An evolving network model with modular growth*

Zou Zhi-Yun(邹志云), Liu Peng(刘 鹏)†, Lei Li(雷 立), and Gao Jian-Zhi(高健智)

School of Civil Engineering & Mechanics, Huazhong University of Science and Technology, Wuhan 430074, *China*

(Received 15 July 2011; revised manuscript received 28 September 2011)

In this paper, we propose an evolving network model growing fast in units of module, according to the analysis of the evolution characteristics in real complex networks. Each module is a small-world network containing several interconnected nodes and the nodes between the modules are linked by preferential attachment on degree of nodes. We study the modularity measure of the proposed model, which can be adjusted by changing the ratio of the number of inner-module edges and the number of inter-module edges. In view of the mean-field theory, we develop an analytical function of the degree distribution, which is verified by a numerical example and indicates that the degree distribution shows characteristics of the small-world network and the scale-free network distinctly at different segments. The clustering coefficient and the average path length of the network are simulated numerically, indicating that the network shows the small-world property and is affected little by the randomness of the new module.

Keywords: evolving, modular growth, small-world network, scale-free network

PACS: 89.75.Hc, 05.10.–a **DOI:** 10.1088/1674-1056/21/2/028904

1. Introduction

Complex networks are currently being studied in many fields of science and engineering. Many systems in nature can be described by models of complex networks, such as an interpersonal relationship network, scientific collaboration network, computer network, protein network, transportation network and electric power network, etc.[1]

The complex network can describe the evolution mechanisms and statistical properties of systems based on graph theory and statistical physics. The empirical study on the real complex networks[2−6] indicated that the most networks each showed small-world property or had a power-law degree distribution, while the small-world property means that the network has a high degree of clustering coefficient and a small average path length. In 1998, Watts and Strogatz[7] and Newman and Watts[8] proposed the WS small-world network model (WS model), in which the transition from regular network to random network is realized. This model describes the small-world property successfully. However, the degree distribution of nodes for this model does not conform with the real networks. The BA scale-free network model (BA model) proposed by Barabási and Albert[9] in 1999, constructs a network with the power-law distribution of nodes degree. However, the clustering coefficient of the BA model is small for a large-size network, indicating that this model does not show significant small-world property.

In addition to the small-world property and the scale-free property, many real networks have module structures,[10−18] which are composed of several modules. The connections are dense in one module, but are sparser between two modules. In the aspect of the research on network modeling with module structure, Li and Maini[19] (2005) proposed an evolving model of scale-free network with community structure. Nakazako and Arita[20] (2007) proposed an evolving network model with the combination of preferential and non-preferential attachments. Cui D *et al.*[21] (2009) studied the properties of asymmetrical evolved community networks. Most evolving network models do not take the module structure into account. The research on network modeling with module structure is still lacking at present.

In the evolving rules of the BA model, network growth is usually achieved by adding only one node at each time step. However, the network size may grow faster in the real world,[22,23] that is, several nodes may become connected each other to constitute a module and join the network at each time step. For instance in scientific collaboration network, each group

*Project supported by the National Natural Science Foundation of China (Grant No. 51078165) and the Fundamental Research Funds for Central Universities, China (Grant No. HUST 2010MS030).

†Corresponding author. E-mail: lpengn@yahoo.cn

© 2012 Chinese Physical Society and IOP Publishing Ltd

including several numbers can be regarded as a module and cooperate with other groups.

In this paper, we propose an evolving network model with modular growth which shows the small-world property, the scale-free property and the module feature at the same time. At each time step several nodes added constitute a module (Hereinafter referred to as the new module), which shows the small-world property. A network grows fast in module units, between which the nodes are linked by preferential attachment on degree of nodes. According to the mean field theory and a numerical example, we study the network statistic properties, including the degree distribution, the modularity measure, the clustering coefficient and the average path length and finally we discuss the small-world property and the scale-free property of the network.

