A NEW CLASS OF SCALE FREE RANDOM GRAPHS

Zsolt Katona and Tamás F. Móri

Eötvös Loránd University, Budapest

13 May 2005

Consider the following modification of the Barabási–Albert random graph. At every step a new vertex is added to the graph. It is connected to the old vertices randomly, with probabilities proportional to the degree of the other vertex, and independently of each other. We show that the proportion of vertices of degree k decreases at the rate of k^{-3}. Furthermore, we prove a strong law of large numbers for the maximum degree.

1. The model

The examination of complex networks, such as the World Wide Web, led to the emerging role of random networks. Barabási and Albert [1] pointed out that many complex real world networks cannot be adequately described by the classical Erdős-Rényi random graph model, where the possible edges are included independently, with the same probability p. In this model the degree distribution is approximately Poisson with parameter np, while in real networks (for example the WWW) power law degree distributions have been observed. That is, the ratio of vertices with degree k is $P(k) \sim c \cdot k^{-\gamma}$ with some constant c and the γ exponent around 3. Here the parameter is independent of n.

Barabási and Albert [1] proposed a random graph model where they start with a single edge and at every step we add a new vertex and connect it to m of the old vertices by m edges. The old vertices are chosen randomly with probability proportional to their degrees. It was shown by Bollobás, Riordan, Spencer and Tusnády [2] that the proportion of vertices of degree k decreases at the rate k^{-3} as found in the empirical studies. In the most general model Cooper and Frieze [4] allow edges to be established between old vertices and the number of edges established in a step is determined randomly by a given distribution. They show that the degree distribution is power-law with exponent $\gamma > 2$.

However, they do not consider the case when edges from a new vertex are added independently. Consider the following modification of the Barabási–Albert random graph. At every step a new vertex is added to the graph. It is connected to the old

1991 *Mathematics Subject Classification.* 05C80, 60C05.

Key words and phrases. random graph, scale-free disribution.

Research supported by the Hungarian National Foundation for Scientific Research, Grant No. T-43037

vertices randomly, with probabilities proportional to the degree of the other vertex, and independently of each other. Since we are interested in asymptotic analysis, the initial configuration can be arbitrary, but, for the sake of simplicity, let us start from the very simple graph consisting of two points and the edge between them.

Let us number the vertices in the order of their creation; thus the vertex set of the graph after n steps is $\{0, 1, \ldots, n\}$. Let $X[n,k]$ and $Y[n,k]$ denote the number of vertices of degree exactly k and at least k, resp., after n steps. Thus, $X[n,0] + X[n,1] + \cdots = n+1$. Let $S_n = \sum_{k\ge 1} kX[n,k] = \sum_{k\ge 1} Y[n,k]$, the sum of degrees, or equivalently, the double of the number of edges. At the nth step an old vertex of degree k is connected to the new one with probability $\lambda k/S_{n-1}$. This quantity remains below 1, provided the proportionality coefficient λ is less than 2, which will therefore be assumed in the sequel.

Let $\mathcal{F}_n$ denote the σ-field generated by the first n steps. Let $\Delta[n,k]$ be the number of new edges into the set of old vertices of degree k at the nth step, and let $\Delta_n = \sum_{k\ge 1} \Delta[n,k]$ be the total number of new edges. Obviously, the conditional distribution of $\Delta[n+1,k]$ with respect to $\mathcal{F}_n$ is binomial with parameters $X[n,k]$ and $\lambda\,\dfrac{k}{S_n}$, hence $E\big(\Delta_{n+1} \bigm| \mathcal{F}_n\big) = \lambda$.

The aim of the present paper is to study some asymptotic properties of this random graph as the number of vertices tends to infinity. In Section 2 we prove a strong law of large numbers for the maximum degree, and in Section 3 it will be shown that the proportion of vertices of degree k converges a.s. to a constant, which, as a function of k, decreases in the order of k^{-3} as $k \to \infty$.

2. The maximum degree

First we deal with the asymptotics of S_n.

Theorem 2.1. $S_n = 2\lambda n + o\big(n^{1/2+\varepsilon}\big)$, $\forall \varepsilon > 0$.

