Mandelbrot Law of Evolving Networks

Xue-Zao Ren[1], Zimo Yang[2,3], Bing-Hong Wang[1,4,5], and Tao Zhou[2,4*]
[1]*School of Science, Southwest University of Science and Technology, Mianyang 621010, People's Republic of China*
[2]*Web Sciences Center, University of Electronic Science and Technology of China, Chengdu 610054, People's Republic of China*
[3]*Department of Physics, The Chinese University of Hong Kong, Shatin, Hong Kong, People's Republic of China*
[4]*Department of Modern Physics, University of Science and Technology of China, Hefei 230026, People's Republic of China*
[5]*Research Center for Complex System Science, University of Shanghai for Science and Technology, Shanghai 200093, People's Republic of China*
(Dated: July 26, 2011)

Degree distributions of many real networks are known to follow the Mandelbrot law, which can be considered as an extension of the power law that determined by not only the power-law exponent, but also the shifting coefficient. Although the shifting coefficient highly affect the shape of distribution, it receives less attention in the literature and in fact, mainstream analytical method based on backward or forward difference will lead to considerable deviation to its value. In this article, we show that the degree distribution of a growing network with linear preferential attachment approximately follows the Mandelbrot law. We propose an analytical method based on a recursive formula that can obtain a more accurate expression of the shifting coefficient than the previous methods. Simulations demonstrate the advantages of our method. This work provides a possible mechanism leading to the Mandelbrot law of evolving networks, and refines the mainstream analytical methods for the shifting coefficient.

PACS numbers: 89.75.Hc, 89.75.Fb, 02.50.-r

I. INTRODUCTION

Many systems can be described as complex networks [1–4], in which, the nodes correspond to the elements and the links to the relations between elements. Uncovering the mechanisms underlying the structural features of real networks is one of the most interesting challenges in network science. Two pioneering models, respectively for small-world [5] and scale-free networks [6], give explanations for many real phenomena, such as, the logarithmic growth of average distance, the power-law degree distribution, and the high clustering coefficient. With the idea of 'rich get richer', Barabási and Albert proposed the scale-free network model, embedding two mechanisms: growth and preferential attachment. That is, at each time step, a new node is added and connected to a few old nodes with probability proportional to their degree as:

$$\Pi(k_i) = k_i / \sum_j k_j, \tag{1}$$

where k_i is the degree of node i, and j runs over all old nodes. The analytical solution of the degree distribution,

$$p(k) = 2m^2 k^{-3}, \tag{2}$$

can be obtained by applying the mean-field approximation [6, 7], in which $2m$ is the average degree of the network.

Unfortunately, for many real networks, the degree distributions are different from exact power laws [8, 9]. For example, the scientific collaboration networks can be better characterized by the power-law distributions with exponential cutoff [10], the degree distributions of the email networks [11], some collaboration networks [12], and online user-object bipartite networks [13] obey the stretched exponential forms, and the double power-law distribution seems a better way to describe the air transportation networks [14–16]. In this article, we focus on the *Mandelbrot law* or called the *shifted power law* [17], which can be written as:

$$p(k) \propto (k+c)^{-\gamma}, \tag{3}$$

where γ is the power-law exponent and c is the shifting coefficient. Recently, the Mandelbrot law has been applied to characterize the degree distributions of some real networks [18, 19]. Even for the well-known BA model, the degree distribution

$$p(k) = \frac{2m(m+1)}{(k+2)(k+1)k} \approx 2m^2 k^{-3}, \tag{4}$$

obtained by the master equation [20], is not an exactly power-law distribution. This distribution can be approximated as $p(k) \propto (k+1)^{-3}$, which also satisfies the Mandelbrot law with $\gamma = 3$ and $c = 1$.

