Degree Distribution of a Two-Component Growing Network

Jianhong Ke[1,2] and Xiaoshuang Chen[2]

[1] College of Physics and Electronic Information, Wenzhou University, Wenzhou 325027, China
kejianhong@yahoo.com.cn
[2] National Laboratory of Infrared Physics, Shanghai Institute of Technical Physics, Chinese Academy of Sciences, Shanghai 200083, China

Abstract. We propose a two-component growing network model which comprises two kinds of nodes. Such a network is constructed by introducing new nodes of either kind with no immediate links and creating new links between any two nodes. We then investigate the connectivity of the two-component growing network by means of the rate equation approach. For a network system with shifted linear connection rate kernels, the in-degree and out-degree distributions take power-law forms; while for a random growing network, the in-degree and out-degree distributions are both exponential. Moreover, the in-degree and out-degree distributions are correlated each other.

Keywords: growing network, connectivity, scale-free, rate equation approach.

1 Introduction

In terms of random graph, a network consists of nodes and possible links between any two nodes, in which nodes denote the individuals while links represent the interactions between two different individuals [1]. And many real-world natural and social systems, such as the World Wide Web [2], Internet [3], metabolic network [4], and the science co-authorship [5, 6], can be abstracted as complex networks. In recent few years, a great deal of effort has been devoted to understanding various aspects of complex networks from topological structure to dynamics (see [7-9] and references therein). Most interestingly, many real-world complex networks are of scale-free and small-world topological structure [10-13]. It is exhibited that in a wide variety of open growing network systems, the degree distribution n_k approaches a power-law form, $n_k \sim k^{-\nu}$ [14-20].

There are many important models contributing to the interpretation of topological properties of complex networks. Starting from a ring lattice, Watts and Strogatz constructed a complex network through random rewiring procedure, which interpolates between regular and random networks and displays small-world properties [11]. Barabási and Albert introduced a famous growing network

J. Zhou (Ed.): Complex 2009, Part II, LNICST 5, pp. 1838–1845, 2009.

model to show the possible origin of scale-free topology of complex networks, which is well-known as the BA model [12]. The standard BA model yields a power-law degree distribution with a fixed exponent $\nu = 3$, while some extended versions of the BA model can "tune" the exponent ν by augmenting the linear connection kernel to the asymptotically linear connection kernel [16-18].

It should also be pointed out that most of these investigations ignored the diversity of nodes and focused only on the single-component networks that comprise a sole kind of nodes. However, many real-world networks are typically composed of distinct nodes and links. For example, in the web of human sexual contacts [10], there are two kinds of individuals, males and females. In a mammalian cell cycle regulatory system [21], there exist a great diversity of enzymes and substrates as nodes [19]. As far as we know, there are only a few works devoting their effort to network models consisting of more than one kind of nodes [19, 20, 22]. Cheng et al. studied an inhomogeneous network consisting of two kinds of nodes denoted by A and B [19], in which a type-B node is only permitted to be linked by a type-A node while a type-A node can be connected either by an A or B node. Kim et al. investigated the topological properties of a growing network model that incorporates multiple species and initial link probabilities [20]. In our previous work, we proposed a multicomponent growing network model which consists of many kinds of nodes as well as links only between the nodes of different kind [22]. It is found that multicomponent networks can also have scale-free topology when the connection rate kernels are linear or shifted linear.

In this work, we construct a two-component growing network on the basis of the network models proposed in [19, 20]. Assume that there are two kinds of nodes denoted by A and B. At each time step, a new link is added to the network in one of three ways: (i) with probability r a new link is created between two already existing nodes of different kind; (ii) with probability q_1 a new link is created between two already existing type-A nodes; (iii) with probability q_2 a new link is created between two already existing type-B nodes. Meanwhile, new nodes are continuously introduced to the system with no immediate links. A new type-A node is added at a rate p_1 and a new type-B node at a rate p_2. Here, we assume that the parameters r, q_l and p_l ($l = 1, 2$) are all fixed constants. While for some situations, the values of the parameters may change with time, and we shall defer it to a future work. We believe that our model may mimic some real-world complex systems such as the above-mentioned human sexual contact web (including homosexual and heterosexual).

