Scale-Free Networks with Different Types of Nodes

Juan Zhang and Wenfeng Wu

Department of Applied Mathematics, College of Science,
Donghua University, Shanghai, P.R. China
zhangjuan@dhu.edu.cn

Abstract. In many natural and social networks, nodes may play different roles or have different functions. In this paper, we propose a simple model with different types of nodes and deterministic selective linking rule. We investigate the structural properties by theoretical predictions. It is found that the given model exhibits a power-law distribution. In addition, we make the model become the weighted network by giving the links the weight and analyze the probability distribution of the node strength.

Keywords: power-law distribution, degree distribution, clustering coefficient, weighted network, mean-field method.

1 Introduction

In recent years, complex network systems [1-4] have received remarkable attention, and many scale-free networks have been created because of the pioneer work finished by Barabási and Albert [2], and most of these models are homogeneous since they are composed of the same type of nodes. However, in many real networks, nodes can be divided into different types according to their importance or other properties. For example, in the transports network (the cities and roads being nodes and edges, respectively), the province seats are more important than county ones [5]. For another example, a researcher may work in two or more fields and cooperate with different authors. He writes papers in one field with collaborators who are always engaged in this field while he may also contribute to the other field by participating in another group. He may also have an independent friendship network [6].

Recently, Shi-Jie Yang and Hu Zhao [7] have developed a heterogenous network in which the nodes are catalogued into two types according to the interactions between them, and they found that the node degree exhibits a multi-scaling law distribution with the scaling exponent of each type of nodes. However, they did not consider the clustering coefficient and the average path length of the network. Shou-liang Bu, Bing-Hong Wang and Tao Zhou [5] have gained a scale-free and high clustering complex networks with three types of nodes by large simulations, however, they did not give the theoretic compute.

In this paper, we develop the network with s ($s \geq 1$) types of nodes, and we consider the relations among not only the same types of nodes but also the different types of nodes. That is to say, every type of nodes can be linked with their own type or linked

J. Zhou (Ed.): Complex 2009, Part I, LNICST 4, pp. 991 – 1000, 2009.

with their different types. An evident real system obeying this rule is that people of the different sex can have relations in the network, while people of the same sex can also establish relations. Noticeably, the theoretic calculations of two important topological properties (the degree distribution and the clustering coefficient) are developed in this paper. Lastly, we make the model become the weighted network and compute the probability distribution of node strength.

2 The Model

We consider a growing network as follows:

(i) There are s types of nodes in the network, with fractions $\left(p_1, p_2, p_3, \cdots, p_s\right)$, respectively, here $\sum_{i=1}^{s} p_i = 1$.

(ii) Selective rule: for every step, a new node of the type j with probability $p_j (1 \le j \le s)$ and m new edges are added to the network. The new node is preferentially attached to the same type of existing nodes with probability q and to the different types of existing nodes with probability $1-q$.

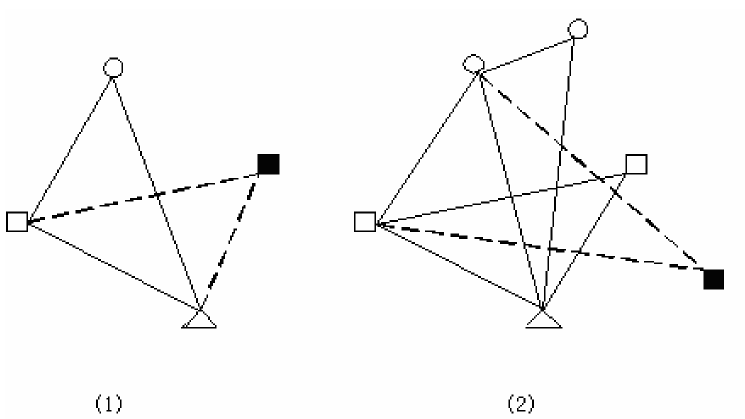

Fig. 1. (1) the network consists of three nodes, denoting as one circle, one triangle and one square, respectively, here; (2) each instant a new node and m new links (dashed line) are added (here, $m = 2, q = 0.5$)

(iii) At each instant, when a new node is generated, we assume that the probability $\prod(k_i)$ with which the new node will be connected to an existing node i depends on the degree k_i of node i:

$$\prod(k_i) = \frac{k_i}{\sum_j k_j}$$

In addition, we assume that the growing network starts with N_0 ($N_0 \geq m$) nodes consisting of different types of nodes, and they link with each other. It is known that after t steps the model leads to a network with $N_t = N_0 + t$ nodes and $E_t = mt + N_0(N_0 - 1)/2$ edges. The average node degree is $\langle k \rangle_t = 2E_t / N_t$ which is approximately equal to $2m$ for infinite network size. Fig 1 is an example when $s = 3, m = 2, q = 0.5$.

