Miskolc Mathematical Notes
Vol. 24 (2023), No. 2, pp. 625–634

HU e-ISSN 1787-2413
DOI: 10.18514/MMN.2023.4133

AN ENUMERATION APPROACH TO NETWORK EVOLUTION

GÁBOR BACSÓ AND JÓZSEF TÚRI

Received 21 February, 2022

Abstract. A simple theoretical model of network evolution is discussed here. In each step, we add a new vertex to the graph and it is allowed to connect it to maximum degree vertices (hubs) only. Given a constant p, the probability of such a connection is p for any hub. The initial (non-random) graph G_1 is arbitrary but here we investigate mostly the case when G_1 has one vertex.

We solve here some particular cases of the problem, using enumeration methods. We obtain not limit theorems but exact results for the parameters discussed.

2010 *Mathematics Subject Classification:* 05C80

Keywords: enumeration, random graphs, preferential attachment rule

1. Introduction

Network evolution is the subject of very intensive research since the starting of World Wide Web [4]. The literature of random graphs is extremely diverse [5]. In [3] the authors examine the size of large subgraphs of the binomial random graphs. From [15] we get a general insight how random graphs model large networks. In [16] a sequential metric dimension was examined in random graphs. Móri and Backhausz [1] study the degree distribution in the lower levels of the uniform recursive tree. Among the numerous further models and applications we mention here the solution of problems in physics [7].

In [8] the random graph dynamics appears, and [12] examines the adversarial deletion in a scale-free random graph process. In [17], the random graphs and the complex networks are examined. [13] actually reports the same results as Erdős and Rényi [9]. In Grenander's book [14], the descriptions of [10] is generalized. [11] details the probability theory information used in the article.

A simple theoretical model of such processes is investigated here. As we know, for numerous processes in large networks, Web, pandemic, and so on, the new vertex will be adjacent mostly to vertices of large degrees. According to that, we define a sequence of random graphs such that the rules of developing yields a model for

Supported in part by the National Research, Development and Innovation Office – NKFIH under the grant SNN 129364.

the properties above. A very frequent problem is to determine the distribution of degrees. We have partial results, concerning that. The model is far from the Erdős-Rényi one ([9], [10]), even the simplest questions are difficult here, thus often it is worth assuming that we start with the one-vertex graph and in every step, a new vertex appears.

Paper of Bollobás ([6]) is also significant on the subject.

Note that even now, for many networks, different from Web, the 'old' model can be applied successfully. In [2], for example, we solve domination problems, using Erdős-Rényi model. Here we would have tried it in vain.

2. Definitions and notation

Here random graphs will be denoted by boldface characters. Let the random graph $\mathbf{G}_N$ be the output of the following process, called Majority Process. Given a constant non-random graph G_1, the vertex set of the t'th graph $\mathbf{G}_t$ is $V(G_1) \cup \{x_2, \ldots, x_t\}$. The maximum possible value of t is denoted by N. $\mathbf{G}_N$, the final graph, represents the entire network. We call $V(\mathbf{G_N})$ by V for brevity. A vertex of a (random or deterministic) graph is called *hub* if it has maximum degree in the underlying graph. In general, G_1 is an arbitrary fixed graph, but we begin the work by the special case $G_1 \cong K_1$, where K_n is the clique on n vertices.

Supposed that the t'th graph is $\mathbf{G}_t$, we define $\mathbf{G}_{t+1}$ in the following way:

For any hub x, v_{t+1} is adjacent to x with probability p. These decisions are completely independent. For any other vertex x, xv_{t+1} is defined to be a non-edge.

Remark 1. At this step, we can see already that the model differs from the Erdős-Rényi one.

Remark 2. N and p are the only non-random values here.

