УДК 621.892 *Original paper*

PREFERENTIAL ATTACHMENT WITH FITNESS DEPENDENT CHOICE

Y.A. Malyshkin
Tver State University, Tver, Russia
yury.malyshkin@mail.ru

DOI: 10.26456/pcascnn/2021.13.483

Abstract: We study the asymptotic behavior of the maximum degree in the preferential attachment tree model with a choice based on both the degree and fitness of a vertex. The preferential attachment models are natural models for complex networks (like neural networks, etc.) and constructed in the following recursive way. To each vertex is assigned a parameter that is called a fitness of a vertex. We start from two vertices and an edge between them. On each step, we consider a sample with repetition of d vertices, chosen with probabilities proportional to their degrees plus some parameter $\beta > -1$. Then we add a new vertex and draw an edge from it to the vertex from the sample with the highest product of fitness and degree. We prove that the maximum degree, dependent on parameters of the model, could exhibit three types of asymptotic behavior: sublinear, linear, and of $n/\ln n$ order, where n is the number of edges in the graph.

Keyword: complex networks, random graphs, preferential attachment, power of choice, fitness.

1. Introduction

The complex networks appear in numerous applications whose study requires analysis of big data (see, e.g. [1]). It is usually used to describe the structure of the internet and various social and communication networks (see, e.g. [2]). The other important application is different neural networks (see, e.g. [3]). It also could be used to model different nano-objects (see, e.g. [4, 5]). One of the properties of such a structure is that it could consist of nodes which have different properties that affects the development of the network. One of the basic models for a complex network is the preferential attachment model. To take into account the different properties of the nodes, one could introduce fitness into the model. It is also possible to add additional decision-making aspects to the model through the addition of choice. The standard preferential attachment graph model was introduced in [6]. The preferential attachment graph is constructed in the following way. First, we start with some initial graph G_1, usually, for simplification purposes, it consists of two vertices and an edge between them. Then on each step we add a new vertex and draw an edge from it to an already existing vertex chosen by some rule. For the standard preferential attachment model, the rule is that we choose a vertex with a probability proportional to its degree. Usually one considers the rule where we choose a vertex with probability proportional to its degree plus some parameter $\beta > -1$ (see, e.g. [7, 8]). Such a model was widely studied (see, e.g., [9], section 8) and different modifications have been introduced.

In the present work, we study the combination of two mentioned above modifications of this model. One of the modifications is the introduction of the fitness to the model (see, e.g., [10]). Fitness is a parameter that assigns to each vertex and affects its probability to be chosen at each step. The other modification is the introduction of a choice to the model (see, e.g., [11,12,13]). In this modification, we consider the sample of d independently chosen vertices and then choose one of them by some rule. Two types of rules have been considered: the degree base rule (see, e.g., [14,15]) and the location (or fitness) based choice (see, e.g., [16,17]). In the present work, we consider choice based on both degree and fitness.

Let us introduce our model. Fixing $\beta > -1$ and $d \in \mathbb{N}$, $d > 1$, we consider a sequence of graphs G_n build recursively as following. We start with the initial graph G_1 that consists of a two vertex v_0, v_1 and an edge between them. Graph G_{n+1} is built from G_n by adding a new vertex v_{n+1} and drawing an edge from it to the vertex of G_n, chosen by following rule. We first consider a sample of d vertices of G_n chosen independently from each other with probabilities proportional to their degrees plus parameter β (we would refer to the degree of vertex plus β as its weight). Then we choose vertex among them that maximize function $W(v,n) := \lambda_i \deg_{G_n} v$, where λ_i is the fitness of v_i. We consider case when λ_i are i.i.d. random variables that take two values, 1 and $\lambda > 1$ with non-zero probabilities. For simplification we also suggest that λ is not rational. So $W(v,n)$ would not take the same value on vertices with different fitness.

Let us formulate our main result. Let $M(n)$ be the maximum degree of vertices of G_n.

Theorem 1. In the defined model,

1. If $d < 2+\beta$ than for any $\epsilon > 0$

$$\mathbb{P}\left(\forall n > n_0 : n^{\frac{d}{2+\beta}-\epsilon} < M(n) < n^{\frac{d}{2+\beta}+\epsilon} \right) \to 1, \text{ as } n_0 \to \infty .$$

2. If $d = 2+\beta$, then almost surely

$$\liminf_{n\to\infty} \frac{M(n)\ln n}{n} \geq \frac{2d}{(d-1)\lambda},$$

$$\limsup_{n\to\infty} \frac{M(n)\ln n}{n} \leq \frac{2d}{d-1}.$$

3. If $d > 2+\beta$, then almost surely

$$\liminf_{n\to\infty} \frac{M(n)}{n} \geq \frac{x^*}{\lambda},$$

$$\limsup_{n\to\infty} \frac{M(n)}{n} \leq x^*,$$

where x^* is a unique positive root of equation $1-\left(1-\frac{x}{2+\beta}\right)^d = x$.

