Scale-free multicomponent growing networks

Jianhong Ke*
School of Physics and Electronic Information, Wenzhou Normal College, Wenzhou 325027, China

(Received 18 November 2003; published 16 March 2004)

We propose a multicomponent growing network model which consists of many types of nodes as well as links only between the nodes of different types. Such a multicomponent network is constructed by (i) introducing a new node of one type and immediately linking it to a preexisting node of the other type, and (ii) creating a new link between two nodes of different types. We then investigate the connectivity of the multicomponent growing networks by means of the rate equations. For a network system with shifted or asymptotically linear connection rate kernels, the degree distributions take scale-free power-law forms, while a random growing network has exponential degree distributions.

DOI: 10.1103/PhysRevE.69.037101 PACS number(s): 89.75.Hc, 05.10.−a, 05.40.−a, 87.23.Ge

In the last few years, complex networks which play an important role in many natural and social fields [1–5] have attracted considerable interest. In terms of random graph, a network consists of nodes and links, where the nodes represent the individuals and the links represent the interactions between two different individuals. Most interestingly, many real-world complex networks are of scale-free degree distributions and small-world properties [6–9]. In particular, recent researches exhibited that for a wide variety of open evolving network systems, such as the World Wide Web, the electrical distribution system, and the biological system, the degree distributions take a power-law form [10–16]. In order to mimic such networks with complex topology, Barabási and Albert introduced a simple growing network (GN) model (well known as the BA model) in which new nodes are continuously added to the network and meanwhile, they attach preferentially to the old nodes that are already well connected [7]. Some modified versions of the BA model (e.g., by introducing the initial attractiveness and aging of nodes) were also investigated carefully [11,17]. However, most of these investigations focused on the single-component networks that comprise a sole type of nodes, and only a few works devoted their efforts to understanding the multicomponent network models [14–16]. In fact, for some real network systems there may exist different types of nodes and different component nodes may have distinct properties. For example, in the web of human sexual contacts [4], there are two kinds of individuals, males and females; moreover, the common sexual partners of males are the females, and *vice versa*.

In this work, we shall construct a multicomponent growing network on the basis of the network models proposed in Refs. [13,15]. Assume that there are L types of nodes, denoted as A_l $(l=1,2,\dots,L)$, in a multicomponent network. At each time step, a new link is added to the network in one of the two ways: with probability q a new link is created between two already existing nodes of different types, or with probability p_l a new node of type A_l is added to the network and immediately attached to an already existing node of any other type A_m $(l,m=1,2,\dots,L$ and $m\neq l)$. Obviously, $q+\Sigma_{l=1}^{L}p_l=1$. Here we only consider the model in which the links among the same type of nodes are forbidden. In some other situations, it may be more sound that the links between two nodes of the same type are also permissive [14,15]. We believe that our model may mimic some real-world complex systems such as the above-mentioned human sexual contact web.

By employing the standard probabilistic method or generating function technique one may readily solve such a growing network model (see, e.g., Refs. [1,18]). Additionally, Krapivsky *et al.* introduced another simpler but useful rate equation approach [12,13], which can be used to study more general evolving graph systems. Here we also investigate the evolution properties of the multicomponent GN model by means of the rate equations. Let the number of the A_l type nodes with k links be N_{lk} $(l=1,2,\dots,L)$. Then the degree distribution $N_{lk}(t)$ evolves according to the rate equation

$$\frac{dN_{lk}}{dt}=\sum_{m=1}^{L}p_m\frac{W_{k-1}^{(l;m)}N_{l,k-1}-W_k^{(l;m)}N_{lk}}{\sum_{l=1}^{L}\sum_k W_k^{(l;m)}N_{lk}}+q\frac{\sum_{1\le m\le L}\sum_{k'}N_{mk'}[V_{k-1,k'}^{(l;m)}N_{l,k-1}-V_{kk'}^{(l;m)}N_{lk}]}{\frac{1}{2}\sum_{l,m=1}^{L}\sum_{k,k'}V_{kk'}^{(l;m)}N_{mk'}N_{lk}}+p_l\delta_{k1},\quad l=1,2,\dots,L, \tag{1}$$