The rest of the paper is organized as follows. In Sec. 2, we investigate the two-component growing network model by using the rate-equation approach and then analyze the scale-free properties of the degree distribution. A brief summary is given in Sec. 3.

2 Analytical Solution of Degree Distribution

One can employ the standard probabilistic method or generating function technique to determine the analytical expression of the degree distribution of a growing network model (see, e.g., [1, 23]). Here, we shall investigate the evolution properties of the two-component growing network model by means of the rate equation approach (see, e.g., [16-18]). Let $N_{ij(1)}$ ($N_{ij(2)}$) be the number of type-A (-B) nodes with i in-links to the nodes of the same kind as well as j out-links to the nodes of different kind. The joint-degree distribution $N_{ij(l)}$ ($l = 1, 2$) evolves according to the rate equation

$$\frac{dN_{ij(l)}}{dt} = r\frac{A_l(i;j-1)N_{i,j-1(l)} - A_l(i;j)N_{ij(l)}}{\sum_{l,m} A_l(l;m)N_{lm(l)}} \\ +2q_l\frac{B_l(i-1;j)N_{i-1,j(l)} - B_l(i;j)N_{ij(l)}}{\sum_{l,m} B_l(l;m)N_{lm(l)}} + p_l\delta_{i0}\delta_{j0}. \qquad (1)$$

Here, $A_l(i;j)$ represents the preferential connection rate at which a newly introduced link between two nodes of differen kind is originated from the preexisting node $V_{ij}^{(l)}$ with i in-degrees and j out-degrees, while $B_l(i;j)$ represents the preferential connection rate at which a newly introduced link between two nodes of the same kind is originated from the preexisting node $V_{ij}^{(l)}$ ($V_{ij}^{(1)}$ denotes a type-A node and $V_{ij}^{(2)}$ denotes a type-B node). In equation (1), the first term accounts for the the gain/loss in the number $N_{ij(l)}$ due to the new link (between two nodes of different kind) originated from the nodes $V_{i,j-1}^{(l)}$, while the second term accounts for the gain/loss in the number $N_{ij(l)}$ due to the new link (between two nodes of the same kind) originated from or ended in $V_{i,j-1}^{(l)}$ or $V_{ij}^{(l)}$. It is worth noting that during each connection operation between two nodes of the same kind, one node can be selected as the origin node as well as the target node and thus, the factor 2 must be introduced in the second term. The last term accounts for the continuous introduction of new nodes with no immediate links. It should be noticed that in the context of this paper $l = 1, 2$.

The total number of nodes of the same kind is $M_{0(l)}(t) = \sum_{i,j} N_{ij(l)}(t)$. Summing up equation (1) over all i and j, we obtain the evolution equation of the total number of nodes of either kind, $\dot{M}_{0(l)}(t) = p_l$, which yields the solution $M_{0(l)}(t) = p_l t + M_{0(l)}(0)$. It is obvious that the total number of the nodes of either kind increases at a fixed rate independent of the connection rate kernels ($A_l(i;j)$ and $B_l(i;j)$). Moreover, let $I_l(t)$ and $J_l(t)$ be the total in-degree and the total out-degree of the nodes of the same kind, respectively. Then we have $I_l(t) = \sum_{i,j} iN_{ij(l)}(t)$ and $J_l(t) = \sum_{i,j} jN_{ij(l)}(t)$. Multiplying equation (1) with i and summing them up, one can find that the total in-link of the network evolves as $\dot{I}_l(t) = 2q_l$, with the exact solution $I_l(t) = 2q_l t + I_l(0)$. Multiplying equation (1) with j and summing them up, we deduce $\dot{J}_l(t) = r$, which gives $J_l(t) = rt + J_1(0)$. These show that both the total in-degree and the total out-degree are also independent of the connection rate kernels. Except for the total number of nodes and the total in(out)-degree, the degree distributions as well as their higher moments may be dependent crucially on the connection kernels.