The result is similar to the Theorem 1.1 of [15] and shows that the addition of fitness to the choice from the sample does not affect the type of asymptotic of the maximal degree. Let us provide an outline of the proof. For each case, the lower and the upper bound is proven separately. The general idea is that we obtain either a lower or upper bound for the conditional probability to increase $M(n)$ (or its modification for the lower bound) as a function of $M(n)$. Then we study the properties of this function to get an estimate for $M(n)$. One of the key factors is the coefficient of the first term of the expansion by degrees of $M(n)/n$, which equals to $d/(2+\beta)$. In case $d<2+\beta$ we would analyze the fraction $M(n+1)/M(n)$ to prove sublinear behavior of $M(n)$. In case $d=2+\beta$ we would construct additional expressions to outline the second term of the expansion to get $n/\ln n$ bounds and for $d>2+\beta$, we would use a stochastic approximation to get linear estimates.

Let us give a short description of the stochastic approximation approach (see, e.g., [18,19] for more details). Process $Z(n)$ is a stochastic approximation process if it could be written as:

$$Z(n+1)-Z(n)=\gamma_n\left(F(Z(n))+E_n+R_n\right),$$

where γ_n, E_n and R_n satisfy the following condition. γ_n is not random and $\sum_{n=1}^{\infty}\gamma_n>\infty$, $\sum_{n=1}^{\infty}(\gamma_n)^2<\infty$, usually one puts $\gamma_n=1/n$ or $\gamma_n=1/(n+1)$. The term E_n is $\mathcal{F}_n$-measurable where $\mathcal{F}_n$ is the natural filtration of $Z(n)$, $\mathbb{E}(E_n\mid\mathcal{F}_n)=0$ and $\mathbb{E}((E_n)^2\mid\mathcal{F}_n)<c$ for some fixed constant c. We consider $E_n=\frac{1}{\gamma_n}(Z(n+1)-\mathbb{E}(Z(n+1)\mid\mathcal{F}_n))$ and therefore the function $F(x)$ could be found from representation $\mathbb{E}(Z(n+1)-Z(n)\mid\mathcal{F}_n)=\gamma_n(F(Z(n))+R_n)$ where R_n is a small error term that satisfies $\sum_{n=1}^{\infty}\gamma_n\mid R_n\mid<\infty$ almost surely. If these conditions hold (they could be easily checked in our case) then $Z(n)$ converges to the zero set of $F(x)$. In case when there is more than one eligible (nonnegative) root of $F(x)$ we would also prove non-convergence to one of the roots.

The other argument we would use is the persistent hub type of argument (see, e.g., [20] and Proposition 1.2 in [15]). This argument is based on the following urn model property. If we have random walk $(x(n),y(n))$ that takes steps up and right with probabilities $\frac{x(n)+\beta}{x(n)+y(n)+2\beta}$ and $\frac{y(n)+\beta}{x(n)+y(n)+2\beta}$ (it also represents the evolution of the urn with $x(n)$ white and $y(n)$ black balls in urn model, see e.g. Theorem 3.2 in [21] or Section 4.2 in [22]) then it converges in distribution to continuous Beta-distributions and hence one of the variables $x(n)$, $y(n)$ would exceed the other after some random moment. Also, if such process starts at point $(1,a)$, then probability that $x(n)$ exceeds $y(n)$ at some moment would decay exponentially with a. Hence, to prove the existence of the persistence hub it is enough to show that the pair ($\deg_{G_n} v_i$, $\deg_{G_n} v_j$) dominates urn model in a sense that it has a higher conditional probability to increase the degree of a vertex with a higher degree. We would use such an argument separately for different fitness.

2. Lower bounds

To prove the lower bound we separately consider maximum degrees among vertices with different fitness and estimate dynamics of the product of fitness and degree. Let $M_1(n)$ be the maximum degree among vertices with fitness 1 in G_n, $M_\lambda(n)$ be the maximum degree among vertices with fitness λ in G_n. Let $F_1(k,n)$ be the total weight of vertices with fitness 1 with degrees more than k in G_n, $F_\lambda(k,n)$ be the total weight of vertices with fitness λ with degrees more than k in G_n. Let $L_1(k,n)$ and $L_\lambda(k,n)$ be the number of vertices in G_n with degree k and fitness 1 and λ correspondingly. Consider functions:

$$f_n(x,y) := \left(1-\frac{y}{(2+\beta)n}\right)^d - \left(1-\frac{x+y}{(2+\beta)n}\right)^d,$$

$$g_n(x,y) := \sum_{k=0}^{d-1}\left(1-\frac{y}{(2+\beta)n}\right)^k \left(1-\frac{x+y}{(2+\beta)n}\right)^{d-1-k}$$

Note that $f_n(x,y)=\frac{x}{(2+\beta)n}g_n(x,y)$ and both functions are decreasing with y when $x,y>0$, $x+y\le(2+\beta)n$. Also, $f_n(x,y)$ is increasing with x while $g_n(x,y)$ is decreasing with x, and hence $\lambda f_n(x,y)\ge f_n(\lambda x,y)$.

Then:

$$\mathbb{E}\left(M_1(n+1)-M_1(n)\,|\,\mathcal{F}_n\right) = f_n\left(\left(M_1(n)+\beta\right)L_1\left(M_1(n),n\right), F_\lambda\left(M_1(n)/\lambda,n\right)\right),$$

$$\mathbb{E}\left(M_\lambda(n+1)-M_\lambda(n)\,|\,\mathcal{F}_n\right) = f_n\left(\left(M_\lambda(n)+\beta\right)L_\lambda\left(M_\lambda(n),n\right), F_1\left(\lambda M_\lambda(n),n\right)\right).$$

Note that either $F_1(\lambda M_\lambda(n),n)$ (if $M_1(n)<\lambda M_\lambda(n)$) or $F_\lambda(M_1(n)/\lambda,n)$ (if $M_1(n)>\lambda M_\lambda(n)$) equals to 0. Moreover, if $F_\lambda(M_1(n)/\lambda,n)\ne 0$ there is a vertex with fitness 1 and a degree of at least $\lambda M_\lambda(n)$. Let us define the process $X_n := \max\{M_1(n),\lambda M_\lambda(n)\}$. Note that $M(n)\ge\frac{X_n}{\lambda}$.