A number of tools have been developed to get the analytical solutions of network degree distributions, including the mean-field approximation, the master equation, the rate equation, and so on [7, 20–23]. Most of these kinds of known analytical methods only concentrate on the power-law exponent, yet paid less attention to the value of shifting coefficient, which, however, plays a significant role in determining the shape of degree distributions. Even worth, we will show later that the widely used difference approximation, no matter forward difference or backward difference, will result in considerable deviation to the real value of the shifting coefficient. We propose an analytical method based on a recursive formula that can obtain a more accurate expression of the shifting coefficient than the previous methods. In addition, we show that the degree

*Electronic address: zhutou@ustc.edu

distribution of a growing network with linear preferential attachment approximately follows the Mandelbrot law. This article is organized as follows. In Section 2, we will present the evolving network model with linear preferential attachment, which can be considered as an extension of the BA model. In Section 3, we will analyze the model, especially propose an analytical method to accurately predict the shifting coefficient. Finally, we will summarize our work in the last section.

II. MODEL

The present model embodies a linear preferential attachment, which can be considered as an extension of the famous BA model. Initially, our model starts with a fully connected network with m_0 nodes and $m_0(m_0-1)/2$ links. If the final size is S, it should satisfy the condition $m_0 \ll S$. After initialization, at each time step, a new node will be added into the network, which will connect to m old nodes. The probability of an old node i to be connected is linearly connected with its degree k_i, say

$$\Pi(k_i) = \frac{\alpha k_i + \beta}{N} = \frac{am(\alpha k_i + \beta)}{2mN} = \frac{am(\alpha k_i + \beta)}{\sum_j k_j}, \quad (5)$$

where α and β are two parameters. This model will degenerate to the BA model if $\beta = 0$. The self-loop and multiple links are not allowed in this model. The parameters α and β satisfy the normalization condition

$$\sum_k \pi(k)p(k) = 1, \quad (6)$$

where $\pi(k)$ is the probability a selected node is of degree k. That is

$$\sum_k \pi(k)p(k) = \sum_k (\alpha k + \beta)p(k) = 2m\alpha + \beta = 1. \quad (7)$$

It is equivalent to:

$$\alpha = \frac{1}{2m}(1-\beta). \quad (8)$$

III. ANALYSIS

The rate equation is based on the assumption that the added nodes and links, during a time step, has no influence on the global degree distribution of the network. That is to say, in the large limit of the network, the degree distribution approaches to a steady form. Denote $p(k)$ the steady degree distribution and N the number of nodes in the current time step, if N is large enough, then the number of nodes with degree k is approximated to $Np(k)$. Analogously, the number of nodes with degree k in the next step is $(N+1)p(k)$. Accordingly, since during a time step in total m links are added with no

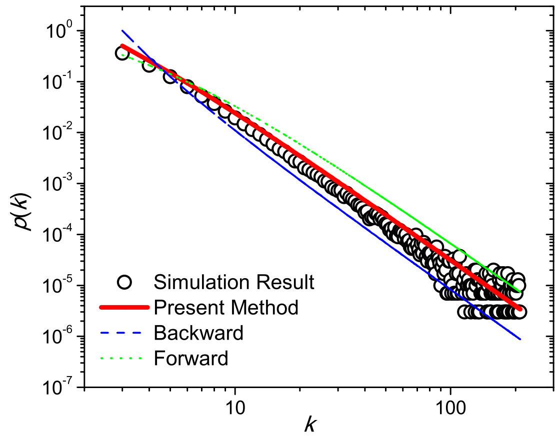

FIG. 1: Degree distribution of the modeled network with $\beta = 0$ and $m = 3$. Since $\beta = 0$, it is equivalent to a BA network. Compared with the results of backward difference approximation (blue dash line) and forward difference approximation (green dot line), the present method based on recursive formula (red solid line) is closer to the simulation result (round donuts). The network size is $S = 10^4$ and the results are obtained by averaging over 100 independent realizations.

