PHYSICAL REVIEW E, VOLUME 63, 025101(R)

Effect of the accelerating growth of communications networks on their structure

S. N. Dorogovtsev[1,2,*] and J. F. F. Mendes[1,†]
[1]Departamento de Física and Centro de Física do Porto, Faculdade de Ciências, Universidade do Porto, Rua do Campo Alegre 687, 4169-007 Porto, Portugal
[2]A.F. Ioffe Physico-Technical Institute, 194021 St. Petersburg, Russia

(Received 6 September 2000; published 18 January 2001)

Motivated by data on the evolution of the Internet and World Wide Web we consider scenarios of self-organization of nonlinearly growing networks into free-scale structures. We find that the accelerating growth of networks establishes their structure. For growing networks with preferential linking and increasing density of links, two scenarios are possible. In one of them, the value of the exponent γ of the distribution of the number of incoming links is between 3/2 and 2. In the other scenario, $\gamma>2$ and the distribution is necessarily nonstationary.

DOI: 10.1103/PhysRevE.63.025101 PACS number(s): 05.10.−a, 05.40.−a, 05.50.+q, 87.18.Sn

The incredible place of the Internet in our civilization evokes the exponentially increasing flow of the studies of evolving networks (Internet, World Wide Web, nets of citations, collaboration networks, etc.) [1–10]. The study of communications networks has a long history (e.g., see the papers of "the father" of the Internet, Baran [11]). Nevertheless, the first data on the structure of the Internet, Web, etc. were obtained only recently [1–3,12–14]. Most of them are on the simplest "one-site" characteristic of networks, site degree distribution. Site degree is the number of connections of a site. The observation of power-law dependences of site degree distribution in these networks puts forward the question, How can a growing network organize itself into a scale-free structure?

The most natural mechanism of such self-organization is the network growth with preferential attachment of new links to sites with a high number of connections [6,12,15]. There are several ways to introduce the preferential linking that lead to different values of scaling exponents of site degree distribution inside of a huge interval [16–18]. Nevertheless, it is still unclear why the exponents of real systems have their specific values, and why the degree distributions have their specific forms.

The last data on the Internet evolution demonstrate that the total number of links increases more quickly than the number of sites [13]. For instance, the study of the interdomain topology of the Internet [13] has shown that the relation of the edges and the nodes equaled 1.71 in November of 1997 (3015 nodes) and 1.88 in December of 1998 (8256 nodes). The number of pages in the World Wide Web was 203×10^6 with 1466×10^6 links between them in May 1999 (the corresponding ratio is 7.22) and it was already 271×10^6 pages and 2130×10^6 links in October 1999 (the ratio is 7.86) [14]. This looks quite natural, as old sites may establish new connections all the time. Old documents on the Web may be updated and new references may be added to them. Collaboration networks [8,10] become more dense while growing due to the increasing possibility of finding collaborators. Thus, we certainly face a wide spread situation: the density of links of a network becomes higher and higher during its evolution, the total number of links grows faster than the number of sites, the growth is nonlinear, i.e., *accelerating*.

This very significant factor is omitted in most of the considered models [6,12,15–18]. One may ask: What is the effect of the accelerating growth on the structure of the networks? What kind of acceleration produces scale-free networks? What is the form of the degree distribution in such networks?

In the present Rapid Communication, we consider these problems using general arguments. Assuming that the networks are scale-free, we describe possible degree distributions and show that, in this case, the total number of links is a power-law function of the network size. We introduce simple models with preferential linking accounting for the accelerated growth and study their evolution both analytically using the continuous approach [15,16,18] and by simulation. The arising distributions cover the range of the forms predicted from the general considerations. Also, we find the situation in which the flow of new links, increasing via a power law, takes the growing network out of the class of free-scale structures.

We start from general considerations. In scale-free networks, a wide range of the degree distribution function is of the power-law form, $P(q)\propto q^{-\gamma}$. It will be clear from the following that, to keep the network in the class of free-scale nets, the flow of new links has to be a power function of the number of sites of the network that plays the role of time, i.e., be proportional to t^{α}.

First, let us assume that the exponent of the distribution is less than 2. The reasonable range is $1<\gamma<2$. To produce the restricted average degree (that is proportional to t^{α}), the distribution has to have a cut-off at large q. Its natural value is of the order of the total number of sites, which is proportional to $t^{\alpha+1}$, although other situations are also possible. It is also clear from the following that, if the distribution is nonstationary, the degree distribution of scale-free networks has to be of the form $t^{z}q^{-\gamma}$, between $q\sim t^{x}$ and $q\sim t^{\alpha+1}$, where z and x are some exponents. It is a restricted function in the range below $q\sim t^{x}$. $z\geqslant0$ since the network grows.

