Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Physica A 335 (2004) 240–248
PHYSICA A
www.elsevier.com/locate/physa

The modeling of scale-free networks☆

Qinghua Chen[a,b,*], Dinghua Shi[a]

[a] Department of Mathematics, College of Science, Shanghai University, Shanghai 200436, China
[b] College of Mathematics and Computer Science, Fujian Normal University, Fuzhou 350007, China

Received 2 September 2003

Abstract

In order to explore further the mechanism responsible for scale-free networks, we introduce two extended models of the BA model. The model A, where the system incorporates the addition of new links between existing nodes, a new node with new links and the rewiring of some links at every time step, all sites are born with some initial attractiveness. We calculate analytically the degree distribution. The system self-organizes into a scale-free network, the scaling exponent $\gamma > 2$. The model B is a new model; we consider that some old links are deleted with the anti-preferential probability. The result indicates that the system evolves itself into a scale-free network, the scaling exponent γ varies from 2 to 3.

PACS: 84.35.+i; 64.60.Fr; 87.23.Ge

Keywords: Scale-free networks; Degree distribution; Scaling exponent

1. Introduction

Complex networks describe a wide range of systems in nature and society, whose nodes are the elements of the system and edges represent the interactions between them. For example, World Wide Web (WWW); internet, genetic networks, ecological networks, citation networks, movie actor collaboration network, etc. Life systems form a huge genetic network, whose vertices are proteins and genes, the links represent the chemical interactions between them. Similarly, the WWW represents the largest network, whose nodes are HTML documents (web pages) and the edges are the hyperlinks

☆ This research is supported by the National Natural Science Foundation of China (Grant No. 70171059).

* Corresponding author. Department of Mathematics, College of Science, Shanghai University, Shanghai 200436, China.

E-mail addresses: qhdchen@yahoo.com.cn (Q. Chen), shidh2001@263.net (D. Shi).

0378-4371/$ - see front matter
doi:10.1016/j.physa.2003.12.014

(URL'S) that point from one document to another. The great challenge today is the accurate and complete description of complex systems. The most basic issues are the statistical mechanics of complex networks. Researchers are interested to unravel the structure and dynamics of complex networks [1–5].

The degree of a node in networks is the number of its edges. The probability $P(k)$ that a node in the network is connected to k other nodes $(k = 0, 1, 2 \ldots,)$ is called the degree distribution or connectivity distribution. The degree distribution is a simple but very important characteristic of networks.

Traditionally, random networks was described with the random graph theory of Erdös and Rényi (ER model) [6,7]. The degree distribution of such random graph is a binomial distribution, it can be approximated by a Poisson distribution for large N. Recently Watts and Strogatz (WS) [8] have introduced a small-world network. The average path lengths between nodes in this network is surprisingly small, leading to a small-world phenomenon. The shape of the degree distribution in the WS model is similar to the Poisson distribution of a random graph (ER model).

A common feature of the ER and WS models is that the degree distribution $P(k)$ decays exponentially for large k, displays an exponential tail. Such networks are called the exponential networks. Thus the topology of the network is relatively homogeneous, all nodes have approximately the same number of edges. Consequently, nodes with large connectivity are practically absent [1,9].

Barabási and Albert (1999) introduced first scale-free networks in their seminal works [3,9,10]. They explored several large databases describing the topology of the large networks, including the WWW, the actor collaboration network and the citation network, etc. The empirical results showed that the degree distribution $P(k)$ in these networks decays as a power law, follows $P(k) \sim k^{-\gamma}$ for large k. The exponent γ is scattered between 2.1 and 3. These results offered first the evidence that some large networks can self-organize into a scale-free state. Such networks are called scale-free (SF) networks. This is a significant discovery in complex networks. The power-law tailed degree distribution is remarkably different from the Poisson distribution. It is a fat-tailed distribution. SF networks are inhomogeneous, leading over time to some vertices that are highly connected, a "rich-get-richer" phenomenon that can be easily detected in real networks.

What is the mechanism responsible for SF networks? Barabási and Albert [9–11] suggested that growth and preferential attachment play important roles in the network evolution, the SF nature of real networks is rooted in these two generic mechanisms. In fact, most real-world networks grow by the continuous addition of new nodes, and exhibit the preferential attachment. These two ingredients, growth and preferential attachment, inspired the introduction of the Barabási–Albert model.