The maximum vertex degree of a graph G is denoted by $\Delta(G)$. We denote the set of hubs in G by $B(G)$, $|B(G)|$ by $b(G)$. A vertex of a graph is *universal*, if it is adjacent to all the other vertices in the graph. For a non-empty graph G, omitting the isolated points, we obtain graph G'. A vertex is *quasi-universal* in G if it is universal in G'. A graph is *primitive* if it consists of a clique and isolated points. A graph is (γ, ε)-*primitive* if the clique has γ vertices and the number of isolated points is ε. Empty graphs and cliques are primitive, by definition. For a graph G_t in the process, the probability of being (γ, ε)-primitive, will be denoted $\pi(t, \gamma, \varepsilon)$.

A cycle (path) on n vertices is denoted by C_n (P_n). We need a basic definition concerning the general model.

Definition 1. A graph is G_1-*relevant* if it appears as $\mathbf{G}_t$ somewhere during the Majority Process, beginning with G_1. G_1 is G_1-relevant, by definition. If $G_1 \cong K_1$, we often write *relevant* for short.

Example 1. C_4 and P_4 are not [K_1-]relevant.

3. RESULTS AND PROOFS FOR THE GENERAL CASE

As we shall see, in the special case the number of hubs and the maximum degree are in thorough connection. (See Corollary 1, for example.) The following statement shows that for the general case, on the contrary, the pair (b,Δ) can be almost arbitrary.

Proposition 1. *Let $b_0 \geq 1$ and $\Delta_0 \geq 0$ be arbitrary integers, except the case "b_0 odd, $\Delta_0 = 1$". Then there exists some graph G_1 with $\Delta(G_1) = \Delta_0$ and $b(G_1) = b_0$.*

Proof. We begin by the case $\Delta_0 \geq 2$. Let H be any graph on $b_0 \geq 3$ vertices with all degrees at most Δ_0. (Such a graph exists, say, the cycle on b_0 vertices.) For a vertex x in H, we join $\Delta_0 - deg_H(x)$ pendant edges to x. For $b_0 = 2$, a P_4 fits.

For $\Delta_0 = 1$, the construction for an even b_0 is a set of independent edges and isolates, which always yields an even number of hubs. The case $\Delta_0 = 0$ is trivial. □

The following statement is valid for every graph G_1.

Proposition 2. *Suppose for a G_1-relevant graph G_j that $|B(G_j)| = \Delta(G_j) + 1$, in words, the number of hubs is exactly one greater than the maximum degree. Then, starting from this graph, there exists even an infinite sequence of increasing b and Δ, keeping their distance 1.*

Proof. We may connect the new vertex with every hub, in every step. □

Remark 3. Clearly, the probability of this phenomenon tends to zero.

We will use the statements below also in the next section.

Proposition 3. *Suppose v_j is universal in $\mathbf{G}_j$. Then $\mathbf{G}_{j-1}$ is a regular graph.*

Proof. In this case, by definition, each vertex in $\mathbf{G}_{j-1}$ is a hub. Thus, it is regular. □

We present here a surprising fact.

Proposition 4. *If a G_1-relevant graph is regular, then it is a clique or an empty graph.*

Proof. Let G be a graph in the statement of the Lemma and let G be r-regular. Let us omit the last vertex y of G. Suppose, by way of contradiction, that y does have both neighbors and non-neighbors. (This is equivalent for the whole graph to be neither a clique nor an empty graph.) Take a neighbor z of y. Clearly, $r \geq 1$ and the degree of z is $r-1$. The non-neighbors of y have degree r in $G-y$, thus z is not a hub in $G-y$. We have got a contradiction since a non-hub must be nonadjacent to y, by definition. □

4. Results and proofs for the special case

Let us give now some further definitions and notation. Let x be an arbitrary element of V and let $\Pi(k)$ be the probability that x has degree k in $\mathbf{G}_N$.

Remark 4. This probability depends on N, p and G_1 but they may be considered as constants.