influence on the degree distribution $p(k)$, the number of nodes with degree k in the time step $N+1$ reads

$$(N+1)p(k) = Np(k) + m\pi(k-1)p(k-1) - m\pi(k)p(k) + \delta_{km}, \quad (9)$$

where $m\pi(k-1)p(k-1)$ and $m\pi(k)p(k)$ represent, respectively, the number of nodes whose degree changes from $k-1$ to k in this time step, and the number of nodes whose degree changes from k to $k+1$ in this time step whose. δ_{km} accounts for the specific degree equal to m, namely $\delta_{km} = 1$ when $k = m$ and $\delta_{km} = 0$ otherwise. Eq. (9) is the usually form of the well-known rate equation [21, 22]. This equation is usually solved by using the difference approximation, however, here we will show a much different method that will lead to a recursive formula. We will later compare our results with the ones obtained by the difference approximation.

Eq. (9) is equivalent to

$$\begin{cases} p(m) = \frac{1}{1+m\pi(m)}, k = m, \\ p(k)\left[1 + m\pi(k)\right] = m\pi(k-1)p(k-1), k > m. \end{cases} \quad (10)$$

Reminding the linear relation

$$\pi(k) = \alpha k + \beta, \quad (11)$$

considering Eq. (8), the probability of a newly added link connecting to an old node with minimum degree is

$$\pi(m) = \frac{1-\beta}{2m}m + \beta = \frac{1+\beta}{2}. \quad (12)$$

Clearly, this probability should be no less than zero and no larger than one, and thus $-1 \le \beta \le 1$. According to Eq. (10), the probability density of m-degree nodes is

$$p(m) = \frac{1}{1+m\pi(m)} = \frac{2}{2+m(1+\beta)}. \tag{13}$$

Substituting Eq. (8) into Eq. (10), we get

$$p(k)\left[k + \frac{2(1+m\beta)}{1-\beta}\right] = \left[k + \frac{2m\beta}{1-\beta} - 1\right] p(k-1). \tag{14}$$

Specifying:

$$\begin{aligned} a &= \frac{2m\beta}{1-\beta} - 1, \\ b &= \frac{2(1+m\beta)}{1-\beta}, \end{aligned} \tag{15}$$

then Eq. (14) can be rewritten in a simple recursive formula as

$$p(k) = \frac{k+a}{k+b} p(k-1). \tag{16}$$

Taking logarithm in both sides of Eq. (16), we get

$$\log \frac{p(k)}{p(k-1)} = \log \frac{k+a}{k+b}. \tag{17}$$

With the ansatz that $p(k)$ follows the Mandelbrot law, substituting Eq. (3) into Eq. (17), we can obtain the relationship between the power-law exponent γ and the shifting coefficient c as

$$\log \frac{k+a}{k+b} = \gamma \log \frac{k-1+c}{k+c}, \tag{18}$$

which is equivalent to:

$$\log \frac{1+a\frac{1}{k}}{1+b\frac{1}{k}} = \gamma \log \frac{1+(c-1)\frac{1}{k}}{1+c\frac{1}{k}}. \tag{19}$$

Under the approximation with large k, through the second order Taylor expansion of Eq. (19) with $1/k$ being the variable, we can get the power-law exponent

$$\gamma = b - a = 1 + \frac{2}{1-\beta}, \tag{20}$$

and the shifting coefficient

$$c = \frac{b+a+1}{2} = \frac{1+2m\beta}{1-\beta}. \tag{21}$$

Eq. (20) and Eq. (21) declare that the power-law exponent γ only depends on the parameter β, while the shifting coefficient c is related to both β and m. When $m\beta$ is very large or $\beta \to 1$, a and b are both very large, and the Taylor expansion cannot be applied on Eq. (19). Under such condition, Eq. (16) can be approximately rewritten as