According to the comprehensive results of [16, 17], we can conclude that for a network model with linear or sublinear connection rate kernels, the solution of equation (1) takes the following form at large times:

$$N_{ij(l)}(t) = n_{ij(l)}t, \tag{2}$$

where $n_{ij(1)}$ is independent of time t. Here, $n_{ij(l)}$ is also called the degree distribution. Otherwise, if at least one of the connection rate kernels is superlinear, the so-called "winner takes all" phenomenon [16], a single dominant gel node linked to almost all other nodes, will emerge in our model. Thus, in order to derive the analytical solution of the degree distribution from equation (1), one should know in detail the concrete forms of all the connection rate kernels. In this work, we focus mainly on the growing network with shifted linear connection kernels. Such a network can exhibit the scale-free properties of the degree distribution, which is in agreement with the realistic topology properties of some real-world systems, such as the World wide web.

We consider the connection kernels $A_l(i;j) = j + \lambda_l$ and $B_l(i;j) = i + \mu_l$. In order to ensure that the corresponding connection rates are positive, the parameters λ_l and μ_l are four constants larger than zero. These parameters are also called additional attractiveness (see, e.g., [24]). In the long-time limit, the degree distribution takes the form (2). Substituting equation (2) into equation (1), we recast the governing rate equation into the following recursion equation:

$$\begin{aligned}[1 + 2q_l b_l^{-1}(i+\mu_l) + r a_l^{-1}(j+\lambda_l)]n_{ij(l)} &= 2q_l b_l^{-1}(i+\mu_l-1)n_{i-1,j(l)} \\ &\quad + r a_l^{-1}(j+\lambda_l-1)n_{i,j-1(l)} + p_l\delta_{i0}\delta_{j0}, \end{aligned} \tag{3}$$

where $a_l = r + \lambda_l p_l$ and $b_l = 2q_l + \mu_l p_l$.

We first determine the in-degree and out-degree distributions, $n_{i(l)}^{in} = \sum_{j\geq 0} n_{ij(l)}$ and $n_{j(l)}^{out} = \sum_{i\geq 0} n_{ij(l)}$. Summing up equation (3) over all $j \geq 0$ and $i \geq 0$, respectively, we obtain the recursion equations for the in-degree and out-degree distributions,

$$[i + \mu_l + b_l(2q_l)^{-1}]n_{i(l)}^{in} = (i+\mu_l-1)n_{i-1(l)}^{in} + p_l b_l(2q_l)^{-1}\delta_{i0}, \tag{4}$$

$$(j + \lambda_l + a_l r^{-1})n_{j(l)}^{out} = (j+\lambda_l-1)n_{j-1(l)}^{out} + p_l a_l r^{-1}\delta_{j0}. \tag{5}$$

By solving the recursion equations (4) and (5) we can obtain the in-degree and out-degree distributions as follows:

$$n_{i(l)}^{in} = \frac{c_{1l}\Gamma(i+\mu_l)}{\Gamma[i+1+\mu_l+b_l(2q_l)^{-1}]}, \quad n_{j(l)}^{out} = \frac{c_{2l}\Gamma(j+\lambda_l)}{\Gamma(j+1+\lambda_l+a_l r^{-1})}, \tag{6}$$

where $c_{1l} = p_l b_l \Gamma[1+\mu_l+b_l(2q_l)^{-1}][(b_l+2\mu_l q_l)\Gamma(\mu_l)]^{-1}$ and $c_{2l} = p_l a_l \Gamma(1+\lambda_l+a_l r^{-1})[(a_l+\lambda_l r)\Gamma(\lambda_l)]^{-1}$. For large i and j, we find that both the in-degree and out-degree distributions take the power-law scale-free forms,

$$n_{i(l)}^{in} \sim i^{-2-\mu_l p_l/2q_l}, \quad n_{j(l)}^{out} \sim i^{-2-\lambda_l p_l/r}. \tag{7}$$

The exponents for the in-degree and out-degree distributions can both vary from 2 to ∞ by augmenting the model parameters.