with the boundary condition $N_{l0}(t)\equiv 0$. Here $W_k^{(l;m)}$ represents the preferential connection rate at which a newly introduced node of type A_m is linked to a preexisting (i.e., old) A_l type node with k links and $V_{k;k'}^{(l;m)}$ denotes the connection rate of a new link created between an old A_l type node with k links and an old A_m type node with k' links. It follows from the connection restriction of our network model that $W_k^{(l;l)}\equiv 0$ and $V_{k;k'}^{(l;l)}\equiv 0$ for all k and k'. In Eq. (1), the term, $p_m W_{k-1}^{(l;m)}N_{l,k-1}/\Sigma_{l=1}^{L}\Sigma_k W_k^{(l;m)}N_{lk}$, on the right-hand side accounts for the gain in the number of the A_l type nodes with k links due to the process in which a newly introduced node of type A_m is connected to an old A_l type node with $k-1$

*Electronic address: kejianhong@yahoo.com.cn

links ($l,m=1,2,\ldots,L$ and $l\neq m$); while the term, $p_m W_k^{(l;m)} N_{lk}/\Sigma_{l=1}^{L}\Sigma_k W_k^{(l;m)} N_{lk}$, accounts for the loss in N_{lk} due to the new A_m type node attached to an old A_l type node with k links. The denominator, $\Sigma_{l=1}^{L}\Sigma_k W_k^{(l;m)} N_{lk}$, is the appropriate normalization factor. Similarly, the second group of terms accounts for the changes caused by the creation of a new link between an old A_l type node with $k-1$ (or k) links and an old node of type A_m ($m\neq l$). The last term accounts for the continuous introduction of new nodes.

Summing up Eq. (1) over all k, we obtain the evolution equations of the total number of the same type of nodes, $\dot{M}_0^{(l)}=\Sigma_k \dot{N}_{lk}=p_l$, which yield the solutions $M_0^{(l)}(t)=p_l t+M_0^{(l)}(0)$, $l=1,2,\ldots,L$. Obviously, the total number of nodes of the same type increases at a fixed rate independent of the connection rate kernels $W_k^{(l;m)}$ and $V_{kk'}^{(l;m)}$. Moreover, multiplying Eq. (1) with k and summing them up, one can then find that the first moment of the degree distributions obeys $\dot{M}_1=\Sigma_{l=1}^{L}\Sigma_k k\dot{N}_{lk}=2$. The solution of the first moment is readily obtained, $M_1(t)=2t+M_1(0)$. This shows that the total number of the links among the nodes of different type is also independent of the connection rate kernels. Except for the total number of nodes and the total number of links, the degree distributions as well as their higher moments may be crucially dependent on the connection kernels.

For a network model in which all the connection rates ($W_k^{(l;m)}$ and $V_{kk'}^{(l;m)}$, $l,m=1,2,\ldots,L$) are sublinear or linear in k and k', we can conclude, according to the comprehensive results of Ref. [12], that the solutions of Eq. (1) take the following forms at large times:

$$N_{lk}(t)=n_{lk}t,\quad l=1,2,\ldots,L, \tag{2}$$

where n_{lk} are independent of time t. In the context of this text, n_{lk} is also called the degree distribution of nodes of type A_l ($l=1,2,\ldots,L$). Otherwise, if at least one of the connection rate kernels is superlinear, the phenomenon that a single dominant gel node is linked to almost all nodes of other types will also arise in our multicomponent GN model. This is the so-called "winner takes all" phenomenon [12]. Thus, all the connection rate kernels should be known in detail before one can derive the analytical solutions of the degree distributions from Eq. (1). In this work, we only focus on the multicomponent networks with linear or sublinear connection kernels. Such networks may exhibit the scale-free properties of the degree distributions and are therefore expected to mimic some real-world systems.

Consider a model with linear or sublinear connection kernels. Thus, the long-time degree distributions evolve according to the forms (2). Consequently, in the long-time limit,

$$\sum_{l=1}^{L}\sum_{k} W_k^{(l;m)} N_{lk}(t)=w_m t,$$

$$\sum_{l,m=1}^{L}\sum_{k,k'} V_{kk'}^{(m;l)}[N_{lk}(t)N_{mk'}(t)]=v t^2,\quad m=1,2,\ldots,L, \tag{3}$$

where w_l and v are independent of time and dependent only on the connection rate kernels. Substituting Eqs. (2) and (3) into Eq. (1), we obtain