Theorem 1. *For the probability of being universal in* $\mathbf{G}_N$*, the following result is valid.*

$$\Pi(N-1) = (p^{N-1}/N)\sum_{j=1}^{N}(p^{\lambda_j} + (1-p)^{\lambda_j})$$

where

$$\lambda_j = (j-1)(j-2)/2.$$

First we prepare the proof of Theorem 1 by six statements.

For a given element u of V, for any j between 1 and N, $\mathbb{P}(u = v_j) = 1/N$. Let us introduce now the following auxiliary variables:

$$a_j := \mathbb{P}(\mathbf{G}_{j-1} \text{ is a clique})$$
$$b_j := \mathbb{P}(v_j \text{ is universal in } \mathbf{G_j})$$
$$c_j := \mathbb{P}(\text{For every } l > j,\ v_l \text{ and } v_j \text{ are adjacent})$$
$$d_j := \mathbb{P}(\mathbf{G}_{j-1} \text{ is an empty graph})$$

The two variables below are of more importance:

$$q_j := \mathbb{P}(\mathbf{G}_{j-1} \text{ is a clique and } v_j \text{ is universal in } \mathbf{G_N})$$
$$r_j := \mathbb{P}(\mathbf{G}_{j-1} \text{ is an empty graph and } v_j \text{ is universal in } \mathbf{G_N})$$

Proposition 5.

$$a_j = p^{(j-1)(j-2)/2}$$

Proof.

$$\mathbb{P}(\mathbf{G}_{t+1} \text{ is a clique} \mid \mathbf{G}_t \text{ is a clique}) = p^t$$

for every $t < j$ since each vertex of $\mathbf{G}_t$ is a hub. Moreover, $\sum_{i=1}^{j-2} i = (j-1)(j-2)/2$.
□

□

Proposition 6.

$$d_j = (1-p)^{(j-1)(j-2)/2}$$

Proof.

$$\mathbb{P}(\mathbf{G}_{t+1} \text{ is an empty graph} \mid \mathbf{G}_t \text{ is an empty graph}) = (1-p)^t$$

for every $t \leq j-2$ since each vertex of $\mathbf{G}_t$ is a hub. We continue similarly as in the proof of Proposition 5. □

Proposition 7.

$$b_j = p^{j-1}$$

Proof. Each vertex of $\mathbf{G}_{j-1}$ is a hub. □

□

Proposition 8.

$$c_j = p^{N-j}$$

Proof. For each $l > j$, v_j is a hub in $\mathbf{G}_{l-1}$, thus $\mathbb{P}(v_l$ and v_j are adjacent$)= p$. □

□

Proposition 9.

$$q_j = p^{(j-1)(j-2)/2+N-1}$$

Proof. Clearly, $q_j = a_j b_j c_j$. From Proposition 5 and Proposition 8, we obtain the result. □

□

Proposition 10.

$$r_j = (1-p)^{(j-1)(j-2)/2} p^{N-1}$$

Proof. Clearly, $r_j = d_j b_j c_j$. From *Proposition* 7 and *Proposition* 6, the statement follows. □

Proof of Theorem 1. Let us apply the statements above, to prove Theorem 1. We denote $\mathbb{P}(v_j$ is universal in $\mathbf{G}_N)$ by s_j. Clearly,

$$s_j = q_j + r_j$$

Finally, it can be easily seen that $\Pi(N-1) = 1/N \sum_{j=2}^{N} s_j$ The proof of Theorem 1 has been established. □

Theorem 2. *Suppose $G_1 \cong K_1$. Then in any non-empty relevant graph, the hubs are quasi-universal.*

Proof. We use induction on $|V(G)|$. The first step is obvious. By the induction hypothesis, the hubs are universal in G'_{j-1}. Let v_j be a hub in G_j. If v_j is isolated in G_j then we are done. Otherwise, taking a neighbor s of v_j, it is a hub of G_{j-1}, by definition. That is, s is universal in G'_{j-1}. The degree of v_j in G_j is at least the degree of s. Consequently, it is quasi-universal in G_j. □

Corollary 1. *For the special case, the number of hubs is at most one more than the maximum degree.*

(This is an example for the difference between special and general case.)
Before stating the theorem below, we give a definition.