We then discuss the joint-degree distribution. It is of interest to determine some special joint-degree distributions $n_{i0(l)}$ and $n_{0j(l)}$. Obviously, $n_{i0(l)}$ denotes the number of nodes only connecting to nodes of the same kind, while $n_{0j(l)}$ denotes the number of nodes only connecting to nodes of different kind. From the recursion equation (3) we can readily deduce

$$n_{i0(l)} = \frac{c_{3l}\Gamma(i+\mu_l)}{\Gamma(i+1+\mu_l+d_l)}, \quad n_{0j(l)} = \frac{c_{4l}\Gamma(j+\lambda_l)}{\Gamma(j+1+\lambda_l+f_l)}, \tag{8}$$

where $d_l = b_l/(2q_l) + rb_l\lambda_1/(2a_lq_l)$, $f_l = a_l/r + 2q_la_l\mu_l/(b_lr)$, $c_{3l} = a_lb_lp_l\Gamma(1+\mu_l+d_l)[\Gamma(\mu_l)(a_lb_l+2a_lq_l\mu_l+b_lr\lambda_l)]^{-1}$, and $c_{4l} = a_lb_lp_l\Gamma(1+\lambda_l+f_l)[\Gamma(\lambda_l)(a_lb_l+2a_lq_l\mu_l+b_lr\lambda_l)]^{-1}$. For large k, both $n_{i0(l)}$ and $n_{0j(l)}$ take the power-law form

$$n_{i0(l)} \sim i^{-1-d_l}, \quad n_{0j(l)} \sim j^{-1-f_l}. \tag{9}$$

When $2q_la_l = rb_l$, one can deduce the exact expression of the joint-degree distribution. In this special case, the recursion equation (3) can be reduced to

$$\begin{aligned}(i+j+\lambda_l+\mu_l+a_lr^{-1})n_{ij(l)} &= (i+\mu_l-1)n_{i-1,j(l)} + (j+\lambda_l-1)n_{i,j-1(l)} \\ &\quad + p_la_lr^{-1}\delta_{i0}\delta_{j0}.\end{aligned} \tag{10}$$

By introducing the substitution,

$$n_{ij(l)} = \frac{\Gamma(i+\mu_l)\Gamma(j+\lambda_l)}{\Gamma(i+j+1+\lambda_l+\mu_l+a_lr^{-1})} m_{ij(l)}, \tag{11}$$

equation (11) can be further simplified as

$$m_{ij(l)} = m_{i-1,j(l)} + m_{i,j-1(l)} + h_l\delta_{i0}\delta_{j0}, \tag{12}$$

where $h_l = p_la_l\Gamma(1+\lambda_l+\mu_l+a_lr^{-1})[r\Gamma(\lambda_l)\Gamma(\mu_l)]^{-1}$. Equation (12) can be exactly solved by employing the generating function technique. Multiplying equation (12) by x^iy^j and summing them up over all i and j, we find that the generating function

$$G_l(x,y) = \sum_{i,j\geq 0} n_{ij(l)}x^iy^j = \frac{h_l}{1-x-y}. \tag{13}$$

Expanding equation (13) in powers of x and y, we then obtain the exact solution of the joint-degree distribution,

$$n_{ij(l)} = \frac{h_l\Gamma(i+j+1)\Gamma(i+\mu_l)\Gamma(j+\lambda_l)}{\Gamma(i+j+1+\lambda_l+\mu_l+a_lr^{-1})\Gamma(i+1)\Gamma(j+1)}. \tag{14}$$

The results show that $n_{ij(l)} \neq n^{in}_{i(l)} \cdot n^{out}_{j(l)}$. Hence, the in-degree and out-degree distributions for either kind of nodes are correlated each other.

Unfortunately, there is no simple consistent expression for the joint-degree distribution in general cases. By analyzing the recursion equation (3) we can asymptotically obtain

$$n_{ij(l)} \sim \begin{cases} i^{-\theta_l}j^{\lambda_l-1} & for \quad i \gg j \gg 1 \\ j^{-\eta_l}i^{\mu_l-1} & for \quad j \gg i \gg 1, \end{cases} \tag{15}$$

where $\theta_l = 1 + b_l(2q_l)^{-1} + r\lambda_lb_l/(2a_lq_l)$ and $\eta_l = 1 + a_lr^{-1} + 2q_la_l\mu_l/(b_lr)$.