Then:

$$\mathbb{E}\left(X_{n+1}-X_n \mid \mathcal{F}_n\right) \geq \mathbf{1}\left\{M_\lambda(n)>M_1(n)/\lambda\right\}\lambda f_n\left(M_\lambda(n)+\beta,0\right)$$
$$+\mathbf{1}\left\{M_\lambda(n)<M_1(n)/\lambda\right\} f_n\left(M_1(n)+\beta,0\right)$$
$$\geq \mathbf{1}\left\{M_\lambda(n)>M_1(n)/\lambda\right\} f_n\left(\lambda\left(M_\lambda(n)+\beta\right),0\right)+\mathbf{1}\left\{M_\lambda(n)<M_1(n)/\lambda\right\} f_n\left(M_1(n)+\beta,0\right)$$
$$\geq f_n\left(X_n+\beta,0\right)=\frac{X_n+\beta}{(2+\beta)n} g_n\left(X_n+\beta,0\right) \geq \frac{X_n}{(2+\beta)n} g_n\left(X_n,0\right)-\frac{1}{n}.$$

Hence

$$\mathbb{E}\left(\frac{X_{n+1}}{X_n} \mid \mathcal{F}_n\right) \geq 1+\frac{1}{(2+\beta)n} g_n\left(X_n,0\right)-\frac{1}{nX_n}.$$

Note that $g_n(x_n,0)\to d$ if $\frac{x_n}{n}\to 0$. Also, $\prod_{k=1}^{n}\left(1+\frac{d}{(2+\beta)k}\right)$ is of order $n^{\frac{d}{2+\beta}}$. Let $\frac{d}{2+\beta}\leq 1$ (the argumentation in this case is similar to the case $\alpha+\gamma<1$ in [23]). Then, for any ε on the event $\left\{X_n<n^{\frac{d}{2+\beta}-\varepsilon}\right\}$ X_n would grow faster than $n^{\frac{d}{2+\beta}-\varepsilon/2}$ and hence for any $n_0\in\mathbb{N}$ with high probability at some time $n\geq n_0$ X_n would exceed $n^{\frac{d}{2+\beta}-\varepsilon}$. Also, due to standard large deviation estimates, if $X_{n_0}>(1-\delta)n_0^{\frac{d}{2+\beta}-\varepsilon}$ then probability that process X_n, $n\geq n_0$ would cross a line $(1-2\delta)n^{\frac{d}{2+\beta}-\varepsilon}$ before it crosses a line $n^{\frac{d}{2+\beta}-\varepsilon}$ does not exceed $ce^{-n_0^{\frac{d}{2+\beta}-\varepsilon}}$ for some $c=c(\delta)$. Therefore (for $\frac{d}{2+\beta}\leq 1$) with high probability $\liminf_{n\to\infty}\frac{X_n}{n^{\frac{d}{2+\beta}-\varepsilon}}\geq 1$.

Similarly, for $\frac{d}{2+\beta}>1$ with high probability $\liminf_{n\to\infty}\frac{X_n}{n}>0$. Also, we get that

$$\mathbb{E}\left(X_{n+1}-X_n \mid \mathcal{F}_n\right)=f_n\left(X_n,0\right)+O\left(\frac{1}{n}\right).$$

Hence, if we define $Z_n=\frac{X_n}{n}$, we would get that

$$\mathbb{E}\left(Z(n+1)-Z(n)\mid \mathcal{F}_n\right)\leq \frac{1}{n+1}\left(\mathbb{E}\left(X_{n+1}-X_n \mid \mathcal{F}_n\right)-Z_n\right)=\frac{1}{n+1}\left(f\left(Z_n\right)-Z_n+O\left(\frac{1}{n}\right)\right),$$

where $f(x)=1-(1-x)^d$.

Note that if $d>2+\beta$ then due to concavity of $f(x)$ equation $f(x)-x=0$ has two roots in $[0,2+\beta]$ (0 and a positive root x^*). Since Z_n does not converge to 0, by stochastic approximation we get that $\limsup_{n\to\infty} Z_n \ge x^*$ almost surely, which gives us the lower bound for $d > 2+\beta$.

3. Upper bounds

To prove the upper bounds, we first study dynamic of pairs ($\deg_{G_n} v_i$, $\deg_{G_n} v_j$) separately for vertices of fitness 1 and λ to prove the existence of persistence hub among vertices of each fitness. Then we would use it to remove terms $L_1(M_1(n),n)$ and $L_\lambda(M_\lambda(n),n)$ and get upper bounds for the increment of the maximum degree and prove upper bounds.

