ISSN 1746-7659, England, UK
Journal of Information and Computing Science
Vol. 13, No. 2, 2018, *pp*.125-130

A Scale-Free Network Evolution Model Based on the Growth Characteristics of Social Networks

Qi Yang[1+], Yulong Shi[1]
[1] *Department of Mathematics, Jinan University, Guangzhou, 510632, China*
[+] *E-mail: jnuyangqi@163.com*

(Received February 13 2018, accepted March 28 2018)

Abstract. In this paper, based on the classic BA scale-free network model, we proposed a new evolution model that gives a more realistic description of the people's behavior on social networks. In the process of growth, there are local preferential attachment mechanisms and random attachment or removal between the old and new edges. We proved that the extended model follows the power-law distribution and the power exponent is between 2 and 3, which provides a theoretical support for analyzing the similar social network. Compared with the classic BA model, the extended model has a smaller average path length and a larger clustering coefficient, which is more consistent with the real social network.

Keywords: BA scale-free network, social network, local preferential attachment, random attachment

1. Introduction

As an important tool to study the complexity problem, complex networks have aroused research upsurge in recent years. A amount of complex networks exist in the real world, such as aeronautical networks, biological networks, social networks and so on. It is found that more and more real networks follow the power-law distribution, called scale-free network [1]. The BA scale- free model [2] focuses on characterizing the power-law distribution of actual networks. In order to be more in line with the logic of real network evolution, it is of great theoretical significance and application value to extend the basic BA scale-free network model. Albert and Barabasi [3] proposed an extended model (EBA model) of network evolution that has more practical significance in the study of local processes. Bianconi and Barabasi [4] assigned a fitness parameter to each node and defined the fitness model. The fitness model have such characteristics of 'first- mover-wins, fit-gets-richer and winner-takes-all'.

There are some other extension models. Watts and Strogatz [5] explored a small-world model, which has short-path, high clustering features and satisfies the characteristics of the small world networks [6-7] that mimic the evolution of a social network process. Barabasi and Albert [8-9] studied the World Wide Web (WWW) and proposed a BA scale-free model based on growth and prioritization. Li, Jin and Chen [10] studied complexity and synchronization of the World Trade Web (WTW), and investigated some scale-free features of the WTW. In [11] , based on the new concept of local-world connectivity, Li and Chen proposed a local-world evolving network model. In the last few years, Wang, Xu and Pang studied the internal structure of online social networks and combined the inside growth, outside growth, and edge replacement base on those of complex networks, then proposed an evolution model in [12].

The model we proposed here is grounded on a modification of the model presented by Barabasi (BA model) [2]. The mathematical definition of BA model:

- Growth: start with a network of n_0 nodes. A new node is added at each timestep with m $(\le n_0)$ edges that connect the new node to m existing nodes.
- Preferential attachment: the probability Π_i of the new node connect the existing node i depends on the degree k_i of node i as

$$\Pi_i = \frac{k_i}{\sum_{j=1}^{n} k_j}, \tag{1}$$

where n is total number of nodes.

2. Model description

Base on BA scale-free network, our algorithm is defined as follows:

Initialization. Start with a network of n_0 nodes. Initial $n_1, n_2, m_1, m_2, m_3, m_4$.

Step 1. Growth: n_1 nodes are added to the network and each new node connected to m_1 existing nodes by preferential attachment probability Π_i (as defined in (1)).

Step 2. Preferential attachment: n_2 new nodes are added simultaneously. And each new node connected to a random existing node, denoted by j. Add $m_2 - 1$ edges between the new node and the neighbor nodes of node j. The edges are selected with probability:

$$\Pi_i^{'} = \frac{k_i}{\sum_{s \in N(j)} k_s}, i \in N(j), \tag{2}$$

where $N(j)$ is the set of neighbor nodes of node j.

Step 3. Aggregation: break m_3 edges randomly, then add m_4 $(m_4 > m_3)$ edges that selected with equal possible probability.

Output. Repeat step 1 to 3 t times. The network has $N(t) = n_0 + (n_1 + n_2)t$ nodes and $M(t) = (n_1 m_1 + n_2 m_2 + m_4 - m_3)t$ edges.

In the above algorithm, as a widespread social network, Step 2 shows that a node recommended by a friend to the network is not only a friend of the recommender, but also a friend of his friend. This is in line with our natural situation. Another point is the preferential attachment in this step is only for the range of friends around the recommender, it is the local world of the node. The purpose of Step 3 is to make the network aggregate and satisfy the characteristics of the small world networks.