2. Network model

2.1. Network morphology of the new module

In a real network, the module itself usually shows some network morphology. For example, in a scientific collaboration network, the members may be fully connected within a team, which is a fully connected module as shown in Fig. 1(a); in an urban road network, the roads are connected by many random factors within a new planned community, which shows a high level of randomness as shown in Fig. 1(b).

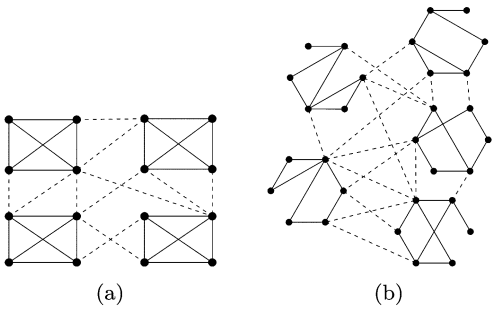


Fig. 1. Schematic diagram of the network morphology of the module.

Under the influence of the randomness, most modules between a fully regular network and fully random network[1] show small-world property significantly. Due to less nodes being included, the new module generally does not show significant scale-free property. Therefore, we take the WS small-world network as the network morphology of the new module in this paper. In the WS model, α denotes the random rewiring probability, which can be adopted to adjust network randomness[7] as shown in Fig. 2. When $\alpha = 0$, the new module shows regular network form; when $\alpha = 1$, the edges of the new module are all rewired randomly which shows ER random network form; when $0 < \alpha < 1$, the new module is a small-world network, which lies between the regular network and random network. With the increase of random rewiring probability α, the randomness of the new module increases gradually.

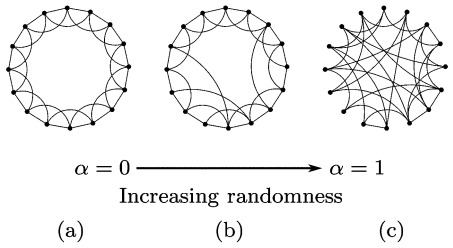


Fig. 2. Random rewiring procedure of WS model: (a) regular network; (b) small-world network; (c) random network.

2.2. Evolving progress

Unlike the previous evolving network models, the proposed model has a higher growth rate. At each time step a module composed of several nodes is added to the network and will experience a complete translation into a small world network (Both the regular network and random network can be regarded as the limited morphology of small-world network). After that, the nodes of the new module would connect to the nodes of the network by preferential attachment.

Figure 3 shows the evolution progress of a network with modular growth. In the initial network the number of nodes is $e_0 = 6$ and the number of edges is $m_0 = 9$. In each module the number of nodes is $s = 10$ and the average degree of nodes is $K = 4$. The new module would be connected to some nodes of the network and the number of connections is $m = 5$.

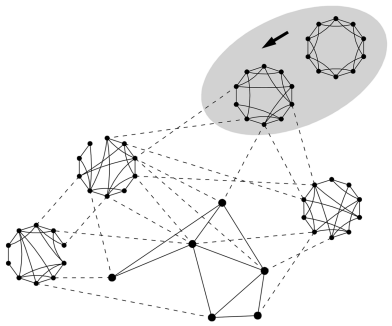


Fig. 3. Evolving progress of a network with modular growth.

The nodes of the new module are chosen based on the mechanism of preferential attachment that the nodes with larger degree priority would be connected to the existing nodes. The existing nodes are also chosen by preferential attachment and become connected to the new module.

The specific evolving process of the proposed network model is presented as follows.

(i) Constructing a new module with small world network morphology

(a) To start with order: Add s nodes to constitute a regular ring lattice, in which each node is connected to the K closest nodes (there are $K/2$ nodes on each side), where K is even. This lattice provides n edges where $n = Ks/2$;

(b) To randomize: Randomly rewire each edge of the lattice with probability α such that self-connections and duplicate edges are excluded. By changing α one can closely monitor the transition between order ($\alpha = 0$) and randomness ($\alpha = 1$).

