Annales Mathematicae et Informaticae
48 (2018) pp. 15–22
http://ami.uni-eszterhazy.hu

Scale-free property of the weights in a random graph model*

István Fazekas, Attila Perecsényi

Faculty of Informatics, University of Debrecen
fazekas.istvan@inf.unideb.hu
perecsenyi.attila@inf.unideb.hu

Submitted March 5, 2018 — Accepted September 13, 2018

Abstract

A new modification of the N interaction model [5], which based on the 3-interactions model of Backhausz-Móri [1]. This is a growing model, what evolves by weights. In every step N verticies will interact by form a star graph. We can choose vertices uniformly or according to their weights (preferential attachment). Our aim is to show asymptotic power-law distributions of the weights. The proofs are based on discrete time martingale methods. Numerical result is also presented.

Keywords: random graph, network, scale-free, power-law

MSC: 05C80, 60G42.

1. Introduction

Barabási and Albert [2] gave an explanation for the frequently observed phenomenon that many real-life networks are scale free, i.e., they have power-law degree distribution. To describe real-life networks such as the WWW, social and biological networks, they introduced a random graph model. They defined an evolving graph using the preferential attachment rule, what leads to scale-free graphs. Preferential attachment rule in a random graph model means, that when a new vertex is born, then the probability that the new vertex will be connected to an old vertex is proportional to the degree of the old vertex.

*Attila Perecsényi was supported through the New National Excellence Program of the Ministry of Human Capacities

In [4] a new network evolution model was introduced. In this paper, we shall study the same model. Consider an increasing sequence of weighted undirected graphs. The evolution of the graphs is based on creations of N-star subgraphs. Throughout the paper we call a graph N-star graph if N vertices form a star, that is it has one central vertex, what is connected with $N-1$ peripheral vertices. We start at time 0, and the initial graph is an N-star graph. This graph and all of its $(N-1)$-star subgraphs and all vertices have initial weights 1. Now we increase the size of the graph as follows. At each step N vertices interact with each other. It means that, we draw all non-existing edges between the peripheral vertices and the center vertex, so that, the vertices will form an N-star graph and the weights are increased by 1. The non-existing elements of the graph have weight 0.

We have two options in every step. On the one hand, with probability p, we add a new vertex, and it interacts with $N-1$ old vertices. On the other hand, with probability $1-p$, we do not add any new vertex, but N old vertices interact. Here $0 < p \leq 1$ is fixed.

When a new vertex is born, we have two possibilities again. With probability r, we choose an $(N-1)$-star graph according to to their weights (i.e. preferential attachment), and the new vertex is connected to its central vertex. Here preferential attachment means that the probability that we choose an $(N-1)$-star subgraph is proportional to its weight. With probability $1-r$, we choose $N-1$ old vertices uniformly at random and they will form an N-star graph with the new vertex, so that, the new vertex will be the center. Here uniform choice means that all subsets of vertices with cardinality $N-1$, have the same chance. Here $0 \leq r \leq 1$ is fixed.

In the other case, when we do not add any new vertex, we have two opportunities again. On the one hand, with probability q, we choose an old N-star graph according to their weights (i.e. preferential attachment). That is the chance of an N-star subgraph is proportional to its weight. Then we increase the weights inside the N-star subgraph chosen. On the other hand, with probability $1-q$, we choose uniformly N old vertices, and they form an N-star graph, so that, we choose the center out of the chosen N vertices uniformly. Here $0 \leq q \leq 1$ is fixed.

In [4] power law distribution of the weights of the vertices was shown. In this paper Theorem 2.1 shows that the weights of the N-stars have power law distribution. In the proof we use the Doob-Meyer decomposition and the method of [3].

2. Power law distribution of the weights of N-stars

Let $S(n, w)$ denote the number of N-stars with weight w, and let S_n denote the number of all N-stars after n steps. Furthermore, V_n denotes the number of vertices after n steps.

Theorem 2.1. *Let $0 < p < 1$ and $0 < q$. For all $w = 1, 2, \ldots$ we have*

$$\frac{S(n,w)}{S_n} \to s_w \tag{2.1}$$

almost surely as $n \to \infty$, where s_w, $w = 1, 2, \ldots$ are positive numbers satisfying the recurrence relation

$$s_1 = \frac{1}{h+1}, \qquad s_w = \frac{h(w-1)}{hw+1} s_{w-1}, \qquad \text{if } w > 1, \tag{2.2}$$

where $h = (1-p)q$. Moreover,

$$s_w \sim C w^{-\left(1+\frac{1}{h}\right)} \tag{2.3}$$

as $w \to \infty$, with $C = \frac{1}{h}\Gamma\left(1+\frac{1}{h}\right)$.