3 Analytical Calculation of Relevant Network Parameters

Topology properties are very important to understand the complex dynamics of real-life systems. Here we focus on the important characteristics: degree distribution and clustering coefficient.

3.1 Degree Distribution

The degree distribution is one of the most important statistical characteristics of a network. In this section, we use the mean-field method [8] to analyze the property of our network model, we assume that the degree k_i of node i is continuous, and node i is the existing node in the network.

In every step, when node i is a node of the type j, it increases its degree with the rate becomes:

$$\begin{aligned}\frac{\partial k_i}{\partial t} &= m[qp_j + (1-q)(p_1 + p_2 + \cdots + p_{j-1} + p_{j+1} + \cdots + p_s)]\frac{k_i}{\sum_l k_l} \\ &= m[(2q-1)p_j + 1 - q]\frac{k_i}{\sum_l k_l}\end{aligned} \quad . \qquad (1)$$

Then when node i is an arbitrary node, we have

$$\begin{aligned}\frac{\partial k_i}{\partial t} &= p_1 m[(2q-1)p_1 + 1 - q]\frac{k_i}{\sum_l k_l} + p_2 m[(2q-1)p_2 + 1 - q]\frac{k_i}{\sum_l k_l} \\ &+ p_3 m[(2q-1)p_3 + 1 - q]\frac{k_i}{\sum_l k_l} + \cdots + p_s m[(2q-1)p_s + 1 - q]\frac{k_i}{\sum_l k_l} \\ &= m[(2q-1)(p_1^2 + p_2^2 + p_3^2 + \cdots + p_s^2) + 1 - q]\frac{k_i}{\sum_l k_l}\end{aligned} \qquad (2)$$

Where $\sum_l k_l = 2mt + N_0(N_0 - 1)$, so Eq. (2) becomes

$$\begin{aligned}\frac{\partial k_i}{\partial t} &= m[(2q-1)(p_1^2 + p_2^2 + p_3^2 + \cdots + p_s^2) + 1 - q]\frac{k_i}{2mt + N_0(N_0-1)} \\ &\approx [(2q-1)(p_1^2 + p_2^2 + p_3^2 + \cdots + p_s^2) + 1 - q]\frac{k_i}{2t}\end{aligned} \tag{3}$$

Denoting $g = (2q-1)(p_1^2 + p_2^2 + p_3^2 + \cdots + p_s^2) + 1 - q$, there is the initial condition that node i was added to the system at time t_i with the expected value of connectivity $k_i(t_i) = m$, so the solution of Eq. (3) is:

$$k_i(t) = m\left(\frac{t}{t_i}\right)^{\frac{g}{2}} . \tag{4}$$

Therefore, the probability with a degree $k_i(t)$ smaller than k, $P(k_i(t) < k)$, can be written as

$$P(k_i(t) < k) = P\left(t_i > \left(\frac{m}{k}\right)^{\frac{2}{g}} t\right) = 1 - P\left(t_i \le \left(\frac{m}{k}\right)^{\frac{2}{g}} t\right) . \tag{5}$$

Assuming that we add the nodes at equal time intervals to the system, the probability density of t_i is [8]:

$$P_i(t_i) = \frac{1}{N_0 + t} \approx \frac{1}{t} . \tag{6}$$

Thus, Eq. (5) can be rewritten as

$$P(k_i(t) < k) = 1 - \left(\frac{m}{k}\right)^{\frac{2}{g}} . \tag{7}$$

Then the degree distribution $P(k)$ is obtained:

$$P(k) = \frac{\partial P(k_i(t) < k)}{\partial k} = \frac{\frac{2}{g} m^{\frac{2}{g}}}{k^{\frac{2}{g}+1}} . \tag{8}$$

Obviously, Eq. (8) exhibits the extended power-law form as

$$P(k) \sim k^{-\gamma} . \tag{9}$$

where $\gamma = \frac{2}{g} + 1 = \frac{2}{(2q-1)(p_1^2 + p_2^2 + p_3^2 + \cdots + p_s^2) + 1 - q} + 1$.