On the other hand, it is also instructive to determine the total degree distribution $P_{k(l)}(t) = \sum_{i+j=k} N_{ij(l)}(t)$, which counts the number of the nodes having the same link number without regard to the objects of their links. Letting $P_{k(l)}(t) = p_{k(l)}t$, we have $p_{k(l)} = \sum_{i+j=k} n_{ij(l)}$. By making use of equation (14) we obtain the asymptotical expression of the total degree distribution at $k \gg 1$,

$$p_{k(l)} \sim k^{-2-\lambda_l p_l/r}, \tag{16}$$

which takes the power-law form. However, it is difficult to deduce the explicit expression of the total degree distribution for general cases, because we cannot derive the exact solution of $n_{ij(l)}$ from equation (3).

In order to verify that the connection rate kernels being shifted linear is the key to bringing the scale-free degree distribution, we here investigate another special case with complete random connection kernels. That is, $A_l(i;j) = B_l(i;j) \equiv 1$ for all i and j. Equation (2) is also valid for this case. Then equation (1) can be rewritten as

$$(p_l + 2q_l + r)n_{ij(l)} = 2q_l n_{i-1,j(l)} + rn_{i,j-1(l)} + p_l^2\delta_{i0}\delta_{j0}. \tag{17}$$

Equation (17) can also be solved by introducing the generating function $H_l(x,y) = \sum_{i\geq 0, j\geq 1} n_{ij(l)}x^i y^j$. Multiplying equation (17) with $x^i y^j$ yields

$$H_l(x,y) = \frac{p_l^2}{p_l + 2q_l + r - 2q_l x - ry}. \tag{18}$$

Expanding equation (18) in powers of x and y, we deduce the exact solution of the joint-degree distribution as follows:

$$n_{ij(l)} = \left(\frac{1}{p_l + 2q_l + r}\right)^{i+j+1} p_l^2 (2q_l)^i r^j \frac{\Gamma(i+j+1)}{\Gamma(i+1)\Gamma(j+1)}. \tag{19}$$

We then turn to determine the in-degree and out-degree distributions. From equation (17) we obtain

$$(p_l + 2q_l)n_{i(l)}^{in} = 2q_l n_{i-1(l)}^{in} + p_l^2\delta_{i0}, \quad (p_l + r)n_{j(l)}^{out} = rn_{j-1(l)}^{out} + p_l^2\delta_{j0}. \tag{20}$$

The recursion relations for in- and out-degrees can be easily solved to give

$$n_{i(l)}^{in} = \frac{p_l^2}{p_l + 2q_l}\left(\frac{2q_l}{p_l + 2q_l}\right)^i, \quad n_{j(l)}^{out} = \frac{p_l^2}{p_l + 2q_l}\left(\frac{r}{p_l + r}\right)^j. \tag{21}$$

For large i and j, the in-degree and out-degree distributions approach the exponential forms as follows:

$$n_{i(l)}^{in} \sim \exp[-k\ln(1 + (2q_l)^{-1}p_l)], \quad n_{j(l)}^{out} \sim \exp[-k\ln(1 + r^{-1}p_l)]. \tag{22}$$

For this random network system, we also have $n_{ij(l)} \neq n_{i(l)} \cdot n_{j(l)}$. Hence, the in-degree and out-degree of nodes are correlated each other.

3 Summary

We have proposed a two-component growing network model, in which new nodes of either kind are continuously added to the system with no immediate links and new links are created between any two nodes. By means of the rate equation approach we have obtained the analytical solutions of the degree distributions for the two-component growing network with shifted linear or complete random connection rate kernels. Obviously, the network has two kinds of links, in-links between nodes of the same kind and out-links between nodes of different kind. We have analyzed in detail the connectivity of the two-component growing network system, including the in-degree, out-degree, and joint-degree distributions.

For a network with all the connection rates being shifted linear, the in-degree and out-degree distributions take the power-law forms. And the joint-degree distribution $n_{ij(l)}$ may take a power-law form in the case of $i \gg j \gg 1$ or $j \gg i \gg 1$. Moreover, the total-degree distribution for either kind of nodes may also take a power-law form in some special cases. For the random network with constant connection rate kernels, the in-degree and out-degree distributions both take the exponential forms. It is also exhibited that the in-degree and out-degree are correlated each other.