Proof. With $\Delta_1 = 1$ define $\zeta_n = \sum_{j=1}^{n} (\Delta_j - \lambda) = S_n/2 - n\lambda$. Then $(\zeta_n, \mathcal{F}_n)$ is a square integrable martingale, and the increasing process associated with ζ_n^2 by the Doob decomposition is

$$A_n = \sum_{j=2}^{n} \operatorname{Var}\big(\Delta_j \bigm| \mathcal{F}_{j-1}\big) = \sum_{j=2}^{n-1} \sum_{k\ge 1} X[j,k]\,\frac{k\lambda}{S_j}\Big(1 - \frac{k\lambda}{S_j}\Big) \le n\lambda.$$

It is well known ([7], Proposition VII-**2**-4) that $\zeta_n = o\big(A_n^{1/2+\varepsilon}\big)$ a.e. on the event $A_n \to \infty$. This completes the proof. $\square$

Let us turn to $M_n = \max\{k : X[n,k] > 0\}$, the maximum degree after n steps.

Theorem 2.2. *We have* $\lim_{n\to\infty} M_n/\sqrt{n} = \mu$ *with probability* 1*, where the limit* μ *differs from zero with positive probability.*

Proof. We follow the same lines as in the proof of Theorem 3.1 of [6].

Let $W[n,j]$ denote the degree of vertex j after the nth step, with the initial values $W[n,j] = 0$ for $n < j$, $W[1,0] = W[1,1] = 1$, $W[j,j] = \Delta_j$. Then $M_n =$

$\max\{W[n,j] : j \geq 0\}$. Let us introduce the normalizing terms

$$c[n,k] = \prod_{i=1}^{n-1} \frac{S_i}{S_i+k\lambda}, \quad n \geq 1,\ k \geq 1.$$

For $n \to \infty$, with probability 1 we have

$$c[n,k] = \exp\biggl(-k\lambda \sum_{i=1}^{n-1} \frac{1}{S_i} + \frac{k^2\lambda^2}{2} \sum_{i=1}^{n-1} \frac{1+o(1)}{S_i^2}\biggr). \tag{2.1}$$

Since $\dfrac{1}{S_i} = \dfrac{1}{2\lambda i}\bigl(1 + o\bigl(i^{-1/2+\varepsilon}\bigr)\bigr)$, we obtain that the exponent in (2.1) differs from $-\dfrac{k}{2}\log n$ only by a term converging with probability 1. Thus $c[n,k] \sim \gamma_k\, n^{-k/2}$, with an appropriate positive random variable γ_k.

We clearly have that

$$E\bigl(W[n+1,j] \bigm| \mathcal{F}_n\bigr) = W[n,j] + \lambda\,\frac{W[n,j]}{S_n} = W[n,j]\,\frac{S_n+\lambda}{S_n}.$$

Hence

$$\bigl(Z[n;j,1] = c[n,1]\,W[n,j],\ \mathcal{F}_n\bigr), \quad n \geq \max\{j,1\}$$

is either a positive martingale or constant zero, thus it converges a.s. to some ζ_j. To estimate the moments of ζ_j, consider

$$Z[n;j,k] = c[n,k]\binom{W[n,j]+k-1}{k}.$$

Since $W[n+1,j] - W[n,j]$ is equal to either 1 or 0, we can write

$$\begin{aligned}
\binom{W[n+1,j]+k-1}{k} &= \\
&= \binom{W[n,j]+k-1}{k} + \bigl(W[n+1,j] - W[n,j]\bigr)\binom{W[n,j]+k-1}{k-1} = \\
&= \binom{W[n,j]+k-1}{k}\biggl(1 + \frac{k(W[n+1,j] - W[n,j])}{W[n,j]}\biggr).
\end{aligned}$$

By taking conditional expectation with respect to $\mathcal{F}_n$ we can see that

$$E\biggl(\binom{W[n+1,j]+k-1}{k} \biggm| \mathcal{F}_n\biggr) = \binom{W[n,j]+k-1}{k}\Bigl(1 + \frac{k\lambda}{S_n}\Bigr),$$

hence

$$\bigl(Z[n;j,k],\ \mathcal{F}_n\bigr), \quad n \geq \max\{j,1\}$$

is also a (convergent) martingale. Since $c[n,1]^k \leq c[n,k]$, we can majorize $Z[n;j,1]^k$ by $k!\,Z[n;j,k]$.

Now, $c[n,1]M_n = \max\{Z[n;j,1] : 0 \le j \le n\}$, being the maximum of an increasing number of nonnegative martingales, is a submartingale. The proof can be completed by showing that this submartingale is bounded in L_k, for some $k \ge 1$.