(ii) Network growth

Starting with a small number (m_0) of nodes and a small number (e_0) of edges, at every time step, add a new module of small-world network containing s nodes and n edges, which would become connected to m existing nodes in the network. After t time steps, in the network the total number of nodes is $(m_0 + st)$ and the total number of edges is $[e_0 + (m + n)t]$.

(iii) Preferential mechanism of the new module attaching to the entire network

a) To choose the nodes of the new module preferentially: Choosing the nodes of the new module which would connect to the nodes existing in the network, we assume that the probability $\Pi(k_i)$ that node i was chosen in the module depends on its degree k_i as

$$\Pi(k_i) = \frac{k_i}{\sum\limits_{x}^{s} k_x}. \tag{1}$$

b) To choose the nodes existing in the network preferentially: When choosing the nodes existed in the network which would connect to the chosen nodes of the new modules, we assume that the probability $\Pi(k_j)$ that node j was chosen in the network depends on its degree k_j:

$$\Pi(k_j) = \frac{k_j}{\sum\limits_{y}^{m_0+st} k_y}. \tag{2}$$

3. Network properties

3.1. Modularity measure

There is a modular structure in the network model proposed, modularity measure is proposed to describe the strength of module structure in a network. According to the Newman and Girvan's study[10] the network modular measure is expressed as

$$Q = \sum_r (e_{rr} - a_r^2), \tag{3}$$

where e_{rr} indicates the fraction of edges that links the vertices in module r to the vertices in module r, a_r indicates the fraction of edges that link to the vertices in module r. For the network model proposed, we can write

$$e_{rr} = \frac{n}{m_0 + (n+m)t}, \tag{4}$$

$$a_r = \sum\nolimits_j e_{rj} = \frac{n+m}{m_0 + (n+m)t}, \tag{5}$$

then

$$\begin{aligned} Q &= \frac{n}{m_0/t + n + m} - \frac{n^2}{(m_0/t)^2 + (n+m)t} \\ &\approx \frac{n}{n+m}. \end{aligned} \tag{6}$$

When the time step t is large enough so that the network size is large, m_0 and the second formula can be ignored. Then we can adjust modularity measure by changing the ratio of inner-module edges n to inter-module edges m.

Take the initial network as shown in Fig. 3 as an example. The relevant parameters can be given as $e_0 = 6$, $m_0 = 9$, $s = 10$, $K = 4$, and $n = 20$. Keeping $n = 20$ unchanged, the values of network modular measure Q are respectively equal to 0.8, 2/3, 0.5, and 1/3 for $m = 5$, 10, 20, 40, indicating that the modular measure Q decreases with m increasing.

3.2. Degree distribution

In this paper, we analyse the degree distribution of the new model using the mean-field theory, by assuming that the change of node degree is continuous. When the time step is large enough, the initial network size is negligible and the change rate of node degree can be expressed as

$$\frac{\partial k_i}{\partial t} = m\frac{k_i}{\sum_{k_l}} = m\frac{k_i}{2m_0 + 2(n+m)t} \approx \frac{m}{2(n+m)}\frac{k_i}{t}. \tag{7}$$

Assuming that the initial value $k_i(t_i) = k_x$, where k_x represents the initial degree of node i belonging to

the new module after being added to the network, equation (7) leads to

$$k_i(t) = k_x \left(\frac{t}{t_i}\right)^{m/(2(n+m))}. \quad (8)$$

The probability density function of time t_i can be written as $P_i(t_i) = 1/(t + e_0/s)$, therefore we can obtain the probability that the degree of node i of the network is equal to k:

$$P(k) = \frac{\partial P(k_i(t) < k)}{\partial(k)} \approx \left(\frac{2n}{m} + 2\right)(k_x)^{2n/m+2} k^{-(2n/m+3)}, \quad (9)$$

where $k_x \le k$.