1.1. The BA model

The algorithm of the BA model is the following:

(1) *Growth*: Starting with a small number (m_0) of nodes, at every time step, we add a new node with $m(\leqslant m_0)$ edges that link the new node to m different nodes already present in the system.

(2) *Preferential attachment*: When choosing the nodes to which the new node connects, we assume that the probability Π that a new node will be connected to node i depends on the degree k_i of node i, such that

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j} . \tag{1}$$

Barabási et al. [9,10] developed a continuum theory to calculate analytically the degree distribution $P(k)$. The theory can address the dynamic properties of SF networks. Assuming that k is continuous, and thus the probability $\Pi(k_i)$ can be interpreted as a rate of continuous change of $k_i(t)$. Consequently, $k_i(t)$ satisfies the dynamic equation:

$$\frac{\partial k_i}{\partial t} = m\Pi(k_i) = m\frac{k_i}{\sum_j k_j} = \frac{k_i}{2t} \tag{2}$$

with the initial condition that node i was added to the system at time t_i with connectivity $k_i(t_i) = m$.

The theory predicted two major results:

(1) First, the degree k_i of a node i depends on time t as

$$k_i(t) = m\left(\frac{t}{t_i}\right)^{\beta}, \quad \beta = \frac{1}{2} , \tag{3}$$

where the exponent β is called the dynamic exponent. Eq. (3) indicates that the degree of all nodes evolves in the same way, following a power law.

(2) Second, the probability density for $P(k)$ (i.e., the degree distribution) follows

$$P(k) = \frac{\partial P(k_i(t) < k)}{\partial k} = \frac{2m^2 t}{m_0 + t}\,\frac{1}{k^3} \sim 2m^2 k^{-\gamma}, \quad \gamma = 3 , \tag{4}$$

where the exponent γ is called the scaling exponent, predicting $\gamma = 3$, independent of the parameter m. Eq. (4) indicates that despite its continuous growth, the network self-organizes into a stationary SF state.

All these results are agreement with the numerical simulations. The BA model offers first the successful mechanism accounting for the SF nature of real networks.

The BA model generates a SF network with a fixed scaling exponent 3, while the exponents measured for real networks vary between 1.05 and 4 [1,9]. How can we change the scaling exponent?

Albert and Barabási [11] introduced an extended model of the network evolution that gives a more realistic description of the local processes. This model incorporates the linked new edges between existing nodes, the rewiring of links and the addition of new nodes with new edges. At each time step, one of the above three operations is performed with respective probabilities. They calculated analytically the degree distribution $P(k)$ by using the continuum approach. The result shows that the networks can evolve two different topologies. In the first regime, $P(k)$ has a power-law tail, but the scaling exponent γ depends on the parameters of the model, a range of γ between 2 and 3. In the second regime, however, the numerical simulations indicate that $P(k)$ approaches an exponential.

The goal of this paper is to explore further the mechanism responsible for SF networks. We introduce the two models of evolving networks that give more realistic descriptions of the local processes. In the model A, we incorporate the three local processes at every time step: the addition of new links between old nodes, a new node with new links and the rewiring some existing links. All sites are born with some initial attractiveness. We show that the system evolves itself into a SF network, the scaling exponent γ is greater than 2, its super boundary (greater than 3) depends on the parameters. In model B, we perform three operations at every time step: the addition of a new node with new edges, new edges between old nodes and some old links are deleted. The two ends of the new edges between old nodes are chosen all with preferential attachment (different from the model A). A node is selected as a end of the deleted line, with the anti-preferential probability, it is more reasonable for deleting links. We show that this system self-organizes into a SF network, the scaling exponent γ varies from 2 to 3.

2. The two models for scale-free networks

The BA model is a minimal model that captures the mechanism responsible for the scale-free networks. However, there are discrepancies between the BA model and real networks. The modeling of SF networks will put the emphasis on capturing the network dynamics. Some microscopically evolving processes may influence the topology of networks. Here, we introduce the two models that give more realistic descriptions of the local processes in evolving networks.

In some real networks, a series of microscopic events may shape the evolution of the networks, including the addition of new nodes, new links between old nodes and the rewiring of some links. Consequently, we introduce the following model.

2.1. The model A

Starting with m_0 isolated nodes, and at each time step we perform the following three operations:

(1) We add l new links between existing vertices: select randomly a node as the starting point of the new link, while a node i is selected as the other end of the new link, with the preferential probability

$$\Pi(k_i,\alpha) = \frac{k_i + \alpha}{\sum_j (k_j + \alpha)} . \tag{5}$$

This operation is repeated l times, where all nodes are born with some initial attractiveness $\alpha \geqslant 0$. The introduced parameter α governs the probability for "young" sites to get new links.