Let us start from the estimation

$$\begin{aligned} E\big(c[n,1]M_n\big)^k &= E\big(\max\{Z[n;j,1]^k : 0 \le j \le n\}\big) \le k! \sum_{j=0}^{n} EZ[n;j,k] = \\ = k!EZ[1;0,k] + k! \sum_{j=1}^{\infty} EZ[j;j,k] &= k! + k! \sum_{j=1}^{n} E\left(c[j,k]\binom{\Delta_j + k-1}{k}\right). \end{aligned} \tag{2.2}$$

Here

$$\begin{aligned} E\left(c[j,k]\binom{\Delta_j+k-1}{k}\right) &= E\left[E\left(c[j,k]\binom{\Delta_j+k-1}{k} \,\middle|\, \mathcal{F}_{j-1}\right)\right] = \\ &= E\left[c[j,k]E\left(\binom{\Delta_j+k-1}{k} \,\middle|\, \mathcal{F}_{j-1}\right)\right]. \end{aligned} \tag{2.3}$$

Next we show that, independently of j,

$$E\left(\binom{\Delta_j+k-1}{k} \,\middle|\, \mathcal{F}_{j-1}\right) \le E\binom{\pi+k-1}{k}, \tag{2.4}$$

where π stands for a Poisson(λ) random variable. Remembering that $\Delta_j = \Delta[j,0] + \dots + \Delta[j,j-1]$, we can write

$$\binom{\Delta_j+k-1}{k} = \sum_{\ell_0+\dots+\ell_j=k} \binom{\Delta[j,0]}{\ell_0} \cdots \binom{\Delta[j,j-1]}{\ell_{j-1}} \binom{k-1}{\ell_j}. \tag{2.5}$$

The binomial coefficients on the right-hand side are conditionally independent. The conditional distribution of each $\Delta[j,i]$ is binomial. Let ξ be a Binomial(n,p) random variable and η a Poisson one, with the same expectation. Then

$$E\binom{\xi}{\ell} = \binom{n}{\ell} p^\ell \le \frac{(np)^\ell}{\ell\,!} = E\binom{\eta}{\ell}.$$

Thus, if all random variables $\Delta[j,i]$ on the right-hand side of (2.5) are replaced by conditionally independent Poisson variables, the conditional expectation cannot decrease. Hence (2.4) follows.

By (2.2) and (2.3), for the L_k-boundedness of the submartingale $c[n,1]M_n$ it is sufficient to check that $\sum_{j=1}^{\infty} Ec[j,k] < \infty$. The convergence of $\sum_{j=1}^{\infty} c[j,k]$ when $k > 2$ is clear from the asymptotics obtained for $c[n,k]$, but the integrability does not follow immediately.

Let $k = 8$ and $N = \max\{n : S_n > 4\lambda n\}$. Then for $j > N$ we have

$$c[j,8] = \prod_{i=1}^{j-1}\Big(1 - \frac{8\lambda}{S_i + 8\lambda}\Big) \le \prod_{i=N+1}^{j-1}\Big(1 - \frac{2}{n+2}\Big) = \frac{(N+1)(N+2)}{j(j+1)},$$

but this is obviously true even for $j \le N$. Thus, for the finiteness of $\sum_{j=1}^{\infty} Ec[j,8]$ it is sufficient to prove that $EN^2 < \infty$.

By the usual large deviation arguments we have

$$P(N=n) \le P\big(S_n > 4\lambda n\big) = P\big(2^{S_n/2} > 4^{\lambda n}\big) \le E\big(2^{S_n/2}\big)4^{-\lambda n}. \tag{2.6}$$

Thus, we have to estimate the moment generating function of S_n. With $\Delta_1 = 1$ we can write $S_n/2 = \sum_{i=1}^{n} \Delta_i$, and

$$\begin{aligned} E\big(2^{S_n/2}\big) &= E\Big(E\big(2^{S_n/2} \mid \mathcal{F}_{n-1}\big)\Big) = E\Big(2^{S_{n-1}/2}E\big(2^{\Delta_n} \mid \mathcal{F}_{n-1}\big)\Big) = \\ &= E\left(2^{S_{n-1}/2}\prod_{j=1}^{n-1}\Big(1+\frac{\lambda j}{S_{n-1}}\Big)^{X[n-1,j]}\right) \le \\ &\le E\left(2^{S_{n-1}/2}\exp\left\{\sum_{j=1}^{n-1}\frac{\lambda j X[n-1,j]}{S_{n-1}}\right\}\right) = e^{\lambda}E\big(2^{S_{n-1}/2}\big). \end{aligned}$$

Therefore $E\big(2^{S_n/2}\big) \le e^{\lambda n}$, which, combined with (2.6), implies that $P(N=n) \le (e/4)^{\lambda n}$. Thus $EN^2 < \infty$, indeed. □

3. Degree distribution

In this section we prove that the degree distribution of our graph stabilizes almost surely, as $n \to \infty$, around a power law with exponent -3.