The probability that the initial degree of node i is k_x depends on the small-world evolution and preferential mechanism in the new module. According to the Newman and Watts's study[8] on the WS model, when $\alpha = 0$, all the edges of the new module would not be rewired and the degree distribution of the new module is the same as the regular network, which is for all K edges; when $\alpha \ge 0$, each edge keeps an endpoint unchanged in the process of rewiring, so each node is linked by $K/2$ edges at least. Therefore, after the small-world evolution of the new module, the probability that the degree of node i of the new module is k' can be written as

$$u(k') = \left\{ \begin{array}{ll} 0, & k' < K/2 \\ \sum\limits_{n=0}^{\min(k'-K/2, K/2)} C_{K/2}^{n} (1-\alpha)^n \alpha^{K/2-n} \dfrac{(\alpha K/2)^{k'-K/2-n}}{(k'-K/2-n)!} \mathrm{e}^{-\alpha K/2}, & k' \ge K/2 \end{array} \right\}. \quad (10)$$

The nodes of the new module will then be chosen preferentially, which would be repeated m times. The probability that a node with k' degree is chosen k'' times (namely its degree increases by k'' and $k'' \le m$) can be written as

$$w(k'') = C_m^{k''} \left(\frac{k'}{sK}\right)^{k''} \left(\frac{sK - k'}{sK}\right)^{m-k''}. \quad (11)$$

Then, we can obtain the probability that the initial degree of the node i of the new module is k_x after the node i has been added into the network, as

$$p(k_x) = \left\{ \begin{array}{ll} C_m^{k_x-K} \left(\dfrac{1}{s}\right)^{k_x-K} \left(\dfrac{s-1}{s}\right)^{m-(k_x-K)}, & \alpha = 0, \\ \sum\limits_{k'=K/2} \left[u(k') \cdot w(k_x - k')\right], & 0 < \alpha \le 1. \end{array} \right\} \quad (12)$$

As seen from Eq. (12), the initial degree of node i of the new module is k_x after the node i has been added into the network. The probability equation is similar to Binomial distribution, related to the small-world evolution and the preferential mechanism in the new module.

According to Eqs. (9) and (12), the probability that the degree of node i in the network is k can be written as

$$P(k) = Ak^{-(A+1)} f(k) - k^{-A} \frac{\partial f(k)}{\partial k}, \quad (13)$$

where $A = 2n/m + 2$ and

$$f(k) = \sum_{k_x=K/2}^{k} (k_x)^A p(k_x).$$

The formula $f(k)$ contains factorial expression, which is very difficult to derive, so we keep the expression $\partial f(k)/\partial k$ in formula $P(k)$.

Figure 4 shows the analytical results of $f(k)$ with double logarithmic coordinates. The value of $f(k)$ increases with k increasing, but the increasing rate gradually decreases. Under the assumption of the continuous change of node degree k, when the node degree k is small, there holds $\partial f(k)/\partial k > 0$ and the probability $P(k)$ does not obey the power-law distribution. When the node degree k is large, $f(k)$ remains unchanged, there are

$$f(k) \approx \sum_{k_x=K/2} (k_x)^A p(k_x), \quad \frac{\partial f(k)}{\partial k} \approx 0,$$

and the probability

$$P(k) \approx \left[A \sum_{k_x=K/2} k_x^A p(k_x)\right] k^{-(A+1)},$$

which obeys a power-law distribution with exponent $\gamma = A + 1 = 2n/m + 3$, where $\gamma > 3$ and the value of γ depends on m and n.