3.2 Clustering Coefficient

Clustering coefficient is another important statistic characteristic for a network because it can reflect the clustering extent of the network. Most real-life networks show a cluster structure quantified by it. By definition, clustering coefficient C_i of a node i is $C_i = \frac{2e_i}{k_i(k_i-1)}$, where e_i is the total number of existing edges between all its nearest neighbors and $\frac{k_i(k_i-1)}{2}$ is the number of all possible edges between them.

Using the mean-field rate-equation theory, we can calculate C_i analytically.

When the new node is a node of the type u $(1 \le u \le s)$,

(i) when node i is a node of the type u,

$$\begin{aligned}\frac{\partial e_i}{\partial t} &= \frac{mqk_i}{\sum_l k_l}\sum_{n\in\Omega}[p_u\frac{(m-1)qk_n}{\sum_l k_l} + (1-p_u)\frac{(m-1)(1-q)k_n}{\sum_l k_l}] \\ &= \frac{m(m-1)qk_i}{(\sum_l k_l)^2}[p_u q + (1-p_u)(1-q)]\sum_{n\in\Omega}k_n \\ &= \frac{m(m-1)qk_i}{(\sum_l k_l)^2}[(2q-1)p_u + 1 - q]\sum_{n\in\Omega}k_n\end{aligned} \quad (10)$$

Where k_n denotes the degree of a neighbor of node i, $\sum_{n\in\Omega}k_n$ is the sum of the degrees of all neighbors of node i.

(ii) when node i is a node of the type $j(j=1,2,\cdots,u-1,u+1,\cdots,s)$,

$$\frac{\partial e_i}{\partial t} = \frac{m(m-1)(1-q)k_i}{(\sum_l k_l)^2}[(2q-1)p_u + 1 - q]\sum_{n\in\Omega}k_n \quad . \quad (11)$$

Then when node i is an arbitrary node, we have

$$\begin{aligned}\frac{\partial e_i}{\partial t} &= \frac{m(m-1)k_i}{(\sum_l k_l)^2}\{p_u q[(2q-1)p_u + 1 - q] \\ &\quad + (1-p_u)(1-q)[(2q-1)p_u + 1 - q]\}\sum_{n\in\Omega}k_n \\ &= \frac{m(m-1)k_i}{(\sum_l k_l)^2}[(2q-1)p_u + 1 - q]^2\sum_{n\in\Omega}k_n\end{aligned} \quad . \quad (12)$$

Therefore, when the new node is an arbitrary node, we have

$$\begin{aligned}\frac{\partial e_i}{\partial t} &= \sum_{u=1}^{s}\{p_u \frac{m(m-1)k_i}{(\sum_l k_l)^2}[(2q-1)p_u+1-q]^2\sum_{n\in\Omega}k_n\} \\ &= \frac{m(m-1)k_i}{(\sum_l k_l)^2}[(2q-1)^2\sum_{u=1}^{s}p_u^3+2(3q-2q^2-1)\sum_{u=1}^{s}p_u^2+s(1-q)^2]\sum_{n\in\Omega}k_n\end{aligned} \tag{13}$$

Denoting $a=(2q-1)^2\sum_{u=1}^{s}p_u^3+2(3q-2q^2-1)\sum_{u=1}^{s}p_u^2+s(1-q)^2$.

In addition, for uncorrelated random networks we have [2]

$$\sum_{n\in\Omega}k_n = k_i\frac{\langle k\rangle}{4}\ln t = k_i\frac{m}{2}\ln t \ . \tag{14}$$

So Eq. (13) becomes,

$$\frac{\partial e_i}{\partial t} = \frac{m(m-1)ak_i}{(\sum_l k_l)^2}\sum_{n\in\Omega}k_n = \frac{m(m-1)ak_i}{(\sum_l k_l)^2}k_i\frac{m}{2}\ln t = \frac{m^2(m-1)a}{2}\frac{k_i^{\,2}}{(\sum_l k_l)^2}\ln t \tag{15}$$