Definition 2. If a graph consists of a clique and isolated points, we call it *primitive*.

Theorem 3. *Suppose $G_1 \cong K_1$. Then, in the Majority Process, two phases can be distinguished.*

Phase 1 Every graph in this first sequence is primitive. In some steps the clique is growing, in other steps new isolates appear.

Phase 2 In this phase, we find two types of steps.

Step type a) The set B of hubs (which is always a clique, by Theorem 2) is replaced by one of its proper subsets (but it remains non-empty of course).

Step type b) The set B does not change.

Proof. Suppose G_{j-1} is not primitive. We state that in this case, the step from $j-1$ to j is in *Phase 2* and from this subscript on, we stay in this phase. From the supposition, there exist vertices which are neither isolates nor quasi-universal. As we know from Theorem 2, even they are not hubs. Consequently, v_j is not adjacent to them and thus it is not universal.

If v_j is adjacent to each point of B_{j-1} then this step is of type b). Otherwise, it is of type a). We stay in this second phase since the new graph G_j is not primitive. □

Theorems 2 and 3 show that for the special case the structure of the graphs occurring in the process is too poor. However, the calculation of $\Pi(0)$ is even more difficult than that of $\Pi(N-1)$. Let

$$S_j := \mathbb{P}(v_j \text{ is isolated in } G_N)$$

Obviously

$$\Pi(0) = 1/N \sum_{j=1}^{N} S_j$$

Our aim is to determine S_j for all j's.

We begin by two cases, $j = 1$ and $j = N$. The general solution will be a mixture of them.

Proposition 11.

$$S_1 = (1-p)^{\binom{N}{2}} + \sum_{\iota=2}^{N-1} (1-p)^{\binom{\iota}{2}+1} - \sum_{\iota=2}^{N-1} (1-p)^{\binom{\iota+1}{2}}$$

Proof. Let A_ι be the event that G_ι is an empty graph and $G_{\iota+1}$ is not. (Clearly, in this case all the graphs G_ν with $\nu \leq \iota$ are empty graphs as well.)

Let F_ι be the product of the events A_ι and "v_1 is isolated in G_N". Clearly

Claim 1.

$$S_1 = (1-p)^{\binom{N}{2}} + \sum_{\iota=2}^{N-1} \mathbb{P}(F_\iota)$$

Proof. The events F_ι and the event "G_N is an empty graph" are pairwise exclusive. $\square$

Take G_ι and the neighborhood of $v_{\iota+1}$ as random variables. By the definition of the process, they are independent. The neighborhood does not contain v_1 but it contains at least one vertex v_μ with $2 \le \mu \le \iota$. This implies

$$\mathbb{P}(F_\iota) = (1-p)^{\binom{\iota}{2}}((1-p)-(1-p)^\iota)$$

and, using $\binom{\iota}{2}+\iota = \binom{\iota+1}{2}$,

$$S_1 = (1-p)^{\binom{N}{2}} + \sum_{\iota=2}^{N-1}(1-p)^{\binom{\iota}{2}+1} - \sum_{\iota=2}^{N-1}(1-p)^{\binom{\iota+1}{2}}$$

The calculation of S_1 has been achieved. $\square$

We need

$$B_i := \{x \in V(G_i) \mid x \text{ hub in } G_i\},$$

to define a quantity helping in the calculation of the S_j's.

$$\kappa(i,h) := \mathbb{P}(|B_i| = h)$$

Remark 5. By Theorem 2, instead of 'hub', we could write 'quasi-universal vertex'.