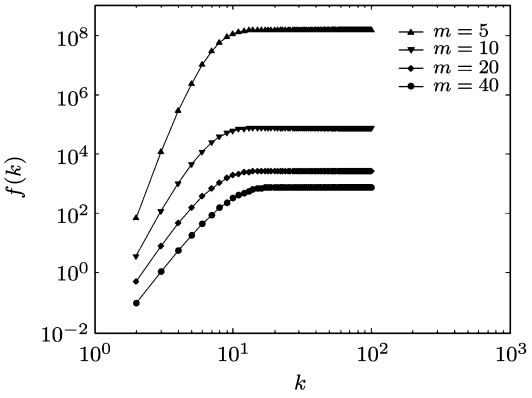


Fig. 4. Dependences of $f(k)$ on node degree k with $m = 5$, 10, 20, 40.

In Fig. 5, a comparison is made between the analytical results of degree distribution and its numerical results in double logarithmic coordinates. For simplicity, let $\partial f(k)/\partial k = 0$ and

$$P(k) = Ak^{-(A+1)} \sum_{k_x=K/2}^{k} (k_x)^A p(k_x).$$

Under this condition, when the node degree k is small, the analytical results shown in the figure are too large compared with the accurate analytical results. However, when the node degree k is large, the analytical results shown in the figure are close to the accurate analytical results. As shown in Fig. 5, we find that when the node degree k is small, the analytical results are larger than the numerical results significantly and with the node degree k increasing, the analytical results becomes close to the numerical results gradually. Therefore, we assert that the accurate analytical results are consistent with the numerical results.

In Fig. 5 the degree distributions clearly show the characteristics of the small-world network and the scale-free network are different at different segments. When the node degree k is small, the probability $P(k)$ has a distribution similar to the degree distribution of the small-world network model and there is a significant peak at the node degree $k = [K + m/s]$ and then the probability $P(k)$ gradually decreases with the increase in node degree. However, when the node degree k is large, the probability $P(k)$ obeys the power-law distribution, and the exponent γ decreases from 11, 7, 5, to 4, with the value of m increasing from 5, 10, 20 to 40.

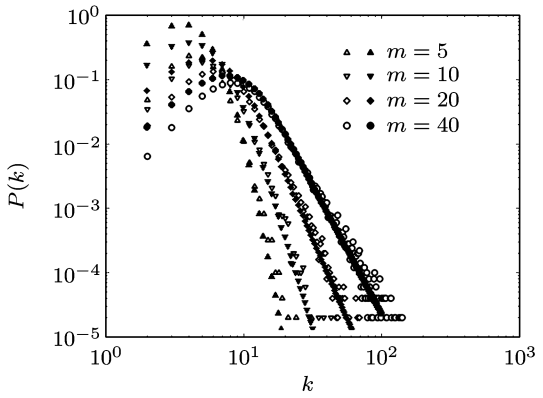


Fig. 5. Degree distributions with $m = 5$, 10, 20, 40, and the relevant parameters $N = 5000$, $\alpha = 0.5$, $e_0 = 6$, $m_0 = 9$, $s = 10$, $K = 4$, and $n = 20$. The solid marks represent the analytical results, while the open marks denote the numerical results.

3.3. Clustering coefficient and average path length

Clustering coefficient is a statistic measure of network clustering, which is defined as the ratio between the actual number of edges among the given neighbour nodes and the theoretical number of edges.[1] For a selected node i in the network, there are k_i edges linked to other node k_i edges. Among the k_i edges there are actually exist E_i edges. The clustering coefficient of node i is then

$$C_i = \frac{2E_i}{k_i(k_i - 1)}. \tag{14}$$

The clustering coefficient C of the network with N nodes is the average of C_i and expressed as

$$C_i = \frac{1}{N} \sum_{i=1}^{N} C_i. \tag{15}$$

Average path length is an important statistic characteristic. Small average path length means that the network shows small-world property.[1] For a connected undirected network, path length from node i to node j is defined as the number of edges on the shortest path between them, while the average path length L is defined as the average value of path lengths between any two nodes in a network and expressed as

$$L = \frac{1}{N(N+1)/2} \sum_{i \geq j} d_{ij}. \tag{16}$$

We study the dependences of clustering coefficient and average path length on the probability α (from 0 to 1) and the network size N that increases from 500 to 5000 with the value of m increasing from 5, 10, 20, to 40. The relevant parameters of the network are $e_0 = 6$, $m_0 = 9$, $s = 10$, $K = 4$, and $n = 20$. The numerical results for a series of computer simulations by using the model are shown in Figs. 6 and 7. In Fig. 6 logarithmic abscissa is used.