*Electronic address: sdorogov@fc.up.pt
†Electronic address: jfmendes@fc.up.pt

From the normalization condition, $\int_0^\infty dq\, P(q,t)=1$, we get $x=z/(\gamma-1)$. The average degree of the network, $\overline{q}$, is of the order $t^{\alpha+1}/t$, then $t^\alpha \sim \int^{t^{\alpha+1}} dq\, q\, t^z q^{-\gamma} \sim t^{z+(2-\gamma)(\alpha+1)}$ (the value of the integral is defined by its upper limit). Therefore, $\alpha=z+(2-\gamma)(\alpha+1)$, and, finally, $\gamma=1+(1+z)/(1+\alpha)$. One sees that $z<\alpha$ to keep γ below 2 as it was assumed. Also, one sees that the lower boundary for γ, $1+1/(1+\alpha)$, is approached for stationary distribution, $z=0$. In this case, the distribution form is completely fixed by the accelerating growth, the exponent γ depends only on α.

Obviously, α cannot exceed 1 (the total number of links has to be smaller than $t^2/2$ since one may forbid multiple links). Hence, $\gamma>3/2$. The value $\alpha=0$ corresponds to permanent density of links [16]. The density of connections in real networks remains rather low all the time [13], so one may reasonably assume that α is small. Therefore, the lower boundary of the possible values of γ is close to 2.

The other possibility is $\gamma>2$. In this case, again $x=z/(\gamma-1)$ but the integral for the average degree is defined by its lower limit, $t^\alpha \sim \int_{t^{z(\gamma-1)}} dq\, q\, t^z q^{-\gamma} \sim t^{z-z\gamma-2)/(\gamma-1)}$. Hence, $\gamma=1+z/\alpha$, and $z>\alpha$. Thus, we have described the possible forms of the degree distribution.

Let us demonstrate how these distributions may arise in the nonlinearly growing networks with the preferential linking. We do not restrict ourselves to some model generating increasing flow of new links dynamically in the process of the growth, since it would narrow the range of possible dependences on time for this flow. (We do not discuss here possible mechanisms of such generation. One of them was considered in Ref. [19].) Instead of that, we prefer to introduce this dependence *ab initio*.

Let us introduce the following simple model of a growing network with *directed* links. We consider only the distribution of incoming links, so, here, in-degree is the number of incoming links q_s of the site s. At each increment of time, a new site is added to the network. Let its initial in-degree be n, where $n\geq 0$. This means that it has n incoming links from some nonspecified old sites. Let extra $c_0 t^\alpha$ new directed links be distributed between old sites. We assume that each of these links goes out from a nonspecified site and is directed to some site s chosen with probability proportional to the sum of its in-degree and some constant, q_s+A. Here, the sum $n+A$ characterizes the initial ''attractiveness'' of a site for new links [16]. It has to be positive.

In fact, we generalize the model [16] to the case of the increasing number of new links. The introduced power-law dependence of the number of new links keeps the network in the class of free-scale networks but changes crucially the in-degree distribution.

We use the continuous approach [15,16,18], which gives exact values for the scaling exponents of such systems, as was demonstrated in [16]. Each site is labeled by the time of its birth, $0<s\leq t$. The in-degree distribution of the site s in the continuous approximation is of the form $p(q,s,t)=\delta(q-\overline{q}(s,t))$ [16], where $\delta[\ ]$ is the δ-function and $\overline{q}(s,t)$ is the average in-degree of the site s at time t. Then,

$$\frac{\partial \overline{q}(s,t)}{\partial t}=c_0\, t^\alpha \frac{\overline{q}(s,t)+A}{\int_0^t du[\overline{q}(u,t)+A]}, \tag{1}$$

$\overline{q}(0,0)=0$, $\overline{q}(t,t)=n$. Equation (1) describes distribution of flow of new links between sites according the introduced rules. Applying $\int_0^t ds$ to both sides of Eq. (1), one gets $\int_0^t du\, \overline{q}(u,t)=nt+c_0 t^{\alpha+1}/(\alpha+1)$, and accounting for the boundary condition, $\overline{q}(t,t)=n$, we obtain the solution,

$$\frac{\overline{q}(s,t)+A}{n+A}=\left[\frac{1+(n+A)(1+\alpha)t^{-\alpha}/c_0}{1+(n+A)(1+\alpha)s^{-\alpha}/c_0}\right]^{1+1/\alpha}\left(\frac{s}{t}\right)^{-(\alpha+1)}. \tag{2}$$