First, we can easily show the fact

Proposition 12.

$$S_N = \sum_{z=1}^{N-1} \kappa(N-1,z)(1-p)^z$$

Proof.

$$S_N = \sum_{z=1}^{N-1} \mathbb{P}(v_N \text{ is isolated in } G_N \text{ and } |B_{N-1}| = z)$$

$\square$

We need further notation.

$$\theta(\iota, j) := \mathbb{P}(v_j \text{ is isolated in } G_N \text{ and } A_\iota \text{ occurs})$$

Now we are in the position to state the theorem determining S_j for arbitrary j. (S_0 and S_N are determined already.)

Theorem 4. *For* $3 \le j \le N-1$, $S_j = T_j + U_j + (1-p)^{\binom{N}{2}}$ *where*

$$T_j = \sum_{z=1}^{j-1} \kappa(j-1,z)(1-p)^z$$

and

$$U_j = \sum_{\iota=j}^{N-1} \theta(\iota, j)$$

Moreover, for $j = 2$, $T_2 = 0$ but the equaity for U_2 remains valid.

Proof. For the subscript of the effectively occurring event A_ι, there are three cases. $\iota \le j-1$, $j \le \iota \le N-1$ or $\iota = N$. According to that, S_j is the sum of three probabilities. □

Claim 2. For $\iota \le j-2$, if v_j is isolated in G_j then it is isolated in G_N as well.

Proof. G_j is not an empty graph, thus, having degree zero, v_j is not a hub in any further graph and consequently it will not get any new neighbor. □

Proposition 13. *Let $j \le \iota \le N-1$. Then*

$$\theta(\iota, j) = (1-p)^{\binom{\iota}{2}}((1-p)-(1-p)^{\iota}))$$

Proof. The new vertex is not adjacent to v_j but it has to be adjacent to some other vertex. □

We will present a recursion formula for κ.

Let us add one vertex to a graph in the process. The number of hubs may increase, namely, when the original graph is a specific primitive one. This is the reason that the recursion formula for κ is in connection also with primitive graphs.

Proposition 14.

$$\kappa(i,h) = \mathbb{P}(G_{i-1} \textit{ is } (h-1,t-h)\textit{-primitive})p^{h-1} + \sum_{z=h}^{N} \kappa(i-1,z)\binom{z}{h} p^h (1-p)^{z-h}$$

To obtain the probability of being primitive, a recursion formula is needed. This can be obtained by finding a recursion for the probability of being (γ, ε)- primitive. Let the underlying graph have t vertices. Thus $\gamma + \varepsilon = t$.

For $\gamma = 0$ we have an empty graph and for $\varepsilon = 0$ we have a clique. The calculation of the probabilities in these cases is straightforward. We consider them as starting values. In most of the cases, the following recursion can be used.

Proposition 15. *For $\gamma \ge 3$ and $\varepsilon \ge 1$*

$$\begin{aligned}\mathbb{P}(G_t \textit{ is } (\gamma,\varepsilon)\textit{-primitive}) &= \mathbb{P}(G_{t-1} \textit{ is } (\gamma-1,\varepsilon)\textit{-primitive})p^{\gamma-1} \\ &\quad + \mathbb{P}(G_{t-1} \textit{ is } (\gamma,\varepsilon-1)\textit{-primitive})(1-p)^{\gamma}\end{aligned}$$

Proof. We can obtain the large graph in two ways: by increasing the clique or the empty part. The two addends here represent exclusive events. □

Instead of $(1,\varepsilon)$-primitiveness, $(0,\varepsilon+1)$-primitiveness will be used. But this means that the recursion formula for $\gamma = 2$ must have an exceptional form:

Proposition 16.

$$\mathbb{P}(G_t \textit{ is } (2,\varepsilon)\textit{-primitive}) = (t-1)\,\mathbb{P}(G_{t-1} \textit{ is } (0,t-1)\textit{-primitive})p(1-p)^{t-2} + \mathbb{P}(G_{t-1} \textit{ is } (2,t-3)\textit{-primitive})(1-p)^2$$

Proof. In the first graph at the right hand side, we have to choose exactly one neighbor for v_t. □

Summarising, we expressed the S_j's by two quantities, θ and κ. An explicit formula for θ has been found, while for κ a recursion only, moreover, we needed another auxiliary parameter, not explicitly known but using a recursion again.