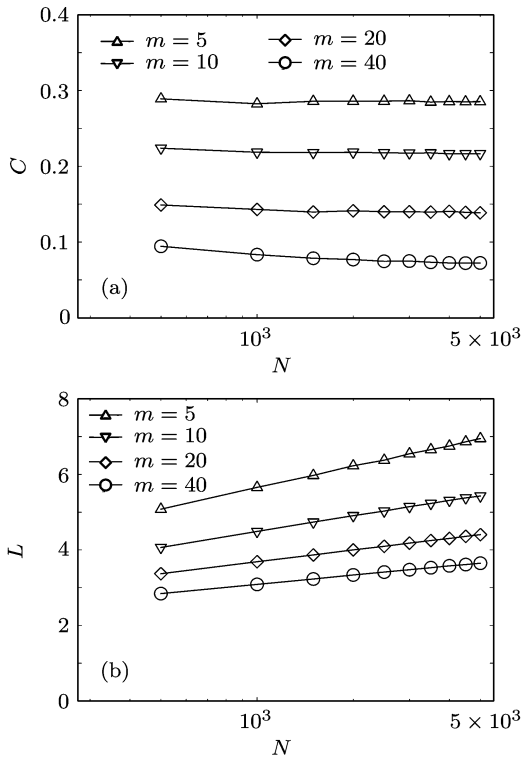


Fig. 6. (a) Dependences of clustering coefficient on network size N with $\alpha = 0.5$ and $m = 5$, 10, 20, 40 and (b) dependences of average path length on network size N with $\alpha = 0.5$ and $m = 5$, 10, 20, 40, respectively.

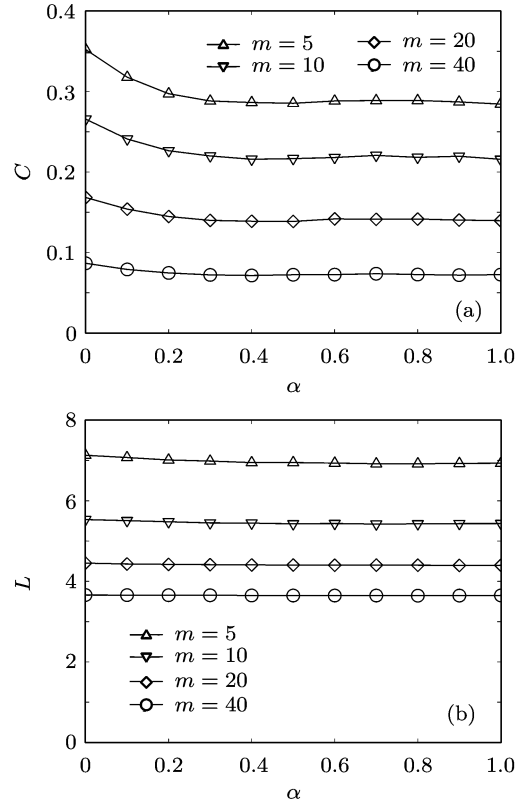


Fig. 7. (a) Dependences of clustering coefficient on probability α with $N = 5000$ and $m = 5$, 10, 20, 40 and (b) dependences of average path length on probability α with $N = 5000$ and $m = 5$, 10, 20, 40.

As shown in Fig. 6, the probability α for this case keeps at 0.5. With the network size increasing, clustering coefficient remains at a high level and average path length grows in logarithmic manner. This indicates that the network proposed shows small-world property.