3. Main result

In this section, we prove that the extended network model follows the power-law distribution which is the property of scale-free networks. And we obtain some statistical properties of the extended model. We present some numerical results to performance of the extended model that is better than BA model. Degree distribution

Let $k_i(t)$ denote the degree of node i at time step t. Node i is added to the network at timestep t_i, we suppose that

$$k_i(t_i) = c_0, \tag{3}$$

where c_0 is a constant. The rate at which the node i acquires new edges is given by our algorithm:

$$\frac{dk_i(t)}{dt} = n_1 m_1 \Pi_i + n_2 \frac{1}{N(t)} + n_2(m_2 - 1)(1 - \frac{1}{N(t)}) \frac{1}{< k_{nn} >} + (m_4 - m_3) \frac{1}{N(t)}, \tag{4}$$

where $< k_{nn} >$ denotes the residual average degree of node i (that is the average degree of the neighbors of node i) [13]. According to [14], we suppose that $< k_{nn} >= k_i(t) + b$ and b is a small constant.

The network has $M(t) = (n_1 m_1 + n_2 m_2 + m_4 - m_3)t$ edges after t time steps. Then we have

$$\begin{aligned}\frac{dk_i(t)}{dt} &= \frac{n_1 m_1 k_i(t)}{2M(t)} + \frac{n_2+m_4-m_3}{N(t)} + \frac{n_2(m_2-1)(N(t)-1)}{N(t)} \times \frac{1}{k_i(t)+b} \\ &= \frac{Ak_i(t)}{t} + \frac{n_2+m_4-m_3}{n_0+(n_1+n_2)t} + \frac{C([n_0+(n_1+n_2)t]-1)}{[n_0+(n_1+n_2)t](k_i(t)+b)},\end{aligned} \tag{5}$$

where

$$A = \frac{n_1 m_1}{2(n_1 m_1 + n_2 m_2 + m_4 - m_3)}, C = n_2(m_2-1). \tag{6}$$

Let $t \to \infty$, it follows from (3) and(5) that

$$\begin{cases} \dfrac{dk_i(t)}{dt} = \dfrac{Ak_i(t)}{t} + \dfrac{C}{k_i(t)+b}, \\ k_i(t_i) = c_0, t > 0. \end{cases} \tag{7}$$

We solve (7) by using Euler method. The form of solution is as following,

$$k_i(t) = c_1(\frac{t}{t_i})^{\alpha} + c_2, \tag{8}$$

where $A \le \alpha < 1$, and c_1, c_2 are two constants. When $n_2 = m_2 = m_3 = m_4 = 0$, we can get $\alpha = A = 0.5$, $c_1 = m_1$, and $c_2 = 0$.

Then we obtain the probability distribution function

$$\begin{aligned} P(k_i(t) < k) &= P(c_1(\frac{t}{t_i})^{\alpha} + c_2 < k) \\ &= P(t_i > t(\frac{k-c_2}{c_1})^{-\frac{1}{\alpha}}) \\ &= 1 - P(t_i \le t(\frac{k-c_2}{c_1})^{-\frac{1}{\alpha}}). \end{aligned} \tag{9}$$

Since the evolution of the network nodes are added at equal intervals, the probability density of t_i is

$$\rho(t_i) = \frac{1}{t+n_0}. \tag{10}$$

Substituting (10) into (9), we get

$$P(k_i(t) < k) = 1 - \frac{t}{t+n_0}(\frac{k-c_2}{c_1})^{-\frac{1}{\alpha}}. \tag{11}$$

So, the distribution of node degree is

$$p(k,t) = \frac{\partial P(k_i(t) < k)}{\partial k} = \frac{t}{(t+n_0)\alpha}(\frac{k-c_2}{c_1})^{-\frac{1}{\alpha}-1}. \tag{12}$$

Let $t \to \infty$, we obtain the node degree distribution when the network is stable,

$$p(k) = \lim_{t\to\infty} p(k,t) = \frac{1}{\alpha}(\frac{k-c_2}{c_1})^{-\frac{1}{\alpha}-1}. \tag{13}$$

By calculation we can see that the degree of each node increases following a power-law. Note that $A<1,$ then the power exponent

$$\gamma = \frac{1}{\alpha} + 1 > 2. \tag{14}$$

Set $n_2 = m_2 = m_3 = m_4 = 0$, we can note that the network is BA scale-free network and compute the power exponent $\gamma_{BA} = 3$.

Initial $n_0 = 3, n_1 = 5, m_1 = 3, n_2 = 8, m_2 = 5, m_3 = 5, m_4 = 10.$ We can obtain several networks under different network scales by using R software. In Fig 1(a), We can know that the networks follow the power-law distribution and appear fat-tail phenomenon. As it can be seen in Fig 1(b), We adjusted the parameters of our model and obtain degree distribution under network scale $N = 1498$. We have calculated the power exponents by using regression fitting respectively (they are shown in Table 1).

