Evsey V. Morozov
Alexander S. Rumyantsev
Oleg V. Lukashenko (Eds.)

Proceedings
of the Third International
Workshop **SMARTY'22**

SMARTY 22

Stochastic Modeling & Applied Research of TechnologY

Third International Workshop
Petrozavodsk, Karelia
August 21-25, 2022

Pavlov, Yu.

Asymptotics of the structure of conditional configuration graphs with bounded number of links

Stochastic Modeling and Applied Research of Technology, Vol. 3, Pp. 22-29.
DOI: 10.57753/SMARTY.2023.49.82.003

Asymptotics of the structure of conditional configuration graphs with bounded number of links

Yury Pavlov

IAMR KRC RAS,
Petrozavodsk, 185910, Russia,
pavlov@krc.karelia.ru

Abstract. A model of a configuration graph on N vertices is considered where the number of edges is at most n. The degrees of the vertices are independent random variables identically distributed to the power law that depends on a slowly varying function with remainder term. We obtained the limit distributions of the maximum vertex degree and the number of vertices with a given degree as $N, n \to \infty$.

Keywords: Configuration graph, Vertex degree, Limit theorems

1 Introduction

The study of random graphs has been gaining attention with the wide use of these models for the description of different complex networks (see e. g. [1]). Observations on real networks showed that their topology can be described by random graphs with vertex degrees being independent identically distributed random variables. Furthermore, it turned out that the number of vertices with degree k is proportional to $k^{-\tau}$, where $\tau > 1$. This means that the distribution of random variable ξ, being equal to an arbitrary vertex degree, can be defined as follows:

$$p_k = \mathbf{P}\{\xi = k\} = \frac{h(k)}{k^\tau}, \qquad k = 1, 2, \dots, \tag{1}$$

where $h(k)$ is a slowly varying function [2]. One of the most appropriate graphs for modeling the networks is so called configuration graph [3]. The random variable ξ in such graph takes natural values equal to the number of vertex semiedges, i. e. edges for which the adjacent vertices are not yet specified. All semiedges are numbered in an arbitrary order. The sum of vertex degrees in any graph has to be even, so if the sum is odd we add one extra vertex with degree one. The graph is constructed by joining all semiedges pairwise equiprobably to form edges. Noted [4] that an additional vertex together with its semiedge does not influence the graph behaviour as the number of graph vertices tends to infinity. So in what follows we shall consider the degrees only for the initial vertices even though an extra vertex is introduced. It is easy to see that such graphs may have loops and multiple edges.

The papers [5, 6] were concerned with conditional configuration graphs with a known number of edges. The vertex degree distributions in these studies are fully known, since the slowly varying function $h(x)$ in (1) is a constant. In [7] it was assumed that

$$p_k \sim \frac{d}{k^g (\ln k)^h}$$

as $k \to \infty$, where $d > 0, g \geqslant 1, h \geqslant 0, g + h > 1$.

Networks are more naturally described by models in which the number of edges in the graph is bounded from above. Configuration graphs with this condition were studied for the first time in [8] where it was assumed that

$$p_k = \frac{1}{k^\tau} - \frac{1}{(k+1)^\tau}, \quad k = 1, 2, \ldots, \quad \tau > 0.$$

This paper considers configuration graphs consisting of N vertices in which the random variable ξ has the distribution (1), where $\tau > 3$. We denote by $\xi_1, \ldots, \xi_N$ the degrees of vertices $1, \ldots, N$ respectively. These random variables are independent and equidistributed with ξ. Such random graph naturally induces a probability measure on the subset of realisations in which $\xi_1 + \ldots + \xi_N \leqslant n$, that is, the number of the semiedges of the graph is not greater than n.. We consider the conditional random graph arising in this way with a bounded total sum of vertex degrees. Denote by $\eta_1, \ldots, \eta_N$ the random variables defined as the degrees of vertices $1, \ldots, N$ in this conditional configuration graph. It is evident that these random variables are dependent.