(2) We add a new node with m new links: a new node is connected to node i already present in the system, with the probability $\Pi(k_i,\alpha)$ given by (5).

(3) We rewire n links that have existed in the network: select randomly a node i and a link l_{ij} connected to it. Next we rewire this link and replace it with a new link

$l_{i'j}$ that connects node j and node i' chosen with the probability $\Pi(k_{i'}, \alpha)$ given by (5). This operation is repeated n times.

The system parameters satisfy the condition: $l, m, n; \alpha \geqslant 0$, and $l + m > 0$.

According to the continuum theory, we assume that k_i changes continuously, and the probability $\Pi(k_i, \alpha)$ can be interpreted as the rate at which k_i changes. Consequently, the processes (1)–(3) all contribute to k_i, each being explained as follows:

(1) The addition of l new edges between the old nodes:

$$\left(\frac{\partial k_i}{\partial t}\right)_{(1)} = l\,\frac{1}{N} + l\left(1 - \frac{1}{N}\right)\frac{k_i + \alpha}{\sum_j (k_j + \alpha)}\,, \tag{6}$$

where N is the size of the system. The first term on the right-hand side corresponds to the random selection of one end of the new link, while the second term reflects the preferential attachment used to select the other end of the link.

(2) The addition of a new node with m new edges:

$$\left(\frac{\partial k_i}{\partial t}\right)_{(2)} = m\frac{k_i + \alpha}{\sum_j (k_j + \alpha)}\,. \tag{7}$$

The term on the right-hand side represents the increasing degree of the node that is connected to the new node.

(3) The rewiring of n links:

$$\left(\frac{\partial k_i}{\partial t}\right)_{(3)} = -n\,\frac{1}{N} + n\left(1 - \frac{1}{N}\right)\frac{k_i + \alpha}{\sum_j (k_j + \alpha)}\,. \tag{8}$$

The first term incorporates the decreasing degree of the node from which the link was rewired, and the second term represents the increasing degree of the node that the link is reconnected to. The total degree does not change during the rewiring process, where the system size N and the total number of degrees $\sum_j k_j$ vary with the time t as: $N(t) = m_0 + t \approx t$, $\sum_j k_j = 2(l + m)t$ and $\sum_j (k_j + \alpha) = [2(l + m) + \alpha]t$.

By adding the contributions of the three processes, we obtain the following dynamical equation:

$$\begin{aligned}
\frac{\partial k_i}{\partial t} &= (l - n)\frac{1}{N} + (l + m + n)\frac{k_i + \alpha}{\sum_j (k_j + \alpha)} - (l + n)\frac{1}{N}\,\frac{k_i + \alpha}{\sum_j (k_j + \alpha)} \\
&= \frac{l - n}{m_0 + t} + \frac{l + m + n}{2l + 2m + \alpha}\,\frac{k_i + \alpha}{t} - \frac{l + n}{2l + 2m + \alpha}\,\frac{k_i + \alpha}{t(m_0 + t)} \\
&\approx \frac{l - n}{t} + \frac{l + m + n}{2l + 2m + \alpha}\,\frac{k_i + \alpha}{t}, \quad \text{for large } t
\end{aligned} \tag{9}$$

with the initial condition that node i was added to the system at time t_i with the degree $k_i(t_i) = m$.

The solution of Eq. (9) has the form

$$k_i(t) = A\left(\frac{t}{t_i}\right)^{\beta} - A + m\,, \tag{10}$$

where the dynamic exponent

$$\beta = \beta(l,m,n;\alpha) = \frac{l+m+n}{2l+2m+\alpha}\,, \tag{11}$$

the coefficient

$$A = A(l,m,n;\alpha) = \frac{(2l+2m+\alpha)(l-n)+(l+m+n)(m+\alpha)}{l+m+n}\,. \tag{12}$$

The condition $A > 0$ holds if and only if $n < l+m+\alpha$.