$$n_{lk}=[\overline{W}_{l,k-1}n_{l,k-1}-\overline{W}_{lk}n_{lk}]+[\overline{V}_{l,k-1}n_{l,k-1}-\overline{V}_{lk}n_{lk}]+p_l\delta_{k1},\quad l=1,2,\ldots,L, \tag{4}$$

with the shorthand notations $\overline{W}_{lk}=\Sigma_{1\leqslant m\leqslant L}(p_m/w_m)W_k^{(l;m)}$ and $\overline{V}_{lk}=\Sigma_{1\leqslant m\leqslant L}\Sigma_{k'}[(2q/v)V_{k,k'}^{(l;m)}n_{mk'}]$. From the recursion formula (4) one then derives the following implicit solutions of the degree distributions:

$$n_{lk}=\frac{p_l}{\overline{W}_{lk}+\overline{V}_{lk}}\prod_{j=1}^{k}\left(1+\frac{1}{\overline{W}_{lj}+\overline{V}_{lj}}\right)^{-1},\quad l=1,2,\ldots,L. \tag{5}$$

Equation (5) gives the universal solutions of the degree distributions of the multicomponent network systems with arbitrary sublinear or linear connection kernels. Hence, once the connection rate kernels are given, the explicit expressions of the degree distributions can be directly derived from Eq. (5).

Since the two-component GN model contains the generic structure of the multicomponent network system, we limit our investigations of the explicit solutions of Eq. (5) in the two-component situation. Consider a two-component network with the general connection kernels $W_k^{(1;2)}=W_{1k}$, $W_k^{(2;1)}=W_{2k}$, and $V_{kk'}^{(1;2)}=V_{1k}V_{2k'}$. Then Eq. (5) reduces to

$$n_{1k}=\frac{p_1 w_2 v_1}{p_2 v_1 W_{1k}+q w_2 V_{1k}}\prod_{j=1}^{k}\left(1+\frac{w_2 v_1}{p_2 v_1 W_{1j}+q w_2 V_{1j}}\right)^{-1},$$

$$n_{2k}=\frac{p_2 w_1 v_2}{p_1 v_2 W_{2k}+q w_1 V_{2k}}\prod_{j=1}^{k}\left(1+\frac{w_1 v_2}{p_1 v_2 W_{2j}+q w_1 V_{2j}}\right)^{-1}, \tag{6}$$

where v_1 and v_2 are two constants satisfying the equations $v_l=\Sigma_k V_{lk}n_{lk}$, $l=1,2$. From Eq. (6) one can derive the explicit solutions of the degree distributions for two-component GN systems.

We first consider a simple network system with the shifted linear connection rates $W_{1k}=k+\lambda_1$, $W_{2k}=k+\lambda_2$, $V_{1k}=k+\lambda_3$, and $V_{2k}=k+\lambda_4$. Here the four parameters λ_m ($m=1,2,3,4$) are constants larger than -1 so as to ensure that all the corresponding connection rates are positive. The four parameters are also called additional attractiveness (see, e.g., Ref. [19]). In this network, one can easily find the relations $v_1=1+p_1\lambda_3$, $v_2=1+p_2\lambda_4$, $w_1=1+p_2\lambda_2$, and $w_2=1+p_1\lambda_1$. By expanding Eq. (6) we then obtain the exact solutions

$$n_{1k}=c_1\frac{\Gamma(k+b_1)}{\Gamma(k+a_1+1)},\quad n_{2k}=c_2\frac{\Gamma(k+b_2)}{\Gamma(k+a_2+1)}, \tag{7}$$

where

$$a_1=\left(1+\frac{p_2\lambda_1}{1+p_1\lambda_1}+\frac{q\lambda_3}{1+p_1\lambda_3}\right)\left(\frac{p_2}{1+p_1\lambda_1}+\frac{q}{1+p_1\lambda_3}\right)^{-1},$$

$$a_2=\left(1+\frac{p_1\lambda_2}{1+p_2\lambda_2}+\frac{q\lambda_4}{1+p_2\lambda_4}\right)\left(\frac{p_1}{1+p_2\lambda_2}+\frac{q}{1+p_2\lambda_4}\right)^{-1},$$

$$b_1=\left(\frac{p_2\lambda_1}{1+p_1\lambda_1}+\frac{q\lambda_3}{1+p_1\lambda_3}\right)\left(\frac{p_2}{1+p_1\lambda_1}+\frac{q}{1+p_1\lambda_3}\right)^{-1},$$