A well-known elementary property (see, f. e., [2]) of a slowly varying function is that as $x \to \infty$ and any $\delta > 0$

$$x^{-\delta} < h(x) < x^\delta. \tag{2}$$

Let $\tau > 3$. From (1), (2) we get that the random variable ξ has a finite expectation $m = \mathbf{E}\xi$ and variance $\sigma^2 = \mathbf{D}\xi$:

$$m = \sum_{k=1}^{\infty} \frac{h(k)}{k^{\tau-1}}, \qquad \sigma^2 = \sum_{k=1}^{\infty} \frac{h(k)}{k^{\tau-2}} - m^2. \tag{3}$$

Assume that the function $h(x)$ is a slowly varying function with a remainder term. The definition of such a function is given in [9] and consists in the following. Let $\varphi(x)$ be a positive increasing function such that $\varphi(x) \to \infty$ as $x \to \infty$ and for some positive numbers θ, X function $\varphi(x)/x^\theta$ does not increase on the interval (X, ∞). A positive measurable function $h(x)$ is called a slowly varying function with remainder term $\varphi(x)$ if for all $\lambda > 0$

$$\frac{h(\lambda x)}{h(x)} = 1 + O\left(\frac{1}{\varphi(x)}\right) \tag{4}$$

as $x \to \infty$. A well-known example of a slowly varying function with a remainder term is the logarithm.

We denote by $\eta_{(N)}$ and μ_r the maximum vertex degree and the number of vertices with degree r respectively. Below we prove limit theorems for $\eta_{(N)}$ and μ_r for different behaviours of the parameters N, n, r.

The paper is organised as follows. The main results (Theorems 1 - 3) are formulated in Section 2. Section 3 discusses the connection of the problem under consideration with the generalized scheme of allocating particles to cells. Section 4 deals with some auxiliary results, which are used later in Section 5 to prove Theorems 1 - 3.

2 Statement of the main results

The next theorems are proved.

Theorem 1. *Let* $N, n \to \infty$, *in such a way that*

$$\frac{n - Nm}{\sqrt{N}} \geqslant C > -\infty \tag{5}$$

and $r = r(N, n)$ *is chosen such that*

$$\frac{Nh(r)}{(\tau - 1)r^{\tau-1}} \to \gamma, \tag{6}$$

where γ *is a positive constant. Then*

$$\mathbf{P}\{\eta_{(N)} \leqslant r\} \to e^{-\gamma}.$$

Theorem 2. *Let* $N, n \to \infty, Np_r(1 - p_r) \to \infty$ *and*

$$\frac{n - Nm}{\sqrt{N}} \to \infty. \tag{7}$$

Then

$$\mathbf{P}\{\mu_r = k\} = \frac{1 + o(1)}{\sqrt{2\pi Np_r(1 - p_r)}} e^{-u_r^2/2}$$

uniformly in the integer k *such that* $u_r = (k - Np_r)/\sqrt{Np_r(1 - p_r)}$ *lies in any fixed finite interval.*

Theorem 3. *Let* $N, n, r \to \infty$ *and suppose that the condition* (5) *is satisfied. Then, for a nonnegative integer* k,

$$\mathbf{P}\{\mu_r = k\} = \frac{(Np_r)^k}{k!} e^{-Np_r}(1 + o(1))$$

uniformly with respect to $(k - Np_r)/\sqrt{Np_r}$ *in any fixed finite interval.*

3 Connection to the generalized scheme of allocation

The technique of obtaining Theorems 1 - 3 is based on the generalized scheme of allocations suggested by V. F. Kolchin [10]. It is easy to see that if $\eta_1+\ldots+\eta_N = n$, then

$$\mathbf{P}\{\eta_1 = k_1, \ldots, \eta_N = k_N\} = \mathbf{P}\{\xi_1 = k_1, \ldots, \xi_N = k_N | \xi_1 + \ldots + \xi_N = n\}.$$

This equality means that the conditions of generalized scheme are valid. It is also clear that in our conditional configuration graph

$$\mathbf{P}\{\eta_1 = k_1, \ldots, \eta_N = k_N\} = \mathbf{P}\{\xi_1 = k_1, \ldots, \xi_N = k_N | \xi_1 + \ldots + \xi_N \leqslant n\}. \quad (8)$$

In papers [11, 12] a pair of N-tuples of random variables $(\eta_1, \ldots, \eta_N), (\xi_1, \ldots, \xi_N)$ satisfying the relation (8) was called an analogue of the generalized scheme of allocation.