In the interval $[(n+A)(1+\alpha)/c_0]^{1/\alpha}\ll s\ll t$, we get

$$\overline{q}(s,t)=(n+A)\left(\frac{s}{t}\right)^{-(\alpha+1)}. \tag{3}$$

Thus, the exponent β, defined by the relation $\overline{q}(s,t)\propto s^{-\beta}$, equals $1+\alpha$ and is greater than 1. The dependence, $\overline{q}(s)$, becomes constant,

$$\overline{q}(s,t)=(n+A)^{-1/\alpha}\left(\frac{c_0}{1+\alpha}\right)^{1+1/\alpha} t^{\alpha+1}, \tag{4}$$

for $s\ll[(n+A)(1+\alpha)/c_0]^{1/\alpha}$. One may compare this result with the total number of links in the network, $N(t)\approx c_0 t^{\alpha+1}/(1+\alpha)$.

In the continuous approximation, one may easily find the distribution $P(q,t)$ using the obtained $\overline{q}(s,t)$,

$$P(q,t)=\frac{1}{t}\int_0^t ds\, p(q,s,t)=\frac{1}{t}\int_0^t ds\ \delta(q-\overline{q}(s,t))$$
$$=-\frac{1}{t}\left(\frac{\partial \overline{q}(s,t)}{\partial s}\right)^{-1}. \tag{5}$$

Therefore, in the region $1\ll q/(n+A)\ll\{c_0/[(n+A)(1+\alpha)]\}^{1+1/\alpha}t^{1+\alpha}$, the in-degree distribution has the following form:

$$P(q,t)=\frac{(n+A)^{1/(1+\alpha)}}{1+\alpha}q^{-[1+1/(1+\alpha)]}. \tag{6}$$

Thus, we obtain the stationary in-degree distribution, $P(q)\propto q^{-\gamma}$, with $\gamma=1+1/(1+\alpha)$ that belongs to one of the types described above.

The last result follows also from the relation $\beta(\gamma-1)=1$ between the scaling exponents, which may be obtained from the assumption that $\overline{q}(s,t)$ and $P(q,t)$ are power-law functions of s and q correspondingly [16]. Note that, if $A\to\infty$, the network is out of the class of scale-free networks for any $\alpha\geq 0$.

The introduction of the increasing flow of new links in the problem crucially changes in-degree distribution. Indeed, in the case of a constant density of links, γ varies from 2 to ∞ depending on the network parameters, n and A [16], and the

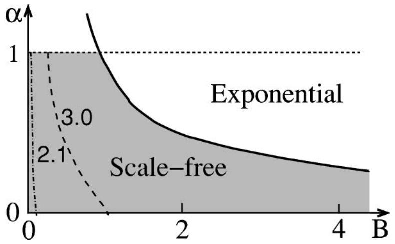

FIG. 1. Phase diagram of the networks with the accelerating growth under consideration. The networks are out of the class of scale-free nets (''exponential'') above the line $\alpha=1/B$. The exponent γ equals 3 on the dashed line and 2.1 (value for the World Wide Web) on the dash-dotted one. $\gamma<2$ on the line $B=0$.

values of the exponent β are between 0 and 1. Here, for the increasing density of links, we obtain γ below 2, and the exponent β exceeds 1. The values of the exponents are independent of n and A.

To demonstrate the other possibility, $\gamma>2$, we consider below the model with a different rule of distribution of new links. We make the only change in the studied above model. Let now a new link be directed at some site s with probability proportional to $q_s(t)/\overline{q}(t)+B$, where $q_s(t)$ is the in-degree of the site s, $\overline{q}(t)$ is average in-degree of sites of the network, and B is some positive constant. $\overline{q}(t)\cong c_0t^\alpha/(1+\alpha)$, so new links are distributed between sites with probability proportional to $q_s(t)+Bc_0t^\alpha/(1+\alpha)$, where a_0 is a positive constant. Hence, the previously introduced A becomes time-dependent.