For any vertex v_i with fitness 1 we get for $n \ge i$

$$\mathbb{E}\left(\deg_{G_{n+1}}(v_i) - \deg_{G_n}(v_i) \mid \mathcal{F}_n\right) =$$

$$= \frac{f_n\left(\left(\deg_{G_n}(v_i)+\beta\right)L_1\left(\deg_{G_n}(v_i),n\right), F_1\left(\deg_{G_n}(v_i),n\right)+F_\lambda\left(\deg_{G_n}(v_i)/\lambda,n\right)\right)}{L_1\left(\deg_{G_n}(v_i),n\right)} =$$

$$= \frac{\deg_{G_n}(v_i)+\beta}{(2+\beta)n} g_n\left(\left(\deg_{G_n}(v_i)+\beta\right)L_1\left(\deg_{G_n}(v_i),n\right), F_1\left(\deg_{G_n}(v_i),\mathrm{n}\right)+F_\lambda\left(\deg_{G_n}(v_i)/\lambda,n\right)\right),$$

and for any vertex v_i with fitness λ we get for $n \ge i$:

$$\mathbb{E}\left(\deg_{G_{n+1}}(v_i) - \deg_{G_n}(v_i) \mid \mathcal{F}_n\right) =$$

$$= \frac{f_n\left(\left(\deg_{G_n}(v_i)+\beta\right)L_\lambda\left(\deg_{G_n}(v_i),n\right), F_1\left(\lambda\deg_{G_n}(v_i),\mathrm{n}\right)+F_\lambda\left(\deg_{G_n}(v_i),n\right)\right)}{L_\lambda\left(\deg_{G_n}(v_i),n\right)} =$$

$$= \frac{\deg_{G_n}(v_i)+\beta}{(2+\beta)n} g_n\left(\left(\deg_{G_n}(v_i)+\beta\right)L_\lambda\left(\deg_{G_n}(v_i),n\right), F_1\left(\lambda\deg_{G_n}(v_i),n\right)+F_\lambda\left(\deg_{G_n}(v_i),n\right)\right).$$

For vertices v_i and v_j with the same fitness let us estimate the probability to draw an edge to v_i conditioned on the event that edge is drawn to one of them. Let $\deg_{G_n} v_i > \deg_{G_n} v_j$ (if vertices have the same degree the probability is 1/2). Note that $g(x,y)$ is decreasing with y and $x+y$. Also, for vertices with fitness 1 we get that:

$$F_1\left(\deg_{G_n}(v_i),n\right)+F_\lambda\left(\deg_{G_n}(v_i)/\lambda,n\right) \le F_1\left(\deg_{G_n}(v_j),n\right)+F_\lambda\left(\deg_{G_n}(v_j)/\lambda,n\right),$$

$$F_1\left(\deg_{G_n}(v_i),n\right)+F_\lambda\left(\deg_{G_n}(v_i)/\lambda,n\right)+\left(\deg_{G_n}(v_i)+\beta\right)L_\lambda\left(\deg_{G_n}(v_i),n\right)$$
$$\le F_1\left(\deg_{G_n}(v_j),n\right)+F_\lambda\left(\deg_{G_n}(v_i)/\lambda,n\right)+\left(\deg_{G_n}(v_j)+\beta\right)L_\lambda\left(\deg_{G_n}(v_j),n\right).$$

Therefore

$$\mathbb{P}\left(\deg_{G_{n+1}}(v_i)-\deg_{G_n}(v_i)=1 \mid \mathcal{F}_n \deg_{G_{n+1}}(v_i)-\deg_{G_n}(v_i)+\deg_{G_{n+1}}(v_j)-\deg_{G_n}(v_j)=1\right)$$
$$\ge \frac{\deg_{G_n}(v_i)+\beta}{\deg_{G_n}(v_i)+\deg_{G_n}(v_j)+2\beta}.$$

Hence, after some random time N both $L_1(M_1(n),n)$ and $L_\lambda(M_\lambda(n),n)$ would be equal to 1.

Therefore, for $n>N$ we would get

$$\mathbb{E}(M_1(n+1)-M_1(n)\mid\mathcal{F}_n)=f_n((M_1(n)+\beta),F_\lambda(M_1(n)/\lambda,n))\le f_n((M_1(n)+\beta),0),$$

$$\mathbb{E}(M_\lambda(n+1)-M_\lambda(n)\mid\mathcal{F}_n)=f_n((M_\lambda(n)+\beta),F_1(\lambda M_\lambda(n),n))\le f_n((M_\lambda(n)+\beta),0).$$

Recall that if $d>2+\beta$, then equation $f(x)-x=0$ has two roots in $[0,2+\beta]$ (0 and a positive root x^*), and if $d\le 2+\beta$, then 0 is the only root in $[0,2+\beta]$.

Let us define $Z(n):=M(n)/n$. Then

$$\mathbb{E}(Z(n+1)-Z(n)\mid\mathcal{F}_n)\le\frac{1}{n+1}(f_n(M_n+\beta)-Z_n)=\frac{1}{n+1}\left(f(Z_n)-Z_n+O\left(\frac{1}{n}\right)\right).$$

For $d\le 2+\beta$ by stochastic approximation, we get $\limsup_{n\to\infty}Z_n=0$ almost surely. For $d>2+\beta$ by stochastic approximation, we get that $\limsup_{n\to\infty}Z_n\le x^*$ almost surely, which gives us an upper bound for $d>2+\beta$.