5. Concluding Remarks

In the subject, the most frequent question is the degree distribution of the network. For general degrees, to obtain exact formulae like here for $\Pi(0)$ and $\Pi(N-1)$ is not hopeful. Inspite of that, we ask:

Open problem How does the degree distribution depend on the initial graph G_1?

References

[1] Á. Backhausz and T. F. Móri, "Degree distribution in the lower levels of the uniform recursive tree," *Annales Univ. Sci. Budapest, Sect. Comput.*, vol. 36, pp. 53–62, 2012.

[2] G. Bacsó, J. Túri, and Z. Tuza, "Connected domination in random graphs," *Indian Journal Pure and Applied Mathematics*, Accepted for publication, doi: 10.1007/s13226-022-00265-2.

[3] J. Balogh and M. Zhukovskii, "On the sizes of large subgraphs of the binomial random graph," *Discrete Mathematics*, vol. 345, no. 2, pp. 128–139, 2022, doi: 10.1016/j.disc.2021.112675.

[4] A. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, pp. 509–512, 1999, doi: 10.1126/science.286.5439.509.

[5] B. Bollobás, *Random graphs*, ser. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2011. doi: 10.1017/CBO9780511814068.

[6] B. Bollobás, "The evolution of random graphs," *Transactions of the American Mathematical Society*, vol. 286, no. 1, pp. 257–274, 1984, doi: 10.2307/1999405.

[7] M. Deijfen and M. Lindholm, "Growing networks with preferential deletion and addition of edges," *Physica A: Statistical Mechanics and its Applications*, vol. 388, no. 19, pp. 4297–4303, 2009, doi: 10.1016/j.physa.2009.06.032.

[8] R. Durrett, *Random graph dynamics*. Citeseer, 2007. doi: 10.1.1.720.9457.

[9] P. Erdős and A. Rényi, "On random graphs I." *Publ. Math. Debrecen*, vol. 6, pp. 290–297, 1959.

[10] P. Erdős and A. Rényi, "On the evolution of random graphs," *Publ. Math. Inst. Hung. Acad. Sci*, vol. 5, no. 1, pp. 17–61, 1960, doi: 10.1.1.348.530.

[11] W. Feller, *An introduction to probability theory and its applications*. Wiley, 1971.

[12] A. D. Flaxman, A. M. Frieze, and J. Vera, "Adversarial deletion in a scale-free random graph process," *Combinatorics, Probability and Computing*, vol. 16, no. 2, pp. 261–270, 2007, doi: 10.1017/S0963548306007681.

[13] E. N. Gilbert, "Random graphs," *The Annals of Mathematical Statistics*, vol. 30, no. 4, pp. 1141–1144, 1959.

[14] U. Grenander, *Probabilities on algebraic structures*. Courier Corporation, 2008.

[15] L. Lovász, "Large networks and graph limits," *American Mathematical Soc., Colloquium Publications*, vol. 60, 2012.

[16] G. Odor and P. Thiran, "Sequential metric dimension for random graphs," *Journal of Applied Probability*, vol. 58, no. 4, pp. 909–951, 2021, doi: 10.1017/jpr.2021.16.

[17] R. Van Der Hofstad, "Random graphs and complex networks," *Available on http://www. win. tue. nl/rhofstad/NotesRGCN. pdf*, vol. 11, p. 60, 2009.

Authors' addresses

Gábor Bacsó
Institute for Computer Science and Control, 1111 Budapest, Hungary
E-mail address: tud23sci@gmail.com

József Túri
(**Corresponding author**) University of Miskolc, 3515 Miskolc, Hungary
E-mail address: matturij@uni-miskolc.hu