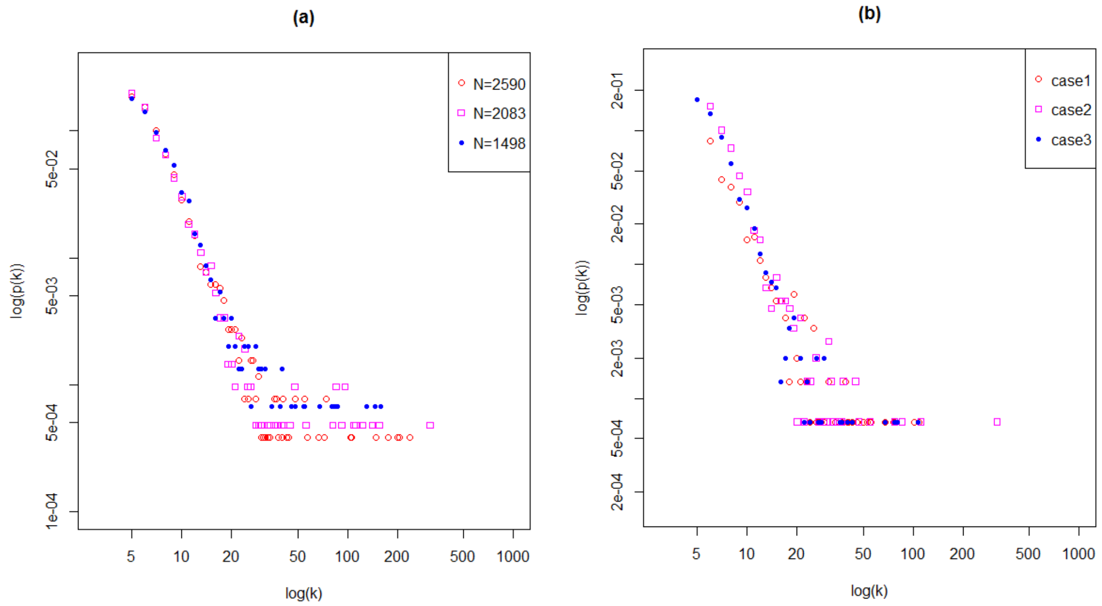

Fig 1: Power-law distribution in double logarithmic coordinate system

Table 1: Power exponent under different parameter values.

case	n_0	n_1	m_1	n_2	m_2	m_3	m_4	γ
1	3	5	3	0	0	0	0	2.8350
2	3	5	3	8	5	5	10	2.7553
3	3	8	3	5	3	7	10	2.3959

3.1. Average path length

The definition of average path length (APL) is the average of the distance between any two nodes, denoted by L. i.e.

$$L = \frac{1}{N(N-1)/2} \sum_{i \geq j} d_{ij}, \tag{15}$$

where N is the number of network nodes, and d_{ij} is the distance between node i and node j.

It can be seen from Fig 2 that the average path length of the extended model is obviously smaller than that the classical BA model, which shows the characteristics of the small world networks.

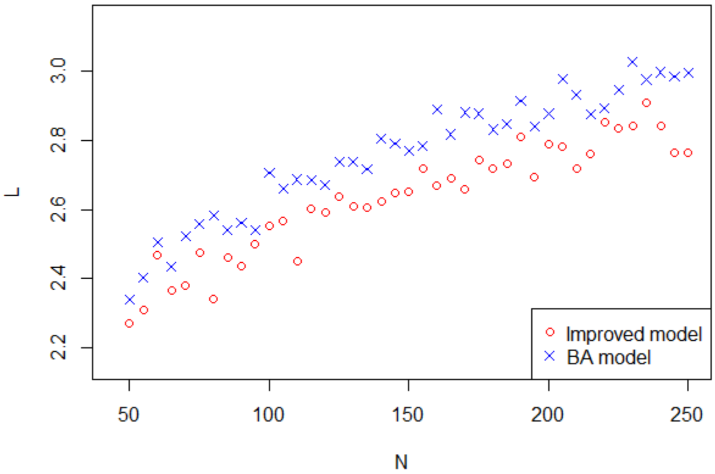

Fig 2: The APL of two models under different network scale.