As shown in Fig. 7 the network size for this case keeps at 5000. With probability α increasing, average path length remains essentially unchanged. When the probability α is in a ranges $[0, 0.3]$, clustering coefficient decreases slightly with α increasing, but remains at a high lever when the probability α increases to 0.3. Overall, the value of the probability α does significantly change neither clustering coefficient nor average path length. That is, the small world property of the network has little effect on the randomness of the new module.

4. Conclusions

Module structure exists in many real networks, which grow not always in units of node but sometimes in units of module. The module usually shows some network morphology.

In this paper we proposed an evolving network model with modular growth. Unlike most of the models studied in the past, a new model grows faster and several interconnected nodes constitute a module of small-world network, joining the network at each time step. We studied the modularity measure of the proposed model, which can be adjusted by changing the ratio between the number of inner-module edges n and the number of inter-module edges m. According to the mean-field theory, we developed an analytical function of the degree distribution, which is verified by a numerical example and indicates that the degree distribution shows that the characteristics of the small-world network and the scale-free network are different at different segments. The degree distribution is similar to that from the small-world model for small node degree k and is a power-law distribution for large node degree k. The clustering coefficient and the average path length were simulated numerically. The results indicates that the network shows the small-world property and is affected little by the randomness of the new module.

References

[1] Albert R and Barabási A L 2002 *Mod. Phys. Rev.* **74** 47
[2] Jiang B and Claramunt C 2004 *Environment and Planning* B **31** 151
[3] Li Y, Cao H D, Shan X M, Ren Y and Yuan J 2009 *Chin. Phys.* B **18** 1721
[4] Montis A D, Barthelémy M, Chessa A and Vespignanni A 2007 *Environment and Planning* B **34** 905

[5] Lammer S, Gehlsen B D and Helbing 2006 *Physica* A **363** 89

[6] Porta S, Crucitti P and Latora V 2006 *Physica* A **369** 853

[7] Watts D J and Strogatz S H 1998 *Nature* (London) **393** 440

[8] Newman M E J and Watts D J 1999 *Phys. Lett.* A **263** 341

[9] Barabási A L and Albert R 1999 *Science* **286** 509

[10] Cui A X, Fu Y, Shang M S, Chen D B and Zhou T 2011 *Acta Phys. Sin.* **60** 038901 (in Chinese)

[11] Newman M E J and Girvan M 2004 *Phys. Rev.* E **69** 026113

[12] Girvan M and Newman M E J 2002 *Proc. Natl. Acad. Sci.* USA **99** 7821

[13] Palla G, Derényi I, Fańkas I and Vicsek T 2005 *Nature* (London) **435** 814

[14] Alves N A 2007 *Phys. Rev.* E **76** 036101

[15] Ravasz E, Somera A L, Mongru Z N, Oltvai Z N and Barabási A L 2002 *Science* **297** 1551

[16] Rives A W and Galitski T 2003 *Proc. Natl. Acad. Sci.* USA **100** 1128

[17] Eriksen K A, Simonsen I, Maslov S and Sneppen K 2003 *Phys. Rev. Lett.* **90** 148701

[18] Guimerá R, Mossa S, Turtschi A and Amaral L A N 2005 *Proc. Natl. Acad. Sci.* USA **102** 7794

[19] Li C G and Maini P K 2005 *J. Phys.* A: *Math. Gen.* **38** 9741

[20] Nakazato K and Arita T 2007 *Physica* A **376** 673

[21] Cui D, Gao Z Y and Zheng J F 2009 *Chin. Phys.* B **18** 516

[22] Dorogovtsev S N and Mendes J F F 2001b *Phys. Rev.* E **63** 056125

[23] Barabási A L, Jeong H, Néda Z, Ravasz E, Schubert A and Vicsek T 2002 *Physica* A **311** 590