On the other hand, by choosing the parameters (such as $q_l = 0$, i.e., the links between any two nodes of the same kind are prohibited) our model can exhibit the power-law degree distribution in accord with the measurements of the human sexual contact web [10]. This two-component network model can be expected to mimic some real-world complex systems, especially, the human sexual contact web in which measurements distinguish homosexual contacts from heterosexual contacts.

Acknowledgement. This project was supported by the National Natural Science Foundation of China under Grant Nos. 10305009 and 10775104.

References

1. Bollobási, B.: Random Graphs. Academic Press, London (1985)
2. Albert, R., Jeong, H., Barabási, A.L.: Diameter of the world-wide web. Nature 401, 130–131 (1999)
3. Pastor-Satorras, R., Vázquez, A., Vespignani, A.: Dynamical and correlation properties of the Internet. Phys. Rev. Lett. 87, 258701 (2001)
4. Jeong, H., Tombor, B., Albert, R., Oltvai, Z.N., Barabási, A.-L.: The large-scale organization of metabolic networks. Nature 407, 651–654 (2000)
5. Newman, M.E.J.: Scientific collaboration networks. I. network construction and fundamental results. Phys. Rev. E 64, 016131 (2001)
6. Newman, M.E.J.: Scientific collaboration networks. II. shortest paths, weighted networks, and centrality. Phys. Rev. E 64, 016132 (2001)
7. Albert, R., Barabási, A.-L.: Statistical mechanics of complex networks. Rev. Mod. Phys. 74, 47–97 (2002)
8. Dorogovtsev, S.N., Mendes, J.F.F.: Evolution of networks. Adv. Phys. 51, 1079–1187 (2002)

9. Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., Hwang, D.-U.: Complex networks: structure and dynamics. Phys. Rep. 424, 175–308 (2006)
10. Liljeros, F., Edling, C.R., Amaral, L.A.N., Stanley, H.E., Åberg, Y.: The web of human sexual contacts. Nature 411, 907–908 (2001)
11. Watts, D.J., Strogatz, S.H.: Collective dynamics of “small-world network”. Nature 393, 440–442 (1998)
12. Barabási, A.-L., Albert, R.: Emergence of scaling in random networks. Science 286, 509–512 (1999)
13. Newman, M.E.J., Moore, C., Watts, D.J.: Mean-field solution of the small-world network model. Phys. Rev. Lett. 84, 3201–3204 (2000)
14. Dorogovtsev, S.N., Mendes, J.F.F., Samukhin, A.N.: Structure of growing networks with preferential linking. Phys. Rev. Lett. 85, 4633–4636 (2000)
15. Dorogovtsev, S.N., Mendes, J.F.F.: Evolution of networks with aging of sites. Phys. Rev. E 62, 1842–1845 (2000)
16. Krapivsky, P.L., Redner, S., Leyvraz, F.: Connectivity of growing random networks. Phys. Rev. Lett. 85, 4629–4632 (2000)
17. Krapivsky, P.L., Rodgers, G.J., Redner, S.: Degree distributions of growing networks. Phys. Rev. Lett. 86, 5401–5404 (2001)
18. Krapivsky, P.L., Redner, S.: Organization of growing random networks. Phys. Rev. E 63, 066123 (2001)
19. Cheng, X., Wang, H., Ouyang, Q.: Scale-free network model of node and connection diversity. Phys. Rev. E 65, 066115 (2002)
20. Kim, J.-W., Hunt, B., Ott, E.: Evolving networks with multispecies nodes and spread in the number of initial links. Phys. Rev. E 66, 046115 (2002)
21. Kohn, K.W.: Molecular interaction map of the mammalian cell cycle control and DNA repair systems. Mol. Biol. Cell 10, 2703–2734 (1999)
22. Ke, J.: Scale-free multicomponent growing networks. Phys. Rev. E 69, 037101 (2004)
23. Krapivsky, P.L., Redner, S.: Finiteness and fluctuations in growing networks. J. Phys. A: Math. Gen. 35, 9517–9534 (2002)
24. Dorogovtsev, S.N., Mendes, J.F.F.: Scaling properties of scale-free evolving networks: continuous approach. Phys. Rev. E 63, 056125 (2001)