Integrating both sides of Eq. (15),

$$\begin{aligned}e_i &= e_{i,0}+\int_1^N\frac{m^2(m-1)a}{2}\frac{k_i^{\,2}}{(\sum_l k_l)^2}\ln t dt = e_{i,0}+\frac{m^2(m-1)a}{2}\int_1^N(\frac{k_i}{\sum_l k_l})^2\ln t dt \\ &= e_{i,0}+\frac{m^2(m-1)a}{2}\int_1^N\frac{1}{m^2g^2}(\frac{dk_i}{dt})^2\ln t dt \\ &= e_{i,0}+\frac{(m-1)a}{2g^2}\int_1^N(\frac{dk_i}{dt})^2\ln t dt \\ &= e_{i,0}+\frac{(m-1)a}{8g^2}\frac{(\ln N)^2}{N}k_i^{\,2}(N)\end{aligned} \tag{16}$$

After neglecting $e_{i,0}$, the clustering coefficient becomes:

$$C_i(k_i) = \frac{e_i}{k_i(k_i-1)/2} \approx \frac{(m-1)a}{4g^2}\frac{(\ln N)^2}{N} \ . \tag{17}$$

Therefore, the clustering coefficient of this network can be obtained:

$$\begin{aligned}C &= \frac{1}{N}\sum_{i=1}^{N}C_i \approx \frac{(m-1)a}{4g^2}\frac{(\ln N)^2}{N} \\ &= \frac{(m-1)[(2q-1)^2\sum_{u=1}^{s}p_u^3+2(3q-2q^2-1)\sum_{u=1}^{s}p_u^2+s(1-q)^2]}{4[(2q-1)\sum_{u=1}^{s}p_u^2+1-q]^2}\frac{(\ln N)^2}{N}\end{aligned} \tag{18}$$

From Eq. (18), C decreases with N according to $\ln^2 N/N$, this result is same to Ref [9].

4 Weighted Network

We propose a weighted network by imposing the essential features of strength preferential attachment on the above model. The rule of weight evolution is based on the notion that ''the rich always gets richer''. In other words, high-weighted link has higher probability to evolve. The weight of the links from the same types of nodes are set to w_1, while The weight of the links from the different types of nodes are set to w_2. At each instant, when a new node is generated, these links are randomly connected to the existing nodes according to the strength preferential probability Λ_i, which is defined as

$$\Lambda_i = \frac{s_i}{\sum_j s_j} \quad . \tag{19}$$

By mean field approximation and by treating all discrete variables as continuous [10-11], we assume that node i is a node, which is added at time t_i.

(1) when the new node and node i have the same type, the strength s_i of the node added at time t_i satisfies at time t,

$$\frac{ds_i(t)}{dt} = mqw_1 \frac{s_i(t)}{\sum_l s_l(t)} \quad . \tag{20}$$

(2) when the new node and node i have the different types, the strength s_i of the node added at time t_i satisfies at time t,

$$\frac{ds_i(t)}{dt} = m(1-q)w_2 \frac{s_i(t)}{\sum_l s_l(t)} \quad . \tag{21}$$

When the new node is a node of the type j $(1 \le j \le s)$,

$$\frac{ds_i(t)}{dt} = p_j mqw_1 \frac{s_i(t)}{\sum_l s_l(t)} + (1-p_j)m(1-q)w_2 \frac{s_i(t)}{\sum_l s_l(t)} \quad . \tag{22}$$

Then when node i is an arbitrary node, we have

$$\begin{aligned}\frac{ds_i(t)}{dt} &= \sum_{j=1}^{s}[p_j{}^2 mqw_1 \frac{s_i(t)}{\sum_l s_l(t)} + p_j(1-p_j)m(1-q)w_2 \frac{s_i(t)}{\sum_l s_l(t)}] \\ &= \frac{s_i(t)}{\sum_l s_l(t)} \sum_{j=1}^{s}[p_j{}^2 mqw_1 + p_j(1-p_j)m(1-q)w_2]\end{aligned} \tag{23}$$

Since one node is added to the network in each time step, the network size $N_t \approx t$. Then we have

$$\sum_l s_l(t) \approx 2mqw_1 t + 2m(1-q)w_2 t \quad . \tag{24}$$

Putting Eq. (24) into Eq. (23) and using the initial condition $s_i(t=t_i) = mqw_1 + m(1-q)w_2$, we conclude that

$$\frac{ds_i(t)}{dt} = \frac{s_i(t)}{t} \frac{\sum_{j=1}^{s}[p_j{}^2 mqw_1 + p_j(1-p_j)m(1-q)w_2]}{2mqw_1 + 2m(1-q)w_2} \quad . \tag{25}$$