When $n < l+m+\alpha$, i.e., $A > 0$, we can obtain from (10):

$$P(k_i(t) < k) = P(t_i > c(k)t)\,.$$

Here,

$$0 < c(k) = \left(\frac{A}{A-m+k}\right)^{1/\beta} < 1, \quad \text{for } k > m\,.$$

Assuming that we add the nodes at equal time intervals to the system, i.e., t_i follows the uniform distribution over interval $(0, m_0+t)$. Hence,

$$P(k_i(t) < k) = 1 - \left(\frac{A}{A-m+k}\right)^{1/\beta}\frac{t}{m_0+t}\,,$$

$$\begin{aligned} P(k) &= \frac{\partial P(k_i(t)<k)}{\partial k} \\ &= \frac{t}{m_0+t}\,\frac{1}{\beta}A^{1/\beta}(k+A-m)^{-\gamma} \\ &\to \frac{1}{\beta}A^{1/\beta}(k+A-m)^{-\gamma},\ (t\to\infty),\ \gamma = 1+\frac{1}{\beta}\,. \end{aligned} \tag{13}$$

The system self-organizes into a SF network, with the scaling exponent γ:

$$2 < \gamma = \frac{3l+3m+n+\alpha}{l+m+n} \leqslant 3 + \frac{\alpha}{l+m+n}\,. \tag{14}$$

This allows us to account for the wide variations seen in real networks, for which γ varies from 2 to 3, or $\gamma > 3$.

We consider the following two particular cases:

(a) If $\alpha = 0, n = 0, l > 0, m > 0$, then $\beta = 0.5, \gamma = 3, A = 2l+m$. This result is same as one in the BA model. It indicates that the addition of new edges between the old nodes at every time step does not change the scaling exponent in the system, but the coefficient of the power-law degree distribution becomes greater.

(b) If $\alpha = 0, m = 0, n = 0, l > 0$, then $\beta = 0.5, \gamma = 3, A = 2l$. The scaling exponent is also same as one in the BA model. We notice where new nodes have no new links, new links are connected only between old nodes.

In the model A, selecting randomly a node as the starting point of the new link, while the other end of the link is selected with the preferential attachment. If the two ends of the new link are all selected with the preferential probability, how does the network evolve? Meanwhile, in some real networks (e.g. WWW), some old links may be deleted, we consider the following model:

2.2. The model B

Starting with m_0 isolated nodes, and at each time step the following three processes are performed:

(1) a new node is added to the system: the new node with $m(\leqslant m_0)$ new edges that are connected to m different nodes. The preferential probability $\Pi(k_i)$ that a new node will be connected to node i is given by (1)
(2) n new edges between old nodes are produced: a node i is selected as a end of a new edge, with the preferential probability $\Pi(k_i)$.
(3) c old links are deleted: we select a node i as a end of a deleted link with the anti-preferential probability:

$$\Pi^*(k_i) = \frac{1}{N(t)-1}\left(1-\Pi(k_i)\right), \tag{15}$$

where $N(t)$ is the size of the system, $(N(t)-1)^{-1}$ is the normalized coefficient for probability, such that $\sum_i \Pi^*(k_i)=1$. The anti-preferential probability is more reasonable for deleting links, it is consistent with the real networks (because some real networks may exhibit anti-preferential deletion), where $m>0,\ n\geqslant 0,\ c\geqslant 0,\ m+n>c$.

By the continuum theory, $k_i(t)$ satisfies the following dynamical equation:

$$\begin{aligned}
\frac{\partial k_i}{\partial t} &= m\Pi(k_i) + n\left[\Pi(k_i)\times 1 + \sum_{j\neq i}\Pi(k_j)\Pi(k_i)\right] \\
&\quad - c\left[\Pi^*(k_i)\times 1 + \sum_{j\neq i}\Pi^*(k_j)\Pi^*(k_i)\right] \\
&\approx m\Pi(k_i) + n[2\Pi(k_i) - (\Pi(k_i))^2] \\
&\quad - c\left[\frac{2(1-\Pi(k_i))}{t} - \frac{(1-\Pi(k_i))^2}{t^2}\right] \\
&\approx \frac{(m+2n)k_i}{2(m+n-c)t} - \frac{nk_i^2}{4(m+n-c)^2t^2} - \frac{2c}{t} + \frac{ck_i}{(m+n-c)t^2} \\
&\approx -\frac{2c}{t} + \frac{m+2n}{2(m+n-c)t}k_i, \quad \text{for large } t \\
&\qquad (\text{because } k_i(t) \propto t^{\beta}, \beta < 1)
\end{aligned} \tag{16}$$

with the initial condition that node i at its introduction has $k_i(t_i) = m$, where $N(t) - 1 = m_0 + t - 1 \approx t$, $\sum_j k_j = 2(m + n - c)t$.