Proof. First we show that

$$\frac{S(n,w)}{n} \to k_w \tag{2.4}$$

almost surely as $n \to \infty$ for any fixed w. Here k_w, $w = 1, 2, \ldots$ are fixed nonnegative numbers.

We compute the conditional expectation of $S(n,w)$ with respect to $\mathcal{F}_{n-1}$ for $w \geq 1$. Let $S(n,0) = 0$ for all n. For $n, w \geq 1$ we have

$$\mathbb{E}\left(S(n,w)|\mathcal{F}_{n-1}\right) = p(n,w-1)S(n-1,w-1) + \left(1-p(n,w)\right)S(n-1,w) +$$
$$+\delta_{1,w}\left[p + (1-p)(1-q)\left(1 - \frac{S_{n-1}}{\binom{V_{n-1}}{N}N}\right)\right], \tag{2.5}$$

where

$$p(n,w) = (1-p)\left[q\frac{w}{n} + (1-q)\frac{1}{\binom{V_{n-1}}{N}N}\right]. \tag{2.6}$$

Let

$$c(n,w) = \prod_{i=1}^{n}\left(1-p(n,w)\right)^{-1}, \qquad w \geq 1. \tag{2.7}$$

It is easy to see that the above random variable is $\mathcal{F}_{n-1}$ measurable. Applying the Marcinkiewicz strong law of large numbers for the number of vertices, we have

$$V_n = pn + \mathrm{o}\left(n^{1/2+\varepsilon}\right) \tag{2.8}$$

almost surely, for any $\varepsilon > 0$.

Using (2.8) and the Taylor expansion for $\log(1+x)$ we obtain

$$\log c(n,w) = -\sum_{i=1}^{n}\log\left(1 - h\frac{w}{i} - \frac{(1-p)(1-q)}{\binom{V_{i-1}}{N}N}\right) = hw\sum_{i=1}^{n}\frac{1}{i} + \mathrm{O}(1),$$

where the error term is convergent as $n \to \infty$. It means

$$c(n,w) \sim h_w n^{hw} \tag{2.9}$$

almost surely as $n \to \infty$ and h_w is a positive random variable.

Let us consider the following process:

$$Z(n,w) = c(n,w)S(n,w) \quad \text{for } w \geq 1.$$

Here $\{Z(n,w), \mathcal{F}_n, n = 1,2,\dots\}$ is a nonnegative submartingale for any fixed $w \geq 1$. By the Doob-Meyer decomposition of $Z(n,w)$, we can write

$$Z(n,w) = M(n,w) + A(n,w)$$

where $M(n,w)$ is a martingale and $A(n,w)$ is a predictable increasing process. The general form of $A(n,w)$ is the following:

$$A(n,w) = \mathbb{E}Z(1,w) + \sum_{i=2}^{n} \left[\mathbb{E}\left(Z(i,w)|\mathcal{F}_{i-1}\right) - Z(i-1,w)\right]. \tag{2.10}$$

Now from (2.5) and (2.10), we have

$$A(n,w) = \mathbb{E}Z(1,w) + \sum_{i=2}^{n} c(i,w) \Bigg[p(i,w-1)S(i-1,w-1) + \\ + \delta_{1,w} \left(p + (1-p)(1-q) \left(1 - \frac{S_{i-1}}{\binom{V_{i-1}}{N} N} \right) \right) \Bigg]. \tag{2.11}$$

Let $B(n,w)$ be the sum of the conditional variances of $Z(n,w)$. In the following we give an upper bound for $B(n,w)$:

$$B(n,w) = \sum_{i=2}^{n} \mathbb{D}^2\left(Z(i,w)|\mathcal{F}_{i-1}\right) = \sum_{i=2}^{n} \mathbb{E}\{\left(Z(i,w) - \mathbb{E}\left(Z(i,w)|\mathcal{F}_{i-1}\right)\right)^2 |\mathcal{F}_{i-1}\} =$$

$$= \sum_{i=2}^{n} c(i,w)^2 \mathbb{E}\{\left(S(i,w) - \mathbb{E}\left(S(i,w)|\mathcal{F}_{i-1}\right)\right)^2 |\mathcal{F}_{i-1}\} \leq$$

$$\leq \sum_{i=2}^{n} c(i,w)^2 \mathbb{E}\{\left(S(i,w) - S(i-1,w)\right)^2 |\mathcal{F}_{i-1}\} \leq$$

$$\leq \sum_{i=2}^{n} c(i,w)^2 = \mathrm{O}\left(n^{2hw+1}\right). \tag{2.12}$$

Above we used that $c(n,w)$ is $\mathcal{F}_{i-1}$ measurable, (2.5) and the fact that, at each step only one N-star is involved in interaction.