First we show that stochastic process of new vertices approaches a stationary regime. More precisely, the random variables Δ_n are asymptotically independent and asymptotically Poisson(λ) distributed. By LeCam's theorem on Poisson approximation [5] we have

$$\sum_{k=0}^{\infty}\left|P\big(\Delta_{n+1}=k \mid \mathcal{F}_n\big) - \frac{\lambda^k}{k!}\,e^{-\lambda}\right| \le 2\sum_{k=1}^{n} X[n,k]\Big(\frac{k\lambda}{S_n}\Big)^2 \le 2\lambda^2\frac{M_n}{S_n} = o(1). \tag{3.1}$$

Similarly, the conditional distribution of $\Delta[n+1,k]$ with respect to $\mathcal{F}_n$ is also asymptotically Poisson with parameter $\lambda k\,\dfrac{X[n,k]}{S_n}$.

Theorem 3.1. *For every $k = 0,1,\dots$ The proportion of vertices of degree k converges a.s. as $n \to \infty$:*

$$P\Big(\lim_{n\to\infty}\frac{X[n,k]}{n+1} = x_k\Big) = 1,$$

where

$$x_0 = p_0, \quad x_k = \frac{2}{k(k+1)(k+2)}\sum_{i=1}^{k} i(i+1)p_i, \quad p_k = \frac{\lambda^k}{k!}\,e^{-\lambda}.$$

Remark.

$$x_k \sim \frac{2\lambda(2+\lambda)}{k^3} \quad \text{as } k \to \infty.$$

Proof. First we will show by induction over k that

$$P\Bigl(\lim_{n\to\infty}\frac{Y[n,k]}{n+1}=y_k\Bigr)=1,$$

where the constants y_k satisfy the following recursion:

$$y_0=1,\quad y_k=\frac{k-1}{k+1}\,y_{k-1}+\frac{2q_k}{k+1},\quad q_k=p_k+p_{k+1}+\dots\;. \tag{3.2}$$

The sequence $X[n,0]$ obeys the strong law of large numbers, because [4], Corollary VII-2-6 implies that

$$X[n,0]=\sum_{i=1}^{n}I(\Delta_i=0)\sim\sum_{i=1}^{n}P\bigl(\Delta_i=0\bigm|\mathcal{F}_{i-1}\bigr)\sim ne^{-\lambda}.$$

Hence we obtain that

$$P\Bigl(\lim_{n\to\infty}\frac{Y[n,1]}{n+1}=1-e^{-\lambda}\Bigr)=1.$$

For the induction step suppose our assertion holds true for $k-1$. This time we introduce the normalizing factors

$$d[n,k]=\prod_{i=1}^{n-1}\biggl(1-I\bigl(S_i\ge 2(k-1)\bigr)\frac{(k-1)\lambda}{S_i}\biggr)^{-1}.$$

Their asymptotic behaviour can be treated similarly to what we have done in (2.1). Thus, with probability 1, we can write

$$d[n,k]=\exp\biggl((k-1)\lambda\sum_{i=1}^{n-1}I\bigl(S_i\ge 2(k-1)\bigr)\frac{1}{S_i}+\frac{(k-1)^2\lambda^2}{2}\sum_{i=1}^{n-1}\frac{1+o(1)}{S_i^2}\biggr).$$

By Theorem 2.1 we have $\dfrac{1}{S_i}=\dfrac{1}{2\lambda i}\bigl(1+o\bigl(i^{-1/2+\varepsilon}\bigr)\bigr)$, thus the exponent differs from $\dfrac{k-1}{2}\log n$ only by an a.s. convergent term. Therefore $d[n,k]\sim\delta_k\,n^{(k-1)/2}$, with some random variable $\delta_k>0$.

