MATHEMATICA APPLICANDA
Vol. 43(2) 2015, p. 115–121
doi: 10.14708/ma.v43i2.425

Krzysztof Choromański (New York)
Jacek Miękisz* (Warsaw)

Scale-free graphs with edge deletion

Abstract We extend the classical Barabási-Albert preferential attachment procedure by allowing edge deletion. We show that unlike in the original model, power-law exponents of degree distribution of scale-free graphs with edge deletion depend on the number of attached edges in one step of the growing process.

2010 Mathematics Subject Classification: 05C80; 05C82; 91D30.

Key words and phrases: random graphs, scale-free graphs, power laws, preferential attachment, edge deletion.

1. Introduction. Many systems of interacting objects or individuals in natural and social sciences can be described by complex networks [1–5]. The structure of neighborhoods may have a quite complex topology resulting from various random processes which describe mechanisms of growing networks. To mimic the "the rich get richer" rule, Barabási and Albert used the preferential attachment rule in growing their networks [2, 6, 7]. It says that a new vertex is linked with already existing ones with a probability proportional to their degrees. Such a procedure leads to a scale-free network with a power-law degree distribution, $P(k) \sim k^{-3}$. This was heuristically understood in [2, 6, 7] and proved mathematically in [8, 9], see also [10] for an introduction to random graphs.

Since then there were proposed many generalizations and extensions of the preferential attachment procedure. In particular, graphs with internal vertex structure given by weights of vertices were considered in [11]. In that model, weight dynamics depends on the current vertex degree distribution and the preferential attachment procedure takes into account both weights and degrees of vertices. It was proved that such a coupled dynamics leads to scale-free graphs with exponents depending on parameters of the weight dynamics.

In many real situations, edges are not only created but they may also be destroyed [12–14]. Here we assume that edges are deleted by a preferential detachment. This describes situations where vertices with high degrees

* JM would like to thank the Ministry of Science and Higher Education for a financial support under the grant N201 362536

are more dynamic, they are more likely to change their connections. Similar model was discussed in [12] where it was shown that power-law exponents depend on the probability of deleting an edge in any time step. In our model we also assume that two vertices which were forced to be disconnected form an edge between themselves. We use a method of "master equations" and show (under the assumption of linear evolution) that our procedure leads to scale free graphs with power-law exponents depending on the number of attached edges in one step of the growing process. This is a novel behaviour not present in the original Barabási-Albert model.

2. Model. We will now define precisely our discrete-time dynamical model. At time $t = m_0$, the graph consists of m_0 vertices and $(m-1)m_0$ edges. At any time $t+1$, we have two sub-steps. At a first sub-step of our procedure, we choose two vertices v_1 and v_2 with the probability proportional to their degrees and then we choose their neighbors w_1 and w_2 with the uniform probability. If vertices (w_1, w_2) are different, then we delete edges (v_1, w_1) and (v_2, w_2), and finally we add an edge (w_1, w_2) (in this way we may create multiple edges), otherwise we repeat the procedure. In the second sub-step, we follow the original preferential attachment procedure, that is we create a new vertex and link it by $m > 2$ edges with existing vertices with the probability proportional to their degrees (we do not take into account changes made in the first sub-step). It is easy to see that at any time $t \geqslant m_0$, our graph has t vertices and $(m-1)t$ edges.

Let $N_k(t)$ be the expected number of vertices of degree k at time t. It follows that

$$\begin{aligned} N_k(t+1) = \quad & N_k(t) - \tfrac{2kN_k(t)}{2(m-1)t} + \tfrac{2(k+1)N_{k+1}(t)}{2(m-1)t} - \\ & \tfrac{mkN_k(t)}{2(m-1)t} + \tfrac{m(k-1)N_{k-1}(t)}{2(m-1)t}. \end{aligned}$$

We assume that like in the original model, the graph evolves in the linear way (we expect this but we are short of a rigorous proof), that is for every k, $\frac{N_k}{t} \to n_k$ when $t \to \infty$. The rates of linear evolution, n_k, satisfy the following linear equations:

$$n_k = \frac{k+1}{m-1} n_{k+1} + \frac{(k-1)m}{2(m-1)} n_{k-1} - \frac{k(m+2)}{2(m-1)} n_k \tag{1}$$

Let us notice that $\sum_{k=0}^{M} n_k = \lim_{t\to\infty} \frac{\sum_{k=0}^{M} N_k(t)}{t} \leqslant 1$ for every M so the sequence n_k is summable. We will prove that that the degree distribution n_k follows the power law with the exponent depending on m.