$$b_2=\left(\frac{p_1\lambda_2}{1+p_2\lambda_2}+\frac{q\lambda_4}{1+p_2\lambda_4}\right)\left(\frac{p_1}{1+p_2\lambda_2}+\frac{q}{1+p_2\lambda_4}\right)^{-1},$$

$$c_1=p_1\frac{\Gamma(2+a_1)}{\Gamma(1+b_1)}\left[1+\frac{p_2(1+\lambda_1)}{1+p_1\lambda_1}+\frac{q(1+\lambda_3)}{1+p_1\lambda_3}\right]^{-1},$$

and

$$c_2=p_2\frac{\Gamma(2+a_2)}{\Gamma(1+b_2)}\left[1+\frac{p_1(1+\lambda_2)}{1+p_2\lambda_2}+\frac{q(1+\lambda_4)}{1+p_2\lambda_4}\right]^{-1}.$$

Since $\Gamma(k+b)/\Gamma(k+a)\sim k^{b-a}$ for large k, from Eq. (7) we find the degree distributions in the $k\gg 1$ region take the pure power-law forms

$$n_{1k}\sim k^{-\nu_1},\quad n_{2k}\sim k^{-\nu_2}, \tag{8}$$

with the exponents

$$\nu_1=1+\left(\frac{p_2}{1+p_1\lambda_1}+\frac{q}{1+p_1\lambda_3}\right)^{-1},$$

$$\nu_2=1+\left(\frac{p_1}{1+p_2\lambda_2}+\frac{q}{1+p_2\lambda_4}\right)^{-1}. \tag{9}$$

These indicate that the two-component network with linear preferential connection kernels is indeed scale-free. We now compare our predictions with the measurements for the web of human sexual contacts. The relevant exponents for the degree distributions are $\nu=2.54\pm 0.2$ for females and $\nu=2.31\pm 0.2$ for males [4]. Equation (9) also shows that the exponents ν_1 and ν_2 are both larger than 2; moreover, one can choose the above-mentioned seven parameters so that the exponents given by Eq. (9) have the same values as those for the real-world web of human sexual contacts. On the other hand, in another two-component network with asymptotically linear connection rate kernels, one can also tune the exponents ν_1 and ν_2. As Krapivsky *et al.* did [12], we assume that $W_{lk}\to a_{l\infty}k$ and $V_{lk}\to b_{l\infty}k$ ($l=1,2$) as $k\to\infty$ while the others are arbitrary. Expanding Eq. (6) also yields the power-law degree distributions (8) with the exponents

$$\nu_1=1+\frac{w_2 v_1}{p_2 v_1 a_{1\infty}+q w_2 b_{1\infty}},\quad \nu_2=1+\frac{w_1 v_2}{p_1 v_2 a_{2\infty}+q w_1 b_{2\infty}}. \tag{10}$$

The results show that the asymptotically linear connection rate kernel may be another possible candidate for describing the structure of the sexual contact web.

The second example is a random GN model. The connection probability of a new node attached to an old one and the creation probability of a new link between two old nodes are both independent of the already existing link number of the target node. We then simply set all the connection rates equal to unit. So, $w_2=v_1=p_1$ and $w_1=v_2=p_2$. Equation (6) then yields

$$n_{1k}=p_1^2(p_2+q)^{k-1},\quad n_{2k}=p_2^2(p_1+q)^{k-1}. \tag{11}$$

The results show that in the two-component random network system, the degree distributions take simple exponential forms.

Finally, we investigate a network in which for either type of nodes one of the two connection rates W_{lk} and V_{lk} ($l=1,2$) is sublinear while another is linear. We assume $W_{1k}=k+\lambda_1$, $W_{2k}=k+\lambda_2$, $V_{1k}=k^{\gamma_1}$, and $V_{2k}=k^{\gamma_2}$ ($0\leqslant\gamma_1,\gamma_2<1$). From Eq. (6) we then obtain the exponential-correction power-law degree distributions

$$n_{1k}\sim k^{-\nu_1'}\exp(C_1k^{\gamma_1-1}),\quad n_{2k}\sim k^{-\nu_2'}\exp(C_2k^{\gamma_2-1}), \tag{12}$$

where $\nu_1'=1+(1+p_1\lambda_1)/p_2$, $\nu_2'=1+(1+p_2\lambda_2)/p_1$, and C_1 and C_2 are two integration constants. Since $\gamma_1-1<0$ and $\gamma_2-1<0$, the exponential correction terms $\exp(C_1k^{\gamma_1-1})$ and $\exp(C_2k^{\gamma_2-1})$ will vanish as $k\to\infty$ and, thus, the degree distributions asymptotically take the power-law forms.