We introduce auxiliary random variables $\xi_1^{(r)}, \ldots, \xi_N^{(r)}$ such that

$$\mathbf{P}\{\xi_i^{(r)} = k\} = \mathbf{P}\{\xi_i = k | \xi_i \leqslant r\}, \quad (9)$$

where $i = 1, \ldots, N, k = 1, \ldots, r$. Let $\tilde{\xi}_1^{(r)}, \ldots, \tilde{\xi}_N^{(r)}$ be random variables with the distribution

$$\mathbf{P}\{\tilde{\xi}_i^{(r)} = k\} = \mathbf{P}\{\xi_i = k | \xi_i \neq r\}, \quad (10)$$

$i = 1, \ldots, N, k = 1, 2, \ldots$ We also set

$$\zeta_N = \xi_1 + \ldots + \xi_N, \quad \zeta_N^{(r)} = \xi_1^{(r)} + \ldots + \xi_N^{(r)}, \quad \tilde{\zeta}_N^{(r)} = \tilde{\xi}_1^{(r)} + \ldots + \tilde{\xi}_N^{(r)} \quad (11)$$

and

$$P_r = \mathbf{P}\{\xi > r\}. \quad (12)$$

It was proved in [11, 12] that (8) implies the following assertions.

Lemma 1. *The following equality holds*

$$\mathbf{P}\{\eta_{(N)} \leqslant r\} = (1 - P_r)^N \frac{\mathbf{P}\{\zeta_N^{(r)} \leqslant n\}}{\mathbf{P}\{\zeta_N \leqslant n\}}.$$

Lemma 2. *The following equality holds*

$$\mathbf{P}\{\mu_r = k\} = \binom{N}{k} p_r^k (1 - p_r)^{N-k} \frac{\mathbf{P}\{\tilde{\zeta}_{N-k}^{(r)} \leqslant n - kr\}}{\mathbf{P}\{\zeta_N \leqslant n\}}.$$

4 Auxiliary results

To estimate the behaviour of $(1 - P_r)^N$ in Lemma 1 we need to consider the asymptotics of NP_r.

Lemma 3. *Under the hypothesis of Theorem 1* $NP_r \to \gamma$.

Proof. It follows from (1) that

$$P_r = \sum_{i=1}^{\infty} \frac{h(r+i)}{(r+i)^\tau} = h(r) \sum_{i=1}^{\infty} \frac{h(r+i)}{h(r)(r+i)^\tau}. \tag{13}$$

It is proved in [9] (see also Lemma A.1.1 in [13]) that (4) implies

$$\frac{h(r+i)}{h(r)} = 1 + O\left(\left(1+\frac{i}{r}\right)\frac{1}{\varphi(r)}\right). \tag{14}$$

Obviously, $r \to \infty$ (see (6)). Then

$$\sum_{i=1}^{\infty} \frac{1}{(r+i)^\tau} \sim \int_1^\infty \frac{dx}{(r+x)^\tau} = \frac{1}{(\tau-1)r^{\tau-1}}. \tag{15}$$

Similarly,

$$\frac{1}{r\varphi(r)} \sum_{i=1}^{\infty} \frac{i}{(r+i)^\tau} < \frac{1}{\varphi(r)} \int_r^\infty \frac{dy}{y^{\tau-1}} \sim \frac{1}{\varphi(r)(\tau-2)r^{\tau-1}} = o\left(\sum_{i=1}^{\infty} \frac{1}{(r+i)^\tau}\right).$$

From this and (13) - (15) we get

$$P_r \sim \frac{h(r)}{(\tau-1)r^{\tau-1}}. \tag{16}$$

Hence, the assertion of Lemma 3 follows from (6).

Let

$$m_r = \mathbf{E}\xi_1^{(r)}, \quad \tilde{m}_r = \mathbf{E}\tilde{\xi}_1^{(r)}, \quad \sigma_r^2 = \mathbf{D}\xi_1^{(r)}, \quad \tilde{\sigma}_r^2 = \mathbf{D}\tilde{\xi}_1^{(r)}.$$

From (1), (8), (9) we have

$$m_r = \frac{m - \sum_{k>r} kp_k}{1-P_r}, \quad \sigma_r^2 = \frac{\sigma^2 + m^2 - \sum_{k>r} k^2 p_k}{1-P_r} - m_r^2,$$

$$\tilde{m}_r = \frac{m - rp_r}{1-p_r}, \quad \tilde{\sigma}_r^2 = \frac{\sigma^2}{(1-p_r)^2}\left(1 - p_r - \frac{(m-r)^2}{\sigma^2} p_r\right). \tag{17}$$

Using the condition $\tau > 3$ and (1) - (3), (17) we find that the first two moments of distributions (1), (9), (10) are finite. Then, the central limit theorem can be applied to the sums (11). The following assertions hold.