THEOREM 2.1 *The distribution of vertex degrees in the preferential attachment model with edge deletion satisfies the power law, that is $n_k k^{\beta} \to c$ when $k \to \infty$, for some positive constant c, where $\beta = \frac{3m-4}{m-2}$, $m > 2$.*

A proof of the Theorem is based on the following Lemma.

Lemma 2.2 *If a sequence* n_k*,* $k = 0, 1, ...$ *of positive real numbers satisfies the following recurrence equations:*

$$\frac{n_{k-1}}{n_k} = 1 + \frac{\beta}{k} + r_k,$$

where the sequence r_k *is summable, that is* $\sum_i r_i < \infty$*, then for some constant* c *we have*

$$n_k k^{\beta} \to c \quad \textit{when} \quad k \to \infty$$

Proof We write $1/n_k$ as a product

$$\frac{1}{n_k} = \frac{1}{n_1}\frac{n_1}{n_2}\frac{n_2}{n_3}...\frac{n_{k-1}}{n_k}$$

hence

$$\frac{1}{n_k} = \frac{1}{n_1}\prod_{j=2}^{k}(1 + \frac{\beta}{j} + r_j) = \frac{1}{n_1}\prod_{j=2}^{k}(1 + \frac{\beta}{j})\prod_{j=2}^{k}(1 + r_j'), \tag{2}$$

where $r_j' = \frac{r_j}{1+\beta/j}$.

Analogous equation is satisfied by the sequence $n_k = k^{-\beta}$ which satisfies the assumption of the Lemma. To see that the second product has a limit, we take the logarithm of it and use that $1 + x \leqslant e^x$ and the summability of r_k. We then multiply Eq. 2 by $n_k = k^{-\beta}$ and the Lemma is proved. ■

Proof (of Theorem 2.1) Now we would like to transform Eq. 1 into the expression present in the Lemma, that is we have to eliminate n_{k+1} . In the identity $n_{k+1} = n_k + n_{k+1} - n_k$ we set $n_{k+1} - n_k = n_k - n_{k-1} + d_k$ hence $n_{k+1} = 2n_k - n_{k-1} + d_k$. We put this expression for n_{k+1} in Eq. 1, divide Eq. 1 by n_k, and after some rearrangements we get

$$\frac{n_{k-1}}{n_k} = 1 + \frac{3m-4}{(m-2)k-(m+2)} + \frac{2(k+1)}{(m-2)k-(m+2)}\frac{d_k}{n_k}$$

Now we have to prove that the sequence

$$\frac{d_k}{n_k} = \frac{n_{k+1}}{n_k} + \frac{n_{k-1}}{n_k} - 2.$$

is summable and then the Theorem would follow from the Lemma.

Let us denote $p_k = \frac{n_{k-1}}{n_k}$. Then from Eq. 1 we get the following recurrence formula for p_k:

$$p_k = \frac{m+2}{m} + \frac{3}{k-1} - \frac{2(k+1)}{m(k-1)p_{k+1}}. \tag{3}$$

We will show that $p_k \geqslant 1$ for every k. We first observe that from $p_{k+1} \geqslant 1$ it follows that $p_k \geqslant 1$. Indeed,

$$p_k = \begin{array}{l} \frac{m+2}{m} + \frac{3}{k-1} - \frac{2(k+1)}{m(k-1)p_{k+1}} \geqslant \\ \frac{m+2}{m} + \frac{3}{k-1} - \frac{2(k+1)}{m(k-1)} = \\ 1 + \frac{3}{k-1} - \frac{4}{m(k-1)} \geqslant 1. \end{array}$$

If there would exists k such that $p_k < 1$, then $p_{k+i} < 1$ for every $i > 0$ and hence $n_j > n_{j-1}$ for every $j \geqslant k$ and n_k would not be summable. Hence we showed that $p_k \geqslant 1$ for every k and in consequence $n_{k+1} < n_k$ for every k.