$$p(k) \approx \frac{a}{b} p(k-1), \tag{22}$$

namely the degree distribution is close to an exponential distribution. It is easy to be understood since when $\beta \to 1$, the selection of old nodes is almost random. When $\beta = 0$, $\alpha = \frac{1}{2m}$, our model degenerates to the BA model, and we can get $a = -1$, $b = 2$, $\gamma = 3$ and $c = 1$, then the degree distribution approaches to

$$p(k) = -\frac{2}{\psi(2, m+1)}(k+1)^{-3}, \tag{23}$$

where

$$\psi(x) = \Gamma'(x)/\Gamma(x) \tag{24}$$

is the *Digamma function* with

$$\Gamma(x) = \int_0^\infty e^{-t} t^{x-1} dt \tag{25}$$

being the *Gamma function* and

$$\psi(n, x) = \frac{d^n \psi(x)}{dx^n}. \tag{26}$$

Hereinafter, we will compare the present method with the traditional method based on the difference approximation. We first introduce the backward difference approximation, which assumes

$$\frac{dp}{dk} = p(k) - p(k-1). \tag{27}$$

Substituting Eq. (27) into Eq. (16), we get

$$p(k) = \frac{k+a}{k+b}\left[p(k) - \frac{dp}{dk}\right], \tag{28}$$

which is equivalent to

$$\frac{dp}{dk} = \frac{a-b}{k+a} p(k) \tag{29}$$

that lead to the solution

$$p(k) \propto (k+a)^{-(b-a)}. \tag{30}$$

Similarly, if we apply the forward difference approximation by assuming

$$\frac{dp}{dk} = p(k+1) - p(k), \tag{31}$$

then Eq. (16) can be rewritten as

$$p(k+1) = \frac{k+a+1}{k+b+1} p(k), \tag{32}$$

which is equivalent to

$$\frac{dp}{dk} = \frac{a-b}{k+1+b} p(k). \tag{33}$$

In this case, the solution is

$$p(k) \propto (k+b+1)^{-(b-a)}. \tag{34}$$

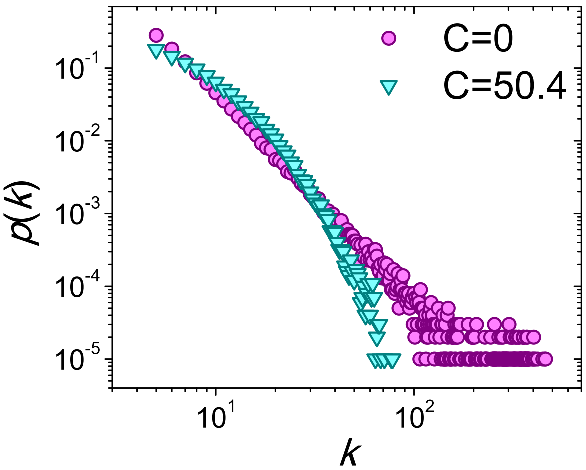

FIG. 2: The comparison of degree distributions with different shifting coefficients given $m = 5$ and $S = 10000$. Compared with the case of $\beta = 0.8$ and shifting coefficient $c = 50.4$ (green down-triangles), the degree distribution of the none-shifting case with $\beta = -0.1$ and $c = 0$ (purple circles) is much more close to a straight line in the log-log coordinates. The results are obtained by averaging over 100 independent realizations.

The three methods all indicate that the Mandelbrot law will emerge from an evolving network with linear preferential attachment, and they all give the same power-law exponent $\gamma = b - a$. In contrast, the shifting coefficient are different: $c^{\texttt{present}} = \frac{a+b+1}{2}$, $c^{\texttt{backward}} = a$ and $c^{\texttt{forward}} = b + 1$. As shown in Fig. 1, we compare the degree distributions obtained by these three methods with the simulation results, which shows that the present method is observably more close to the simulation results.