Now consider the case $d<2+\beta$. Similarly to the lower bound, we get that

$$\mathbb{E}\left(\frac{M(n+1)}{M(n)}\mid\mathcal{F}_n\right)\le 1+\frac{f_n(M(n)+\beta,0)}{M(n)}=1+\frac{g(M(n)+\beta,0)}{(2+\beta)n}\le 1+\frac{d}{(2+\beta)n},$$

where in the last inequality we used that $g(x,0)\le d$ for $x\ge 0$. Hence, as in the lower bound argument, we get that for any $\varepsilon>0$ $\limsup_{n\to\infty}\frac{M(n)}{n^{\frac{d}{2+\beta}+\varepsilon}}=0$, which gives us the upper bound for $d<2+\beta$.

To get the upper bound for the case $d=2+\beta$, for $c>0$ consider variables

$$U_n:=ne^{-c\frac{n}{M(n)}}.$$

We get that

$$\mathbb{E}\left(\frac{U_{n+1}}{U_n}\middle|\mathcal{F}_n\right)=\frac{n+1}{n}\mathbb{E}\left(e^{c\frac{n}{M(n)}-c\frac{n+1}{M(n+1)}}\middle|\mathcal{F}_n\right)=\left(1+\frac{1}{n}\right)\mathbb{E}\left(e^{\frac{cn(M(n+1)-M(n))}{M(n)M(n+1)}-\frac{c}{M(n+1)}}\middle|\mathcal{F}_n\right)$$

$$=\left(1+\frac{1}{n}\right)\mathbb{E}\left(1-\frac{c}{M(n+1)}+\frac{cn(M(n+1)-M(n))}{M(n)M(n+1)}+O\left(\left(\frac{1}{M(n+1)}\right)^2+\left(\frac{n}{(M(n+1))^2}\right)^2\right)\middle|\mathcal{F}_n\right)$$

$$=\left(1+\frac{1}{n}\right)\mathbb{E}\left(1-\frac{c}{M(n)}+\frac{cn(M(n+1)-M(n))}{(M(n))^2}+O\left(\left(\frac{1}{M(n)}\right)^2+\left(\frac{n}{(M(n))^2}\right)^2\right)\middle|\mathcal{F}_n\right)$$

$$= \left(1+\frac{1}{n}\right)\left(1-\frac{c}{M(n)}+\frac{cn\mathbb{E}\left(M(n+1)-M(n)\middle|\mathcal{F}_n\right)}{\left(M(n)\right)^2}+o\left(\frac{1}{n^{3/2}}\right)\right)$$

$$\le \left(1+\frac{1}{n}\right)\left(1-\frac{c}{M(n)}+\frac{cg_n\left(M(n)+\beta,0\right)}{(2+\beta)M(n)}+o\left(\frac{1}{n^{3/2}}\right)\right).$$

Note that $g_n(x,0)=d-\frac{d(d-1)}{2}\frac{x}{(2+\beta)n}+o\left(\frac{x}{n}\right)$ when $\frac{x}{n}\to 0$. Hence if $2+\beta=d$ we get

$$\mathbb{E}\left(\frac{U_{n+1}}{U_n}\middle|\mathcal{F}_n\right)\le 1+\frac{1}{n}-\frac{c}{M(n)}+\frac{c}{(2+\beta)M(n)}\left(d-\frac{d(d-1)}{2}\frac{M(n)}{(2+\beta)n}+o\left(\left(\frac{M(n)}{2}\right)^2\right)\right)+o\left(\frac{1}{n}\right)$$

$$=\left(1+\frac{1}{n}-\frac{c(d-1)}{2dn}+o\left(\frac{1}{n}\right)\right).$$

Hence, if $c>\frac{2d}{d-1}$ then U_n is supermartingale, and by Doob's theorem there is a random variable R, such that $\sup_n U_n<R$ almost surely. Hence for all (large enough) n

$$M(n)<\frac{cn}{\ln n-\ln R}$$

almost surely. Therefore

$$\liminf_{n\to\infty}\frac{\ln nM(n)}{n}\le\frac{2d}{d-1}.$$

almost surely.

4. Lower bound for $d=2+\beta$.

Let now prove lower bound for $d=2+\beta$. The argument is similar to the argument for the upper bound. Recall that $X_n=\max\{M_1(n),\lambda M_\lambda(n)\}$.

For $d/(2+\beta)=1$ and $c>0$ let consider

$$Y_n:=e^{c\frac{n}{X_n}}/n.$$

Note that

$$\mathbb{E}\left(\frac{Y_{n+1}}{Y_n}\middle|\mathcal{F}_n\right)=\frac{n}{n+1}\mathbb{E}\left(e^{c\frac{n+1}{X_{n+1}}-c\frac{n}{X_n}}\middle|\mathcal{F}_n\right)=\frac{n}{n+1}\mathbb{E}\left(e^{\frac{c}{X_{n+1}}-\frac{cn(X_{n+1}-X_n)}{X_nX_{n+1}}}\middle|\mathcal{F}_n\right)$$

$$=\frac{n}{n+1}\mathbb{E}\left(1+\frac{c}{X_{n+1}}-\frac{cn(X_{n+1}-X_n)}{X_nX_{n+1}}+O\left(\left(\frac{1}{X_{n+1}}\right)^2+\left(\frac{n}{X_{n+1}^2}\right)^2\right)\middle|\mathcal{F}_n\right)$$