We assume that $n = \frac{\sum_{j=1}^{s}[p_j{}^2 mqw_1 + p_j(1-p_j)m(1-q)w_2]}{2mqw_1 + 2m(1-q)w_2}$, then

$$s_i(t) \approx [mqw_1 + m(1-q)w_2](\frac{t}{t_i})^n \quad . \tag{26}$$

The probability distribution of node strength can be computed by

$$P(s,t) = -\frac{1}{t}\frac{\partial t_i}{\partial s_i(t)}\bigg|_{s_i(t)=s(t)} \quad . \tag{27}$$

From Eq. (26) we obtain:

$$P(s,t) = s^{-\frac{n+1}{n}} \frac{[mqw_1 + m(1-q)w_2]^{\frac{1}{n}}}{n} . \tag{28}$$

Then $P(s,t) \sim s^{-r_s}$

$$r_s = \frac{n+1}{n} = \frac{\dfrac{\sum_{j=1}^{s}[p_j^{\ 2}mqw_1 + p_j(1-p_j)m(1-q)w_2]}{2mqw_1 + 2m(1-q)w_2} + 1}{\dfrac{\sum_{j=1}^{s}[p_j^{\ 2}mqw_1 + p_j(1-p_j)m(1-q)w_2]}{2mqw_1 + 2m(1-q)w_2}} \quad . \tag{29}$$

$$= \frac{2qw_1 + 2(1-q)w_2 + \sum_{j=1}^{s}[p_j^{\ 2}qw_1 + p_j(1-p_j)(1-q)w_2]}{\sum_{j=1}^{s}[p_j^{\ 2}qw_1 + p_j(1-p_j)(1-q)w_2]}$$

Obviously this weighted model successfully reproduces the scale-free behavior of the probability distributions of strength with a tunable exponent.

5 Conclusions

In conclusion, we have presented a simple model that consists of different types of nodes. One finds that this model is a scale-free network, and the analytical solution of clustering coefficient is obtained. In addition, we have proposed a simple weighted model with different types of nodes. The weighted model successfully reproduces the scale-free behavior of the probability distributions of strength with a tunable exponent that depends on the microscopic mechanism ruling the weight evolution.

References

1. Stoats, S.H., Tao Zhou, M.S.: Exploring Complex Networks. J. Nature (London) 410, 268 (2001)
2. Albert, R., Barabási, A.L., Tao Zhou, M.S.: Statistical mechanics of complex networks. J. Rev. Mod. Phys. 74, 47 (2002)
3. Dorogovtsev, S.N., Mendes, J.F.F., Tao Zhou, M.S.: Evolution of networks. J. Adv. Phys. 51, 1079 (2002)
4. Newman, M.E.J., Tao Zhou, M.S.: The structure and function of complex networks. J. SIAM Rev. 45, 167 (2003)
5. Bu, S.-l., Wang, B.-H., Tao Zhou, M.S.: Gaining scale-free and high clustering complex networks. J. Physica A 374, 864–868 (2007)
6. Palla, G., Derenyi, I., Farakas, I., Vicsek, T., Tao Zhou, M.S.: Uncovering the Overlapping Community Structure of Complex Networks in Nature and Society. J. Nature 435, 814 (2005)

7. Yang, S.-J., Zhao, H., Tao Zhou, M.S.: Generating multi-scaling networks with two types of vertices. J. Physica A. 370, 863–868 (2006)
8. Albert, R., Barabási, A.L., Jeong, H., Tao Zhou, M.S.: Mean-field theory for scale-free random networks. J. Physica A 272, 173–187 (1999)
9. Ben-Naim, E., Frauenfelder, H., Toroczkai, Z., Tao Zhou, M.S.: Complex Networks. Lecture Notes in Physics, vol. 650. Springer, Heidelberg (2004)
10. Leung, C.C., Chau, H.F., Tao Zhou, M.S.: Weighted assortative and disassortative networks model. J. Physica A 378, 591–602 (2007)
11. Masucci, A.P., Rodgers, G.J., Tao Zhou, M.S.: Multi-directed Eulerian growing networks. J. Physica A 386, 557–563 (2007)