Since

$$Y[n+1,k]=Y[n,k]+\Delta[n+1,k-1]+I\bigl(\Delta_{n+1}\ge k\bigr),$$

it is easy to see that

$$E\bigl(d[n+1,k]\,Y[n+1,k]\bigm|\mathcal{F}_n\bigr)=d[n,k]\,Y[n,k]+b[n,k],$$

where

$$\begin{aligned}b[n,k]=d[n+1,k]\,(k-1)\lambda\,\frac{Y[n,k-1]}{S_n}+d[n+1,k]\,P\bigl(\Delta_{n+1}\ge k\bigm|\mathcal{F}_n\bigr)\le\\ \le d[n+1,k](\lambda+1).\end{aligned}$$

In addition, we have

$$\begin{aligned}
&\mathrm{Var}\big(d[n{+}1,k]\,Y[n{+}1,k] \bigm| \mathcal{F}_n\big) = \\
&\qquad = d[n{+}1,k]^2\,\mathrm{Var}\left(\Delta[n{+}1,k{-}1] + I\big(\Delta_{n+1} \ge k\big)\right) \le \\
&\qquad \le 2d[n{+}1,k]^2\left(\mathrm{Var}(\Delta[n{+}1,k{-}1]) + \mathrm{Var}\big(I(\Delta_{n+1} \ge k)\big)\right) \le \\
&\qquad \le 2d[n{+}1,k]b[n,k] = O\big(n^{k-1}\big).
\end{aligned} \tag{3.3}$$

Let us introduce a square integrable martingale by its differences

$$\xi_n = d[n,k]\,Y[n,k] - d[n{-}1,k]\,Y[n{-}1,k] - b[n{-}1,k].$$

The increasing process associated with the square of this martingale is

$$A_n = \sum_{i=1}^{n} E\big(\xi_i^2 \bigm| \mathcal{F}_{i-1}\big) = \sum_{i=1}^{n} \mathrm{Var}\big(d[i,k]Y[i,k] \bigm| \mathcal{F}_{i-1}\big),$$

which is of order $O(n^k)$ by (3.3). Hence

$$\sum_{i=1}^{n} \xi_i = d[n,k]\,Y[n,k] - \sum_{i=1}^{n} b[i{-}1,k] = o\big(n^{k/2+\varepsilon}\big).$$

¿From all these we obtain that

$$\frac{Y[n,k]}{n{+}1} = \frac{1}{(n{+}1)d[n,k]} \sum_{i=1}^{n} b[i{-}1,k] + o(1).$$

Now, by (3.1) and the induction hypothesis,

$$b[i,k] \sim d[i{+}1,k]\Big(\frac{k{-}1}{2}\,y_{k-1} + q_k\Big).$$

By applying the asymptotics we obtained for $b[i,k]$ we arrive at the formula

$$\frac{Y[n,k]}{n{+}1} = \frac{k{-}1}{k{+}1}\,y_{k-1} + \frac{2q_k}{k{+}1} + o(1),$$

which was to be proved.

Next we show that the solution of the recursion (3.2) is $y_k = \dfrac{2}{k(k{+}1)} \displaystyle\sum_{i=1}^{k} iq_i$.

Let $r_k = q_k + q_{k+1} + \dots$ and $z_k = y_k + y_{k+1} + \dots$, then from the recursion for y_k one can derive that $kz_k = (k{-}1)z_{k-1} + 2r_k$. From that we have

$$z_k = \frac{2}{k} \sum_{i=1}^{k} r_i,$$

which easily yields the above mentioned explicite form of y_k, and finally, the desired expression for x_k. $\square$

References

1. Barabási, A.-L., and Albert, R., *Emergence of scaling in random networks*, Science **286** (1999), 509–512.
2. Bollobás, B., Riordan, O., Spencer, J., and Tusnády, G., *The degree sequence of a scale-free random graph process*, Random Structures and Algorithms **18** (2001), 279–290.
3. Chow, Y. S., and Teicher, H., *Probability Theory: Independence, Interchangeability, Martingales*, Springer, New York, 1978.
4. Cooper, C., and Frieze, A., *A general model of web graphs*, Random Structures and Algorithms **22** (2003), 311–335.
5. LeCam, L., *An approximation theorem for the Poisson binomial distribution*, Pacific J. Math. **10** (1960), 1181–1197.
6. Móri, T. F., *The maximum degree of the Barabási–Albert random tree*, Combin. Probab. Comput. **14** (2005), 339–348.
7. Neveu, J., *Discrete-Parameter Martingales*, North-Holland, Amsterdam, 1975.

Department of Probability Theory and Statistics,
Eötvös Loránd University,
Pázmány Péter s. 1/C, Budapest, Hungary H-1117
E-mail address: `zskatona@cs.bme.hu`, `moritamas@ludens.elte.hu`