$$= \frac{n}{n+1}\mathbb{E}\left(1+\frac{c}{X_n}-\frac{cn\left(X_{n+1}-X_n\right)}{X_n^2}+O\left(\left(\frac{1}{X_n}\right)^2+\left(\frac{n}{X_n^2}\right)^2\right)\middle|\mathcal{F}_n\right)$$

$$= \frac{n}{n+1}\left(1+\frac{c}{X_n}-\frac{cn\mathbb{E}\left(X_{n+1}-X_n\middle|\mathcal{F}_n\right)}{X_n^2}+o\left(\frac{1}{n^{3/2}}\right)\right)$$

$$\le \frac{n}{n+1}\left(1+\frac{c}{X_n}-\frac{cg_n\left(X_n,0\right)}{\left(2+\beta\right)X_n}+o\left(\frac{1}{n^{3/2}}\right)\right).$$

Recall that $g_n(x,0)=d-\frac{d(d-1)}{2}\frac{x}{(2+\beta)n}+o\left(\frac{x}{n}\right)$ when $\frac{x}{n}\to 0$. Hence, if $2+\beta=d$, we get

$$\mathbb{E}\left(\frac{Y_{n+1}}{Y_n}\middle|\mathcal{F}_n\right)\le 1-\frac{1}{n}+\frac{c}{X_n}-\frac{c}{X_n}\left(1-\frac{(d-1)X_n}{2dn}+o\left(\frac{X_n}{n}\right)\right)+o(1/\mathrm{n})=1-\frac{1}{n}+\frac{c(d-1)}{2dn}+o(1/\mathrm{n}).$$

Hence, if $c<\frac{2d}{d-1}$ then Y_n is supermartingale, by Doob's theorem there is a random variable R, such that $\sup_n Y_n<R$ almost surely. Hence for all n

$$X_n>\frac{cn}{\ln n+\ln R}$$

almost surely. Therefore

$$\liminf_{n\to\infty}\frac{\ln n\mathrm{X}_n}{n}\ge\frac{2d}{d-1}.$$

almost surely.

5. Conclusion

In the present paper, we considered a generalization of the linear preferential attachment model by adding a choice that is based on both degree and fitness of the vertices in the sample. The generalization provides a two-step process of building a preferential attachment graph, at first, we create a sample of vertices, and then we choose a vertex from that sample. Such a division is natural in terms of representing a decision-making process, when one considers few options and then makes a final decision. Based on the parameters of the model, we proved three types of asymptotic behavior of the maximum degree of the graph. The first is sublinear behavior when all nodes accumulate connection to other vertices over time with similar rates that in long terms vary only by a constant multiplier. The other is the linear behavior, when we have a concentration, i.e. one node accumulates the fraction of all edges so its degree behavior significantly differs from the behavior of other nodes. We also show a transition between these two behaviors. Such effects could be seen in numerous real networks such as neural networks, the internet, social and consumer networks, election models, and others.

The present work was funded by a grant from the Russian Science Foundation (project no. 19-71-00043). Исследование выполнено за счет гранта Российского научного фонда (проект № 19-71-00043).

References:

1. Latora V., Russo G., Nicosia V. *Complex networks: principles, methods and applications*. Cambridge, Cambridge University Press, 2017, 585 p. DOI: 10.1017/9781316216002.
2. Erciyes K. *Complex networks: an algorithmic perspective*. Boca Raton, CRC Press, 2014, 318 p. DOI: 10.13140/2.1.1608.1281.
3. Estrada E. *The structure of complex networks: theory and applications*. Oxford, Oxford University Press, 2012, XII, 465 p. DOI: 10.1093/acprof:oso/9780199591756.001.0001.
4. Fon W., Matheny M.H., Li J. et al. Complex dynamical networks constructed with fully controllable nonlinear nanomechanical oscillators, *Nano Letters*, 2017, vol. 17, issue 10, pp. 5977-5983. DOI: 10.1021/acs.nanolett.7b02026.
5. Yao H., Hsieh Y.-P., Kong J., Hofmann M. Modelling electrical conduction in nanostructure assemblies through complex networks, *Nature Materials*, 2020, vol. 19, issue 7, pp. 745-751. DOI: 10.1038/s41563-020-0664-1.
6. Barabási A., Albert R. Emergence of scaling in random networks, *Science*, 1999, vol. 286, issue 5439, pp. 509-512. DOI: 10.1126/science.286.5439.509.
7. Móri T.F. On random trees, *Studia Scientiarum Mathematicarum Hungarica*, 2002, vol. 39, issue 1-2, pp. 143-155. DOI: 10.1556/SScMath.39.2002.1-2.9.
8. Móri T.F. The maximum degree of the Barabási-Albert random tree, *Combinatorics, Probability and Computing*, 2005, vol. 14, issue 3, pp. 339-348. DOI: 10.1017/S0963548304006133.
9. Hofstag R. *Random graphs and complex networks*. Cambridge, Cambridge University Press, 2016, vol. 1: Cambridge Series in Statistical and Probabilistic Mathematics, Series Number 43, XVIII, 321 p. DOI: 10.1017/9781316779422.
10. Borgs C., Chayes J., Daskalakis C., Roch S. First to market is not everything: an analysis of preferential attachment with fitness, *STOC '07: Proceedings of the thirty-ninth annual ACM symposium on Theory of computing (San Diego, California, USA, June 11-13)*. New York: Association for Computing Machinery, 2007, pp. 135-144. DOI: 10.1145/1250790.1250812.
11. Krapivsky P.L., Redner S. Choice-driven phase transition in complex networks, *Journal of Statistical Mechanics: Theory and Experiment*, 2014, vol. 2014, art. no. P04021, 14 p. DOI: 10.1088/1742-5468/2014/04/P04021.
12. Malyshkin Y., Paquette E. The power of choice combined with preferential attachment, *Electronic Communications in Probability*, 2014, vol. 19, pp. 1-13. DOI: 10.1214/ECP.v19-3461.
13. Malyshkin Y., Paquette E. The power of choice over preferential attachment, *Latin American Journal of Probability and Mathematical Statistics*, 2015, vol. 12, no. 2, pp. 903-915.
14. Haslegrave J., Jordan J. Preferential attachment with choice, Random Structures and Algorithms, 2016, vol. 48, issue 4, pp. 751-766. DOI: 10.1002/rsa.20616.
15. Malyshkin Y. Preferential attachment combined with the random number of choices, *Internet Mathematics*, 2018, vol. 1, issue 1, pp. 1-25. DOI: 10.24166/im.01.2018.
16. Grauer A., Lüchtrath L., Yarrow M. Preferential attachment with location-based choice: Degree distribution in the noncondensation phase. Available at: https://arxiv.org/pdf/1905.08481v3.pdf (accessed 28.03.2021).
17. Haslegrave J., Jordan J., Yarrow M. Condensation in preferential attachment models with location-based choice, Random Structures and Algorithms, 2020, vol. 56, issue 3, pp. 775-795. DOI: 10.1002/rsa.20889.
18. Chen H.F. *Stochastic approximation and its applications. In: Nonconvex Optimization and its Applications.* New York, Boston, Dordrecht, London, Moscow, Kluwer Academic Publishers, 2002, vol. 64, XV, 360 p. DOI: 10.1007/b101987.
19. Pemantle R. A survey of random processes with reinforcement, *Probability Surveys*, 2007, vol. 4, pp. 1-79. DOI: 10.1214/07-PS094.
20. Galashin P.A. Existence of a persistent hub in the convex preferential attachment model, *Probability and Mathematical Statistics*, 2016, vol. 36, fasc. 1, pp. 59-74.
21. Mahmoud H.M. *Polya urn models*. New York, Chapman and Hall/CRC, 2008, 312 p. DOI: 10.1201/9781420059847.
22. Johnson N.L., Kotz S. *Urn models and their application: an approach to modern discrete probability theory*.

New York, London, Sydney, Toronto, John Wiley and Sons, 1977, xiii, 402 p.
23. Malyshkin Y. Sublinear preferential attachment combined with a growing number of choices, *Electronic Communications in Probability*, 2020, vol. 25, pp. 1-12. DOI: 10.1214/20-ECP368.

Библиографический список:

1. **Latora, V.** Complex networks: principles, methods and applications / V. Latora, G. Russo, V. Nicosia. –Cambridge: Cambridge University Press, 2017. – 585 p. DOI: 10.1017/9781316216002.
2. **Erciyes, K.** Complex networks: an algorithmic perspective / K. Erciyes. – Boca Raton: CRC Press, 2014. – 318 p. DOI: 10.13140/2.1.1608.1281.
3. **Estrada, E.** The structure of complex networks: theory and applications / E. Estrada. – Oxford: Oxford University Press, 2012. – XII, 465 p. DOI: 10.1093/acprof:oso/9780199591756.001.0001.
4. **Fon, W.** Complex dynamical networks constructed with fully controllable nonlinear nanomechanical oscillators / W. Fon, M.H. Matheny, J. Li et al. // Nano Letters. – 2017. – V. 17. – I. 10. – P. 5977-5983. DOI: 10.1021/acs.nanolett.7b02026.
5. **Yao, H.** Modelling electrical conduction in nanostructure assemblies through complex networks / H. Yao, Y.-P. Hsieh, J. Kong, M. Hofmann // Nature Materials. – 2020. – V. 19. – I. 7. – P. 745-751. DOI: 10.1038/s41563-020-0664-1.
6. **Barabási, A.** Emergence of scaling in random networks / A. Barabási, R. Albert // Science. – 1999. – V. 286. – I. 5439. – P. 509-512. DOI: 10.1126/science.286.5439.509.
7. **Móri, T.F.** On random trees / T.F. Móri // Studia Scientiarum Mathematicarum Hungarica. – 2002. – V. 39. – I. 1-2. – P. 143-155. DOI: 10.1556/SScMath.39.2002.1-2.9.
8. **Móri, T.F.** The maximum degree of the Barabási-Albert random tree / T.F. Móri // Combinatorics, Probability and Computing. – 2005. – V. 14. – I. 3. – P. 339-348. DOI: 10.1017/S0963548304006133.
9. **Hofstag, R.** Random graphs and complex networks / R. van der Hofstag. – Cambridge: Cambridge University Press, 2016. V. 1: Cambridge Series in Statistical and Probabilistic Mathematics, Series Number 43. – XVIII, 321 p. DOI: 10.1017/9781316779422.
10. **Borgs, C.** First to market is not everything: an analysis of preferential attachment with fitness / C. Borgs, J. Chayes, C. Daskalakis, S. Roch // STOC '07: Thirty-ninth annual ACM symposium on theory of computing, June 11-13, San Diego, California, USA: proceedings. – New York: Association for Computing Machinery, 2007. – P. 135-144. DOI: 10.1145/1250790.1250812.
11. **Krapivsky, P.L.** Choice-driven phase transition in complex networks / P.L. Krapivsky, S. Redner // Journal of Statistical Mechanics: Theory and Experiment. – 2014. – V. 2014. – Art. № P04021. – 14 p. DOI: 10.1088/1742-5468/2014/04/P04021.
12. **Malyshkin, Y.** The power of choice combined with preferential attachment / Y. Malyshkin, E. Paquette // Electronic Communications in Probability. – 2014. – V. 19. – P. 1-13. DOI: 10.1214/ECP.v19-3461.
13. **Malyshkin, Y.** The power of choice over preferential attachment / Y. Malyshkin, E. Paquette // Latin American Journal of Probability and Mathematical Statistics. – 2015. – V. 12. – № 2. – P. 903-915.
14. **Haslegrave, J.** Preferential attachment with choice. / J. Haslegrave, J. Jordan // Random Structures and Algorithms. – 2016. – V. 48. – I. 4. – P. 751-766. DOI: 10.1002/rsa.20616.
15. **Malyshkin, Y.** Preferential attachment combined with the random number of choices / Y. Malyshkin // Internet Mathematics. – 2018. – V. 1. – I. 1. – P. 1-25. DOI: 10.24166/im.01.2018.
16. **Grauer, A.** Preferential attachment with location-based choice: Degree distribution in the noncondensation phase / A. Grauer, L. Lüchtrath, M. Yarrow. – Access mode www.url: https://arxiv.org/pdf/1905.08481v3.pdf. – 28.03.2021.
17. **Haslegrave, J.** Condensation in preferential attachment models with location-based choice / J. Haslegrave, J. Jordan, M. Yarrow // Random Structures and Algorithms. – 2020. – V. 56. – I. 3. – P. 775-795. DOI: 10.1002/rsa.20889.
18. **Chen, H.F.** Stochastic approximation and its applications / H.F. Chen. In: Nonconvex Optimization and its Applications. – New York, Boston, Dordrecht, London, Moscow: Kluwer Academic Publishers. – 2002. – V. 64. – XV, 360 p. DOI: 10.1007/b101987.
19. **Pemantle, R.** A survey of random processes with reinforcement / R. Pemantle // Probability Surveys. – 2007. – V. 4. – P. 1-79. DOI: 10.1214/07-PS094.
20. **Galashin, P.A.** Existence of a persistent hub in the convex preferential attachment model / P.A. Galashin // Probability and Mathematical Statistics. – 2016. – V. 36. Fasc. 1. – P. 59-74.
21. **Mahmoud, H.M.** Polya urn models / H.M. Mahmoud. – New York: Chapman and Hall/CRC, 2008. – 312 p. DOI: 10.1201/9781420059847.
22. **Johnson, N.L.** Urn models and their application: an approach to modern discrete probability theory /

N.L. Johnson, S. Kotz. New York, London, Sydney, Toronto: John Wiley and Sons, 1977. – xiii, 402 p.
23. **Malyshkin, Y.** Sublinear preferential attachment combined with a growing number of choices / Y. Malyshkin // Electronic Communications in Probability. – 2020. – V. 25. – P. 1-12. DOI: 10.1214/20-ECP368.

Оригинальная статья

ПРЕДПОЧТИТЕЛЬНОЕ ПРИСОЕДИНЕНИЕ С ВЫБОРОМ, ЗАВИСЯЩИМ ОТ ПРИГОДНОСТИ

Ю.А. Малышкин
ФГБОУ ВО «Тверской государственный университет»
170002, Россия, Тверь, Садовый пер., 35

DOI: 10.26456/pcascnn/2021.13.483

Аннотация. Исследуется асимптотическое поведение максимальной степени вершины в графе предпочтительного присоединения с выбором вершины, основанном как на ее степени, так и на дополнительном параметре (пригодности). Модели предпочтительного присоединения широко используются для моделирования сложных сетей (таких как нейронные сети и т.д.). Они строятся следующим образом. Мы начинаем с двух вершин и ребра между ними. Затем на каждом шаге мы рассматриваем выборку из уже существующих вершин, выбранных с вероятностями, пропорциональными их степеням плюс некоторый параметр $\beta > -1$. Затем мы добавляем новую вершину и соединяем ее ребром с вершиной из выборки, на которой достигается максимум произведения ее степени на ее пригодность. Мы доказали, что в зависимости от параметров модели возможны три типа поведения максимальной степени вершины – сублинейное, линейное и порядка $n/\ln n$, где n – число вершин в графе.

Ключевые слова: сложные сети, случайные графы, предпочтительное присоединение, случайный выбор, пригодность.

Малышкин Юрий Андреевич – к.ф.-м.н., доцент кафедры информационных технологий ФГБОУ ВО «Тверской государственный университет»
Yury A. Malyshkin – Ph. D., Docent, Applied Mathematics and Cybernetics Department, Tver State University

Поступила в редакцию/received: 01.03.2021; после рецензирования/revised: 11.03.2021; принята/accepted 12.03.2021.