Repeating the previous calculations, one gets

$$\frac{\partial\overline{q}(s,t)}{\partial t}=c_0t^\alpha\frac{\overline{q}(s,t)+Bc_0t^\alpha/(1+\alpha)}{nt+Bc_0t^{\alpha+1}/(1+\alpha)+c_0t^{\alpha+1}/(\alpha+1)}, \tag{7}$$

with the boundary condition $\overline{q}(t,t)=n$. At long times, one obtains

$$\frac{\partial\overline{q}(s,t)}{\partial t}=\frac{1+\alpha}{1+B}\,\frac{\overline{q}(s,t)+Bc_0t^\alpha/(1+\alpha)}{t}. \tag{8}$$

The solution of Eq. (8) is

$$\overline{q}(s,t)=\left[n+\frac{Bc_0s^\alpha}{1-B\alpha}\right]\left(\frac{s}{t}\right)^{-(1+\alpha)/(1+B)}-\frac{Bc_0t^\alpha}{1-B\alpha}. \tag{9}$$

If $B=0$, we obtain the previous result, $\beta=1+\alpha$. For $s^\alpha \gg n(1-B\alpha)/(Bc_0)$,

$$\overline{q}(s,t)\approx\frac{Bc_0t^\alpha}{1-B\alpha}\left\{\left(\frac{s}{t}\right)^{\alpha-(1+\alpha)/(1+B)}-1\right\}. \tag{10}$$

Therefore, the scaling exponents of the growing network are $\beta=(1+\alpha)/(1+B)-\alpha=(1-B\alpha)/(1+B)$ and $\gamma=1+1/\beta=1+[(1+\alpha)/(1+B)-\alpha]^{-1}=2+B(1+\alpha)/(1-B\alpha)$. The in-degree distribution differs sharply from the distribution obtained for the previous model. It is nonstationary and is of the form $P(q,t)\sim t^{-1+(1+\alpha)(1-B\alpha)}q^{-[1+(1+B)/(1-B\alpha)]}$ for $q \gg t^\alpha$. In this case, $\beta<1$ and $\gamma>2$ for any positive α and B. The scaling regime is realized if $B\alpha<1$. The general phase diagram for both considered models is shown in Fig. 1.

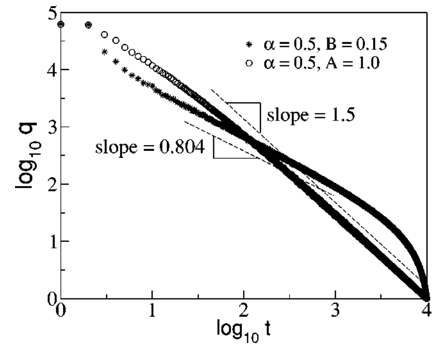

FIG. 2. Log-log plot of the average in-degree of a site (number of incoming links) vs its number (birth time) for the considered models. For the first model, $\alpha=0.5$, $n=1$, $A=1.0$, and $c_0=1.0$. For the second model, $\alpha=0.5$, $n=1$, $B=0.15$, and $c_0=1.0$. The dashed lines have the slopes equal to the values of the scaling exponent β obtained analytically.

Note that, in both of the considered cases, one cannot set $\alpha=0$ directly in the obtained expression for the scaling exponents. In such a situation, we get from Eqs. (1) or (7) $\beta=[1+(A+n)/c_0]^{-1}$ and $\gamma=2+(A+n)/c_0$ [16].

It is known that the used continuous approach gives exact results for the scaling exponents of the growing networks with a constant density of connections [16]. Nevertheless, it is approximate, so we have checked the obtained above results by simulation.

The results of the simulation of the considered models are shown in Figs. 2 and 3. The size of networks in both cases studied is 10 000 sites. The number of the attempts equals 1000. In Fig. 2, we present the log-log plots of the average in-degree versus the number of a site for $\alpha=0.5$, $n=1$, $A=1.0$, $c_0=1.0$ (the first model) and for $\alpha=0.5$, $n=1$, $B=0.15$, $c_0=1.0$ (the second one). In Fig. 3, for these values of parameters of the models, we show the log-log plots of the in-degree distribution.

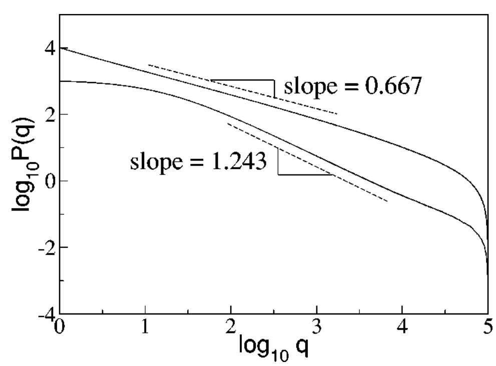

FIG. 3. Log-log plot of the cumulative distribution of a number of incoming links of sites for the considered models. For the first model, $\alpha=0.5$, $n=1$, $A=1.0$, and $c_0=1.0$. For the second model, $\alpha=0.5$, $n=1$, $B=0.15$, and $c_0=1.0$. The dashed lines have the slopes equal to the values of the scaling exponent γ obtained analytically minus 1. For better presentation, the dependences are displaced along the vertical axis.