Eq. (16) has the following solution:

$$k_i(t) = B\left(\frac{t}{t_i}\right)^{\beta} - B + m, \quad \text{for large } t\,, \tag{17}$$

where the dynamic exponent

$$\beta = \beta(m, n, c) = \frac{m + 2n}{2(m + n - c)}\,, \tag{18}$$

the coefficient

$$B = B(m, n, c) = m - \frac{4c(m + n - c)}{m + 2n}\,. \tag{19}$$

The condition $B > 0$ holds if and only if $c < m/2$ or $c > (m+2n)/2$. If $c > (m+2n)/2$, then $\beta > 1$, it is impossible because $k_i(t) \leqslant t$ for large t. Therefore, we consider only the case when $c < m/2$.

Because Eq. (17) is similar to Eq. (10), we can obtain that the degree distribution

$$P(k) = \frac{t}{m_0 + t}\frac{1}{\beta}B^{1/\beta}(k + B - m)^{-\gamma}$$
$$\to \frac{1}{\beta}B^{1/\beta}(k + B - m)^{-\gamma},\ (t \to \infty),\ \gamma = 1 + \frac{1}{\beta}\,. \tag{20}$$

This system self-organizes into a SF network, with the scaling exponent γ:

$$2 < \gamma = \frac{3m + 4n - 2c}{m + 2n} \leqslant 3\,. \tag{21}$$

The scaling exponent γ varies from 2 to 3. The exponent γ is a increasing function for m, but γ is a decreasing function for n and c. This is a interesting phenomenon.

3. Conclusions

In this paper, we have explored further the mechanisms responsible for SF networks. Growth and preferential attachment are mechanisms common to a number of complex systems. We have introduced the two models that give more realistic descriptions of the local processes than the BA model. In the model A, incorporating the addition of new nodes, new links, and the rewiring of some links, all nodes are born with some initial attractiveness. We have calculated analytically the degree distribution by the continuum theory, the system evolves itself into a SF network, the scaling exponent γ depends on the parameters and is greater than 2. In particular, (a) In the BA model, the addition of new links between old nodes does not change the scaling exponent in the system, (b) In the BA model, a new node has no new links, but new links are connected between old nodes, the scaling exponent is same as one in the BA model. In the model B, we perform the operations: the addition of new nodes, new links, and the deletion of some links. A link is deleted with the anti-preferential probability. We think that the anti-preferential probability is more reasonable for deleting links. We have shown that the system self-organizes into a SF network, the scaling exponent γ varies from 2 to 3.

If a new node is added to the system at every time step, then the degree $k_i(t) \leqslant t$ (at time t). Consequently, the dynamical exponent $\beta \leqslant 1$, and the scaling exponent $\gamma = 1 + (1/\beta) \geqslant 2$. However, there are real networks, their scaling exponents are less than 2 (See Ref. [1]). How do we introduce a model with the scaling exponent $\gamma < 2$?

References

[1] R. Albert, A.-L. Barabási, Statistical mechanics of complex networks, Rev. Mod. Phys. 74 (2002) 47–97.
[2] S.H. Strogatz, Exploring complex networks, Nature 410 (2001) 268–276.
[3] R. Albert, H. Jeong, A.-L. Barabási, Diameter of the world-wide web, Nature 401 (1999) 130.
[4] H. Jeong, S.P. Mason, A.-L. Barabási, Z.N. Oltvai, Lethality and centrality in protein networks, Nature 411 (2001) 41.
[5] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, A.-L. Barabási, The large-scale organization of metabolic networks, Nature 407 (2000) 651–654.
[6] P. Erdös, A. Rényi, On the evolution of random graphs, Publ. Math. Inst. Hung. Acad. Sci. 5 (1960) 17–61.
[7] B. Bollobás, Random Graphs, Academic Press, London, 1985.
[8] D.J. Watts, S.H. Strogatz, Collective dynamics of small-world networks, Nature 393 (1998) 440–442.
[9] A.-L. Barabási, R. Albert, Emergence of scaling in random networks, Science 286 (1999) 509–512.
[10] A.-L. Barabási, R. Albert, H. Jeong, Mean-field theory for scale-free random networks, Physica A 272 (1999) 173–187.
[11] R. Albert, A.-L. Barabási, Topology of evolving networks: local events and universality, Phys. Rev. Lett. 85 (2000) 5234–5237.