Now after a series of simple transformations we get from Eq. 3 that

$$(p_{k+2} - p_{k+1}) = \frac{mp_{k+1}p_{k+2}}{2 + \frac{4}{k}}(p_{k+1} - p_k) + O(\frac{1}{k^2}).$$

From the fact that p_k is bounded and $\frac{mp_{k+1}p_{k+2}}{2+\frac{4}{k}} > \frac{15}{14}$ for $k > 4$ it follows that

$$p_{k+1} - p_k = O(\frac{1}{k^2})$$

so p_k converges and it follows from Eq. 3 that it converges to 1. Hence we get that

$$\frac{n_{k+1}}{n_k} + \frac{n_{k-1}}{n_k} - 2 = \frac{1}{p_{k+1}} + p_k - 2 = O(\frac{1}{k^2}).$$

This shows that the sequence is d_k/n_k is summable which proves the Theorem. ■

Now, one may use inequalities concerning sums of random variables [15, 16], see [9], to show that the number of vertices of a given degree is concentrated around its expected value. More precisely, let $Z(k,t)$ be the number of vertices of degree k at time t. It can be shown that $Z(k,t)/t$ converges in probability to n_k as t tends to infinity.

3. Conclusion In summary, we introduced a model of growing scale-free graphs with edge deletion. We showed (under the assumption of linear evolution) that unlike in the standard preferential attachment procedure, the power-law exponent of the degree distribution depends upon the number of added edges at every step.

References

[1] D. J. Watts. Small Worlds: The Dynamics of Networks between Order and Randomness. Princeton University Press, Princeton (1999)

[2] R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. Rev. Mod. Phys. **74**: 47-97 (2002) doi: 10.1103/RevModPhys.74.47

[3] S. N. Dorogovtsev and J F. F Mendes. Evolution of Networks: From Biological Nets to the Internet and WWW. Oxford University Press, Oxford (2003)

[4] S. Boccalettia, V. Latorab, Y. Moreno, M. Chavez, and D.-U. Hwanga. Complex networks: Structure and dynamics. Phys. Rep. **424**: 175-308 (2006) doi: 10.1016/j.physrep.2005.10.009

[5] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes. Critical phenomena in complex networks. Rev. Mod. Phys. **80**: 1275-1335 (2008) doi: 10.1103/RevModPhys.80.1275

[6] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. Science **286**: 509-512 (1999) doi: 10.1126/science.286.5439.509

[7] A.-L. Barabási, R. Albert, and H. Jeong. Mean-field theory for scale-free random networks. Physica A **272**: 173-187 (1999) doi: 10.1016/S0378-4371(99)00291-5

[8] F. Chung and L. Lu. Complex Graphs and Networks. American Mathematical Society (2006).

[9] R. Durrett. Random Graph Dynamics. Cambridge University Press (2007).

[10] S. Janson, T. Łuczak, and A. Ruciński. Random Graphs, Wiley, New York (2000).

[11] K. Choromański, M. Matuszak, and J. Miękisz. Scale-free graph with preferential attachment and evolving internal vertex structure. J. Stat. Phys. **151**: 1175–1183 (2013) doi: 10.1007/s10955-013-0749-1

[12] M. Deifen and M. Lindholm. Growing networks with preferential deletion and addition of edges. Physica A **388**: 4297-4303 (2009) doi: 10.1016/j.physa.2009.06.032

[13] P. Prałat and C. Wang. An edge deletion model for complex networks. Theoretical Computer Science **412**: 5111-5120 (2011) doi: 10.1016/j.tcs.2011.05.016

[14] H. Brot, M. Honig, L. Muchnik, J. Goldenberg, and Y. Louzoun. Edge removal balances preferential attcahment and triad closing. Phys. Rev. E **80**: 042815 (2013) doi: 10.1103/PhysRevE.88.042815

[15] W. Hoeffding. Probability inequalities for sums of bounded random variables. J. Am. Stat. Soc. **58**: 13-30 (1963) doi: 10.1080/01621459.1963.10500830

[16] K. Azuma. Weighted sums of certain dependent random variables. Tohoku Math. J. **3**: 357-367 (1967) doi: 10.2748/tmj/1178243286

Bezskalowe grafy z usuwaniem krawędzi

Krzysztof Choromański i Jacek Miękisz

Streszczenie Praca rozszerza klasyczny model Barabasiego-Alberty o możliwość usuwania krawędzi. Pokazano, że wykładnik w prawie potęgowym rozkładu stopni wierzchołków zależy od liczby krawędzi dodawanych w każdym kroku procesu budowy grafu.