Although we usually refer to the concept of scale-free networks, neither the BA networks nor most real networks have very precise power-law degree distributions. The present method suggests that we can obtain a more precise power-law distribution by setting a right β that corresponds to a zero shifting coefficient. Since the degree distribution is

$$p(k) \propto (k + \frac{1+2m\beta}{1-\beta})^{1+\frac{2}{1-\beta}}, \tag{35}$$

it asks for

$$c = \frac{1+2m\beta}{1-\beta} = 0, \tag{36}$$

namely $\beta = -\frac{1}{2m}$ and $p(k) \propto k^{3-\frac{2}{2m+1}}$. That is, given the linear preferential attachment, the non-shifted power-law exponent is determined by the network's average degree and can never exceed 3. Figure 2 compares two degree distributions, respectively with $c = 0$ and $c = 50.4$, from which one can confirm that the non-shifted power law is indeed much closer to a straight line in the log-log coordinates, and the shifting coefficient largely affects the shape of degree distribution.

IV. CONCLUSION

In this article, we extend the BA model to an evolving model with linear preferential attachment and show that this model will generate networks with Mandelbrot-law degree distributions, which are previously observed in many real systems. As shown in Fig. 2, the shifting coefficient, usually being ignored in the literature, largely affects the shape of degree distribution. In puzzlement, the backward and forward difference approximations will lead to different solutions on shifting coefficient, although they give the same estimation to the power-law exponent. Our analysis indicate that both of them are inaccurate, and we propose an analytical method that results in a more accurate solution. Simulations demonstrate the advantages of our method.

[1] S. H. Strogatz, Nature (London) **410**, 268 (2001).
[2] S. N. Dorogovtsev, Adv. Phys. **51**, 1079 (2002).
[3] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).
[4] M. E. J. Newman, SIAM Rev. **167**, 45 (2003).
[5] D. J. Watts and S. H. Strogatz, Nature (London) **393**, 440 (1998).
[6] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
[7] A. L. Barabási, R. Albert, and H. Jeong, Physica A **272**, 173 (1999).
[8] M. E. J. Newman, Comtemporary Physics **46**, 323 (2005).
[9] A. Clauset, C. R. Shalizi, and M. E. J. Newman, SIAM Rev. **51**, 661 (2009).
[10] M. E. J. Newman, Proc. Natl. Acad. Sci. U.S.A. **98**, 404 (2001).
[11] M. E. J. Newman, S. Forrest, and J. Balthrop, Phys. Rev. E **66**, 035101(R) (2002).
[12] P.-P. Zhang, K. Chen, Y. He, T. Zhou, B.-B. Su, Y.-D. Jin, H. Chang, Y.-P. Zhou, L.-C. Sun, B.-H. Wang, and D.-R. He, Physica A **360**, 599 (2006).
[13] M.-S. Shang, L. Lü, Y.-C. Zhang, T. Zhou, Europhys. Lett. **90**, 48006 (2010).
[14] W. Li and X. Cai, Phys. Rev. E **69**, 046106 (2004).
[15] H.-K. Liu and T. Zhou, Acta Phys. Sin. **56**, 106 (2007).
[16] G. Bagler, Physica A **387**, 2972 (2008).
[17] B. Mandelbrot, *Information theory and psycholinguistics* (New York: Basic Books Publishing Co., 1965).
[18] H. Chang, B. B. Su, Y. P. Zhou, and D. R. He, Physica A **383**, 687 (2007).
[19] Y. L. Wang, T. Zhou, J. J. Shi, J. Wang, and D.-R. He, Physica A **388**, 2949 (2009).
[20] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samuhkin, Phys. Rev. Lett. **85**, 4633 (2000).
[21] P. L. Krapivaky, S. Redner, and F. Leyvraz, Phys. Rev. Lett. **85**, 4629 (2000).
[22] T. Zhou, G. Yan, and B.-H. Wang, Phys. Rev. E **71**, 046141 (2005).
[23] C. Moore, G. Ghoshal, and M. E. J. Newman, Phys. Rev. E **74**, 036121 (2006).