We use induction on w. Let us consider the case when $w = 1$. From (2.9) and (2.11), we have

$$A(n,1) = \mathbb{E}Z(1,1) + \sum_{i=2}^{n} c(i,1) \left[p + (1-p)(1-q) \left(1 - \frac{S_{i-1}}{\binom{V_{i-1}}{N} N} \right) \right] \sim$$

$$\sim \sum_{i=2}^{n} h_1 n^h \left[p + (1-p)(1-q)\left(1 - \frac{S_{i-1}}{i^N}\right)\right] \sim h_1 \frac{n^{h+1}(1-h)}{h+1} \tag{2.13}$$

as $n \to \infty$. Using (2.12), we have

$$B(n,1) = \mathrm{O}\left(n^{2h+1}\right),$$

so

$$B(n,1)^{\frac{1}{2}} \log B(n,1) = \mathrm{O}\left(A(n,1)\right).$$

The conditions of Proposition VII-2-4 of [6] is fulfilled, so we have

$$Z(n,1) \sim A(n,1) \tag{2.14}$$

almost surely on the event $\{A(n,1) \to \infty\}$ as $n \to \infty$. So from (2.9), (2.13) and (2.14), we obtain

$$\frac{S(n,1)}{n} = \frac{Z(n,1)}{c(n,1)n} \sim \frac{A(n,1)}{c(n,1)n} \sim \frac{h_1 n^{h+1}(1-h)}{h_1 n^h n} = \frac{1-h}{1+h} = k_1 > 0, \tag{2.15}$$

as $n \to \infty$.

Let $w > 1$. Suppose that (2.4) is true for all weight less than w. Now from (2.8), (2.9) and (2.11), using the induction hypothesis, we obtain

$$A(n,w) = \mathbb{E}Z(1,w) + \sum_{i=2}^{n} \left(c(i,w)p(i,w-1)S(i-1,w-1)\right) \sim$$

$$\sim \sum_{i=2}^{n} h_w i^{hw} k_{w-1} i \left[h\frac{w-1}{i} + \frac{(1-p)(1-q)}{i^N}\right] \sim k_{w-1} h_w h(w-1) \frac{n^{wh+1}}{wh+1} \tag{2.16}$$

almost surely as $n \to \infty$. We see that the conditions of Proposition VII-2-4 are true, so we have $Z(n,w) \sim A(n,w)$. Therefore, from (2.9) and (2.16), we have

$$\frac{S(n,w)}{n} = \frac{Z(n,w)}{c(n,w)n} \sim \frac{A(n,w)}{c(n,w)n} \sim \frac{k_{w-1} h_w h(w-1)\frac{n^{wh+1}}{wh+1}}{h_w n^{wh} n} =$$

$$= k_{w-1} \frac{h(w-1)}{wh+1} = k_w. \tag{2.17}$$

Now we show that

$$\frac{S_n}{n} \to B, \tag{2.18}$$

almost surely as $n \to \infty$ where $B = 1 - h$.

First we compute the conditional expectation of S_n with respect to $\mathcal{F}_{n-1}$. We can see that the number of N-stars increases if and only if the number of N-stars of weight 1 increases, so we have

$$\mathbb{E}\{S_n | \mathcal{F}_{n-1}\} = S_{n-1} + p + (1-p)(1-q)\left(1 - \frac{S_{n-1}}{\binom{V_{n-1}}{N} N}\right) = \gamma_{n-1} S_{n-1} + B, \tag{2.19}$$

where

$$\gamma_{n-1} = 1 - (1-p)(1-q)\frac{1}{\binom{V_{n-1}}{N}N}.$$

Let

$$G_n = \prod_{i=1}^{n-1}(\gamma_i)^{-1}, \quad n \geq 1. \tag{2.20}$$

Here G_n is an $\mathcal{F}_{n-1}$ measurable random variable. Furthermore, let

$$Z_n = G_n S_n \quad \text{for } 1 \leq n. \tag{2.21}$$

From (2.19), we obtain

$$\mathbb{E}\{Z_n|\mathcal{F}_{n-1}\} = Z_{n-1} + BG_n. \tag{2.22}$$

We can see that $\{Z_n, \mathcal{F}_n, n = 1, 2, \ldots\}$ is a nonnegative submartingale. Applying again the Doob-Meyer decomposition for Z_n, we have

$$Z_n = M_n + A_n,$$

where M_n is a martingale and A_n is a predictable increasing process. From (2.10) and (2.22), we obtain

$$A_n = \mathbb{E}Z_1 + B\sum_{i=2}^{n} G_i. \tag{2.23}$$

By (2.8) and applying the Taylor expansion for $\log(1+x)$, we can give lower and upper bounds for G_i, so we obtain

$$C_1 n < A_n < C_2 n, \tag{2.24}$$

where C_1 and C_2 appropriate positive constants. Let B_n be the sum of the conditional variances of Z_n. In the following we give an upper bound for B_n:

$$B_n = \sum_{i=2}^{n} \mathbb{D}^2(Z_i|\mathcal{F}_{i-1}) = \sum_{i=2}^{n} \mathbb{E}\{(Z_i - \mathbb{E}(Z_i|\mathcal{F}_{i-1}))^2|\mathcal{F}_{i-1}\} =$$

$$= \sum_{i=2}^{n} G_i^2 \mathbb{E}\{(S_i - \mathbb{E}(S_i|\mathcal{F}_{i-1}))^2|\mathcal{F}_{i-1}\} \leq \sum_{i=2}^{n} G_i^2 \mathbb{E}\{(S_i - S_{i-1})^2|\mathcal{F}_{i-1}\} \leq$$

$$\leq \sum_{i=2}^{n} G_i^2 \leq C_3 n, \tag{2.25}$$

where C_3 is a positive constant. Above we used that G_i is $\mathcal{F}_{i-1}$ measurable and the fact that, at each step, at most one N-star can be born. Using (2.25), we have $B_n^{1/2}\log B_n = \mathrm{O}(A_n)$. From (2.24), we can see that $A_n \to \infty$ as $n \to \infty$, so applying Proposition VII-2-4 of [6], we obtain

$$Z_n \sim A_n \tag{2.26}$$

almost surely as $n \to \infty$.

Using (2.26) and (2.23), we have

$$\frac{K_n}{n} = \frac{Z_n}{G_n n} \sim \frac{A_n}{G_n n} = \frac{\mathbb{E}Z_1}{G_n n} + B \frac{1}{G_n} \frac{1}{n} \sum_{i=2}^{n} G_i \to B \tag{2.27}$$

almost surely.

Finally, from (2.4) and (2.18), we obtain

$$\frac{S(n,w)}{S_n} = \frac{S(n,w)}{n} \frac{n}{S_n} \to \frac{k_w}{B} = s_w \tag{2.28}$$

almost surely as $n \to \infty$. By using (2.28) for (2.15) and (2.17), we have the recurrence of s_w (cf. (2.2)). Applying several times (2.2), we obtain

$$s_w = s_1 \prod_{i=2}^{w} \frac{h(i-1)}{hi+1} = \frac{1}{h} \frac{(w-1)!}{\prod_{j=1}^{w} \left(j + \frac{1}{h}\right)} = \frac{1}{h} \frac{\Gamma(w)\Gamma\left(1+\frac{1}{h}\right)}{\Gamma\left(w+1+\frac{1}{h}\right)}. \tag{2.29}$$

Since $\sum_{w=1}^{\infty} s_w = 1$, the sequence $s_1, s_2, \ldots$ is a proper discrete probability distribution.

Now applying Stirling's formula for (2.29), we obtain the power law distribution (2.3). $\square$

3. Numerical result

In this section we present a numerical result. The 4-star model was generated with parameters $p = 0.5$, $q = 0.5$ and $r = 0.5$. We simulated $n = 10^5$ steps. To visualize the power law distribution we used log-log scale. Figure 1 shows that the weight distribution of 4-stars is indeed power law distribution.

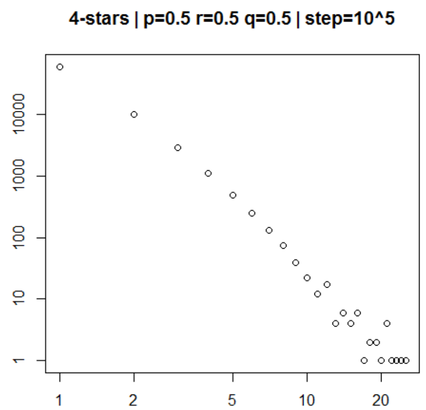

Figure 1: The weight distribution of 4-stars

References

[1] Backhausz Á., Móri T. F., Weights and degrees in a random graph model based on 3-interactions, *Acta Math. Hungar.*, Vol. 143(1) (2014), 23–43.

[2] Barabási A. L., Albert R., Emergence of scaling in random networks, *Science*, Vol. 286 (1999), 509–512.

[3] I. Fazekas, Cs. Noszály, A. Perecsényi, Weights of Cliques in a Random Graph Model Based on Three-Interactions, *Lithuanian Mathematical Journal*, 55 (2015), 207-221.

[4] I. Fazekas, Cs. Noszály, A. Perecsényi, The N-stars network evolution model, *in preparation.*

[5] Fazekas, I., Porvázsnyik, B., Scale-free property for degrees and weights in an N-interactions random graph model, *J. Math. Sci.* Vol. 214 (2016), 69–89.

[6] Neveu, J. Discrete-parameter martingales, *North-Holland, Amsterdam*, 1975.