M. Barthélemy, A. Vespignani, Phys. Rev. Lett. **92**, 228701 (2004)
17. A. Barrat, M. Barthélemy, A. Vespignani, Phys. Rev. E **70**, 066149 (2004)
18. B. Hu, G. Yan, W.-X. Wang, W. Chen, e-print `arXiv:cond-mat/0505417`
19. W.-X. Wang, B.-H. Wang, B. Hu, G. Yan, Q. Ou, Phys. Rev. Lett. **94**, 188702 (2005)
20. W.-X. Wang, B. Hu, T. Zhou, B.-H. Wang, Y.-B. Xie, Phys. Rev. E **72**, 046140 (2005)
21. W.-X. Wang, B. Hu, B.-H. Wang, G. Yan, Phys. Rev. E **73**, 016133 (2006)
22. Z.-X.Wu, X.-J. Xu, Y.-H. Wang, Phys. Rev. E **71**, 066124 (2005)
23. K.-I. Goh, B. Kahng, D. Kim, Phys. Rev. E **72**, 017103 (2005)
24. C.L. da Fontoura, T. Gonzalo, e-print `arXiv:cond-mat/0501252`
25. E. Ravasz, A.-L. Barabási, Phys. Rev. E **67**, 026112 (2003)
26. M.E.J. Newman, Phys. Rev. E **67**, 026126 (2003)
27. P.L. Krapivsky, S. Redner, F. Leyvraz, Phys. Rev. Lett. **85**, 4629 (2000)
28. E. Ravasz, A.L. Somera, D.A. Mongru, Z.N. Oltvai, A.-L. Barabási, Science **297**, 1551 (2002)
29. A.-L. Barabási, Z. Oltvai, Nature Reviews Genetics **5**, 101 (2004)
30. S. Bergmann, J. Ihmels, N. Barkai, PloS Biol. **2**, 85 (2004)