In summary, we have studied a general multicomponent growing network model which combines two processes: (i) introducing a new node of one type and immediately connecting it to an already existing node of any other type, and (ii) creating a new link between two old nodes of different types. By means of the rate equations, we obtained the universal solutions of the degree distributions for the multicomponent GN model with arbitrary linear or sublinear connection rate kernels. We also analyzed in detail the connectivity of the two-component GN systems. For a multicomponent network with all the connection rates being shifted or asymptotically linear, the degree distributions take the pure power-law forms, while the random network with constant connection rate kernels exhibits the exponential degree distributions. An interesting feature of this multicomponent network model is that for the system in which some connection rates are linear while the others sublinear, the degree distribution may have an exponential-correction power-law form and the exponential correction vanishes as $k\to\infty$. On the other hand, by choosing the parameters of the shifted or asymptotically linear kernels, we can construct a two-component GN model which may exhibit the power-law degree distribution in accord with the measurements of the human sexual contact web [4]. Thus, this multicomponent GN model is expected to provide some predictions for the structural properties of some real-world complex systems.

This project was supported by the National Natural Science Foundation of China under Grants Nos. 10305009 and 10275048 and by Zhejiang Provincial Natural Science Foundation of China under Grant No. 102067.

[1] B. Bollobási, *Random Graphs* (Academic Press, London, 1985).
[2] S. Wasserman and K. Faust, *Social Network Analysis* (Cambridge University Press, Cambridge, 1994).
[3] R. Albert, H. Jeong, and A.-L. Barabási, Nature (London) **401**, 130 (1999).
[4] F. Liljeros, C.R. Edling, L.A.N. Amaral, H.E. Stanley, and Y. Åberg, Nature (London) **411**, 907 (2001).
[5] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).
[6] D.J. Watts and S.H. Strogatz, Nature (London) **393**, 440 (1998).
[7] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
[8] M.E.J. Newman, C. Moore, and D.J. Watts, Phys. Rev. Lett. **84**, 3201 (2000).
[9] M.E.J. Newman and D.J. Watts, Phys. Rev. E **60**, 7332 (1999); C. Moore and M.E.J. Newman, *ibid.* **61**, 5678 (2000); M.E.J. Newman, I. Jensen, and R.M. Ziff, *ibid.* **65**, 021904 (2002).
[10] A. Broder, R. Kumar, F. Maghoul, P. Raphavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener, Comput. Netw. **33**, 309 (2000).
[11] S.N. Dorogovtsev, J.F.F. Mendes, and A.N. Samukhin, Phys. Rev. Lett. **85**, 4633 (2000); S.N. Dorogovtsev and J.F.F. Mendes, Phys. Rev. E **62**, 1842 (2000).
[12] P.L. Krapivsky, S. Redner, and F. Leyvraz, Phys. Rev. Lett. **85**, 4629 (2000).
[13] P.L. Krapivsky, G.J. Rodgers, and S. Redner, Phys. Rev. Lett. **86**, 5401 (2001); P.L. Krapivsky and S. Redner, Phys. Rev. E **63**, 066123 (2001).
[14] X. Cheng, H. Wang, and Q. Ouyang, Phys. Rev. E **65**, 066115 (2002).
[15] J.W. Kim, B. Hunt, and E. Ott, Phys. Rev. E **66**, 046115 (2002).
[16] G.J. Lin, X. Cheng, and Q. Ouyang, Chin. Phys. Lett. **20**, 22 (2003).
[17] S.H. Yook, H. Jeong, A.-L. Barabási, and Y. Tu, Phys. Rev. Lett. **86**, 5835 (2001).
[18] P.L. Krapivsky and S. Redner, J. Phys. A **35**, 9517 (2002).
[19] S.N. Dorogovtsev and J.F.F. Mendes, Phys. Rev. E **63**, 056125 (2001).