The obtained values of the scaling exponents are within the error of the simulation from the corresponding ones found analytically. The values $\beta=1.46$ (1.5) are obtained from the simulation and analytically (in brackets) for the first model with the written out parameters, $\beta=0.85$ (0.804) are the corresponding values for the second model. $\gamma=1.70$ (1.667) and $\gamma=2.19$ (2.243) are the values of the critical exponent of the in-degree distribution obtained for the first and second models, relatively. One may see that the correspondence is really good.

The measured value of the scaling exponent of the distribution of incoming links in the World Wide Web is $\gamma=2.1$ [2,12,14] (as far as we know, any data on the exponent β are absent yet). As we have noted, one may assume reasonably that α is small in the real networks. We have shown that, in such a situation, the lower boundary for the possible values of γ is slightly below 2. We have demonstrated that, for $\gamma > 2$, the in-degree distribution has to be nonstationary if the growth of the network is accelerating. There are no data that indicate whether the degree distributions of the World Wide Web and the Internet are stationary or not. Our results make this question intriguing.

The World Wide Web is still in the initial stage of its evolution. Perhaps, the parameters of the accelerating growth will change. In this case, our answers demonstrate the possibility of changing of γ. We have shown that it may become even less than 2 in the future.

To demonstrate all of the existing possibilities, we have considered the models of growing networks with the particular rules of preferential attachment of new links. These models cover the range of possibilities but provide us only with particular values of the scaling exponents. Of course, there exist additional factors (aging [18] and dying [7,18] of sites, etc.) that may change these values.

In summary, we have studied the nonlinear, accelerating growth of the scale-free networks. We have demonstrated that it can be one of the most significant factors defining their structure. We have described the possible in-degree distributions of such networks and have fixed the lower boundary for the scaling exponent γ. Only a power-law time-dependence of the input flow of new links can keep the network inside of the class of scale-free networks. Nevertheless, we have found the region of parameters in which the scale-free structure is impossible. Our results demonstrate the possibility of quite different scenarios of network evolution.

S.N.D. thanks PRAXIS XXI (Portugal) for Research Grant No. PRAXIS XXI/BCC/16418/98. J.F.F.M. was partially supported by Project No. FCT-Sapiens 33141/99. We also thank A.N. Samukhin for many useful discussions.

[1] B.A. Huberman, P.L.T. Pirolli, J.E. Pitkow, and R.J. Lukose, Science **280**, 95 (1998).
[2] R. Albert, H. Jeong, and A.-L. Barabási, Nature (London) **401**, 130 (1999).
[3] B.A. Huberman and L.A. Adamic, Nature (London) **401**, 131 (1999).
[4] J. Lahererre and D. Sornette, Eur. Phys. J. B **2**, 525 (1998).
[5] S. Redner, Eur. Phys. J. B **4**, 131 (1998).
[6] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
[7] R. Albert, H. Jeong, and A.-L. Barabási, Nature (London) **406**, 378 (2000); R. Albert and A.-L. Barabási, Phys. Rev. Lett. **85**, 5234 (2000).
[8] L.A.N. Amaral, A. Scala, M. Barthelemy, and H.E. Stanley, Proc. Natl. Acad. Sci. U.S.A. **97**, 11149 (2000).
[9] C. Moore and M.E.J. Newman, Phys. Rev. E **61**, 5678 (2000).
[10] M.E.J. Newman, e-print cond-mat/0007214.
[11] P. Baran, *Introduction to Distributed Communications Networks*, RM-3420-PR, August 1964, http://www.rand.org/publications/RM/baran.list.html.
[12] R. Kumar, P. Raghavan, S. Rajagopalan, and A. Tomkins (unpublished).
[13] M. Faloutsos, P. Faloutsos, and C. Faloutsos, Comp. Comm. Rev. **29**, 251 (1999).
[14] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener (unpublished).
[15] A.-L. Barabási, R. Albert, and H. Jeong, Physica A **272**, 173 (1999).
[16] S.N. Dorogovtsev, J.F.F. Mendes, and A.N. Samukhin, Phys. Rev. Lett. **85**, 4633 (2000).
[17] P.L. Krapivsky, S. Redner, and F. Leyvraz, Phys. Rev. Lett. **85**, 4629 (2000).
[18] S.N. Dorogovtsev and J.F.F. Mendes, Phys. Rev. E **62**, 1842 (2000); Europhys. Lett. **52**, 33 (2000).
[19] A. Vázquez, e-print cond-mat/0006132.