3.2. Clustering coefficient

The clustering coefficient quantifies the characteristics of small groups in networks. In the network $G=(V,E)$, where V is the set of nodes and E is the set of edges. We suppose that the degree of node i is k_i. The actual number of edges between these k_i nodes is e_i. And the largest possible number of edges is $(k_i-1)k_i/2$. Then the clustering coefficient of node i:

$$C_i = 2e_i/(k_i(k_i-1)). \tag{16}$$

The clustering coefficient of the network is defined as the average of the clustering coefficients of all the nodes, i.e.

$$C=\frac{1}{|V|}\sum_{i\in V}C_i. \tag{17}$$

It can be seen from Fig 3 that the clustering coefficient of the extended model is significantly larger than the classical BA, which is due to the extended model to join the idea of local preferential attachment, making the network more ‘small groups’. Our model has a lager clustering coefficient, which is more in line with the real network.

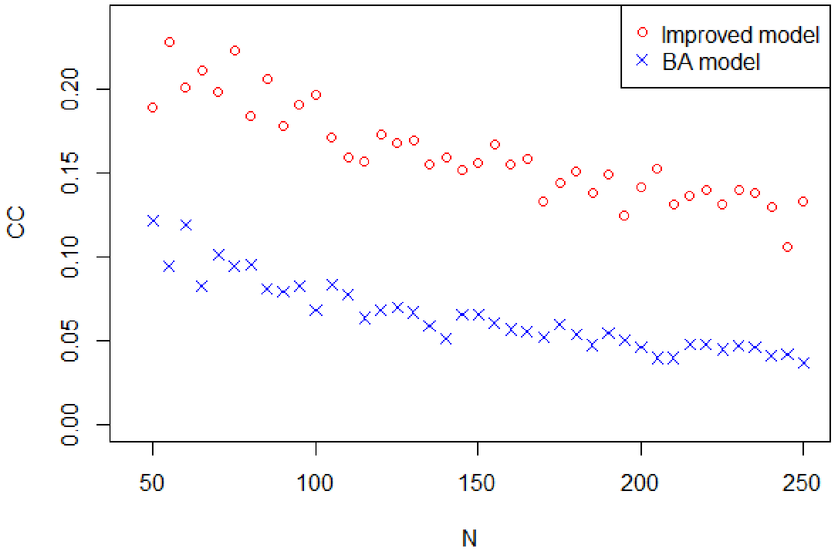

Fig 3: The clustering coefficient of two models under different network scale.

4. Conclusion

Base on BA scale-free network model, we consider the formation process of social network and propose a scale-free network evolution model. In the extended model, local preferential attachment mechanisms and random attachment or removal between the old and new edges are taken into account. When some of the initial parameters in the model are set as 0, the extended model is the classical BA model, so BA scale-free network model is a special case of the extended network model. In addition, we obtain some statistical properties of the extended model. And we use R software to simulate the model, the results prove the correctness of theoretical analysis. Compared with the classical BA model, the scale-free network based on the growth of social networks has characteristics of small average path length and large clustering coefficient, which is more in line with the realistic network.

5. Reference

[1] A. L. Barabasi, *Scale-Free Networks: A Decade and Beyond,* Science, 325: 412-413 (2009).

[2] A. L.Barabasi, *The Barabasi-Albert Model,* Network Science (2014).

[3] R. Albert, A. L. Barabasi, *Topology of evolving networks: Local events and universality,* Phys. Rev. E, 85(24):

[4] G. Bianconi, A. L. Barabasi, *Bose-Einstein condensation in complex networks,* Phys. Rev. Lett, 86(24): 5632-5635 (2001)

[5] D. J. Watts, S.H. Strogatz, *Collective Dynamics of 'Small-World' Network*. Nature, 393(6684): 440-442 (1998).

[6] M. Kochen, *The Small World*, Norwood, N.J (1989).

[7] S. Milgram, *The Small World Problem*, *Psychol*, Today 2, 60-67 (1967).

[8] A. L. Barabasi, R. Albert, H. Jeong, *Diameter of the World Wide Web*, Nature, 401(6749): 130-131 (1999).

[9] A. L. Barabasi, R. Albert, *Emergence of scaling in random network,* Science, 286(5439): 509-512 (1999).

[10] X. Li, Y.Y. Jin, G. Chen, *Complexity and Synchronization of the World Trade Web,* Physica A, 328: 287-296 (2003).

[11] X. Li, G. Chen, *A local world evolving network model,* Physica A, 328: 274-286 (2003).

[12] J. L. Wang, L. B. Xu, Y.C. Pang, *Evolution Model of Online Social Networks Based on Complex Networks. Journal of Intelligent Systems,*10(06): 949-953 (2015).

[13] X. F. Wang, X. LI, G. R. Chen, *Introduction to network science,* Higher Education Press (2012).

[14] S. Feld, *Why your friends have more friends than you do,* Amer. J. of Sociology, 96(6): 1464-1477 (1991).