2010 Klasyfikacja tematyczna AMS (2010): Pierwotna: 05C80; Wtórna: 05C82; 91D30.

Słowa kluczowe: grafy losowe, prawo potęgowe, preferencyjne przyłączenia, usuwanie krawędzi .

Krzysztof Choromanski joined the IEOR department in September 2009 as a first-year PH.D student. He completed his master degree in mathematics and computer science at the University of Warsaw in Poland in 2009. Before starting his PH.D he spent few months as a programmer at IBM in New York (summer internship). Krzysztof obtained his Ph.D in 2013 (advised by Professor Maria Chudnovsky) and then joined Google Research Team in New York. A part of his research paper are mentioned in MathSciNet under ID:1000828.

Krzysztof works in different areas of machine learning and graph theory. Among his interests are: online non-parametric clustering, ranking algorithms, structural graph theory of networks defined by forbidden induced subgraphs, random graph theory, differential privacy. He was the first to prove the Erdos-Hajnal Conjecture for all tournaments on at most five vertices and to use the probabilistic method to give the best-known upper bounds on the so-called Erdos-Hajnal coefficients of prime graphs. He was a member of the team of researchers who gave a complete characteristic of the tournaments satisfying the Conjecture in the strongest linear sense. Together with Maria Chudnovsky and Paul Seymour, he gave a complete structural theorem of tournaments satisfying the Conjecture in the pseudolinear sense and, as a consequence, proved that tournaments with this weird property exist. Krzysztof is an author of patents regarding fast online clustering algorithms. He designed several machine learning algorithms that give strong privacy guarantees (such that differential privacy) as well as efficient algorithms breaking database privacy (also in the setting where a database has a graph structure), generalizing several state-of-the-art results in the field. He is an author of the ranking algorithms adjusted to the heterogeneous setting, where statistics are only locally of good quality and global good quality ranking does not exist. In neural networks he is interested in: models where the network of neural connections is not necessarily a union of bipartite graphs but has a more complex structure, (restricted) Boltzmann machines, new techniques used to speed up the training phase of deep neural networks as well as convolutional neural networks models using hierarchical clustering algorithms and processing input represented by weighted graphs.

Krzysztof plays piano and is an avid salsa dancer.

Jacek Miękisz was hiking Tatra Mountains at the age of minus 6 months and then he was born in Wrocław in 1956. He investigated interacting suspensions in ferrofluids and received Master degree in physics from Wrocław Technical University in 1979. Then he moved to Blacksburg, Virginia Tech, studied interacting spins in Ising models of ferromagnetism and got PhD in mathematics in 1984. He spent postdoctoral years at the University of Texas at Austin, in Louvain-la-Neuve and Leuven, studying interacting particles in lattice-gas models of quasicrystals. A part of his research paper are mentioned in MathSciNet under ID:238656. Now he works at the Institute of Applied Mathematics and Mechanics of the University of Warsaw and deals with interacting agents in evolutionary games and interacting proteins in genetic regulatory networks, sometimes with time delays.

Krzysztof Choromański
Columbia University
School of Engineering and Applied Science
Department of Industrial Engineering and Operations Research
500 W 120th Street, New York, 10027 NY, USA
E-mail: choromanski1@gmail.com

Jacek Miękisz
University of Warsaw
Institute of Applied Mathematics and Mechanics
Banacha 2, 02-097 Warsaw, Poland
E-mail: miekisz@mimuw.edu.pl
URL: https://www.mimuw.edu.pl/miekisz

Communicated by: Łukasz Stettner

(Received: 6th of April 2013; revised: 29th of June 2015)