Lemma 4. *Let $N \to \infty$. Then, the distribution $(\zeta_N - Nm)/(\sigma\sqrt{N})$ converges weakly to the standard normal law.*

Lemma 5. *Let $N \to \infty$. Then, the distribution $(\zeta_N^{(r)} - Nm_r)/(\sigma_r\sqrt{N})$ converges weakly to the standard normal law.*

Lemma 6. *Let $N \to \infty$. Then, the distribution $(\tilde{\zeta}_N - N\tilde{m}_r)/(\tilde{\sigma}_r\sqrt{N})$ converges weakly to the standard normal law.*

5 Proofs of Theorems 1 - 3

It is clear under the conditions of Theorem 1 that $r \to \infty$. By analogy with (16), it is not difficult to see that

$$\sum_{k>r} kp_k = O\left(\frac{h(r)}{r^{\tau-2}}\right). \tag{18}$$

By (6),

$$\frac{h(r)}{r^{\tau-2}} \sim \frac{\gamma(\tau-1)r}{N}. \tag{19}$$

It follows from (4), (6) that for sufficiently small $\delta > 0$

$$r = O(N^{1/(\tau-1-\delta)}).$$

Since $\tau > 3$, we have $r = o(\sqrt{N})$. Thus, by virtue of (18), (19)

$$\sum_{k>r} kp_k = o\left(\frac{1}{\sqrt{N}}\right). \tag{20}$$

It therefore follows from (17) and Lemma 3 that

$$m_r = m(1 + o(N^{-1/2})). \tag{21}$$

Employing (1) and (2), we find that

$$r^2 p_r \to 0. \tag{22}$$

We put

$$z_N(n) = \frac{n - Nm}{\sigma\sqrt{N}}, \quad z_N^{(r)}(n) = \frac{n - Nm_r}{\sigma_r\sqrt{N}}, \quad \tilde{z}_N^{(r)}(n) = \frac{n - N\tilde{m}_r}{\tilde{\sigma}_r\sqrt{N}}. \tag{23}$$

The relation (22) combined with (5), (17), (20), (21), (23) implies that under the conditions of Theorem 1 the values of $z_N(n)$ and $z_N^{(r)}(n)$ behave asymptotically the same. Hence, from Lemmas 4, 5 we find that

$$\frac{\mathbf{P}\{\zeta_N^{(r)} \leqslant n\}}{\mathbf{P}\{\zeta_N \leqslant n\}} = 1 + o(1).$$

Now the conclusion of Theorem 1 easily follows from these relations and Lemmas 1, 3.

Under the hypothesis of Theorem 2, $k = Np_r + u_r\sqrt{Np_r(1-p_r)}$, and from (23) we have

$$\tilde{z}_{N-k}^{(r)}(n - kr) = \frac{n - kr - (N-k)\tilde{m}_r}{\tilde{\sigma}_r\sqrt{N}} = \frac{n - Nm}{\tilde{\sigma}_r\sqrt{N}} - \frac{u_r(r - \tilde{m}_r)\sqrt{p_r(1-p_r)}}{\tilde{\sigma}_r}. \tag{24}$$

From (1), (2), (23) and (24) we deduce the estimate

$$z^{(r)}_{N-k}(n-kr) = z_N(n)\frac{\sigma}{\tilde{\sigma}_r} + O(1).$$

Hence, using (7) we see that $z_N(n)$ and $z^{(r)}_{N-k}(n-kr)$ simultaneously tend to infinity, and from Lemmas 4, 6 we get

$$\frac{\mathbf{P}\{\tilde{\zeta}^{(r)}_{N-k} \leqslant n-kr\}}{\mathbf{P}\{\zeta_N \leqslant n\}} \to 1. \tag{25}$$

Using the normal approximation for a binomial distribution as $Np_r(1-p_r) \to \infty$ we find that

$$\binom{N}{k} p_r^k (1-p_r)^{N-k} = \frac{1+o(1)}{\sqrt{2\pi N p_r(1-p_r)}} e^{-u_r^2/2}. \tag{26}$$

Hence, the assertion of Theorem 2 follows from (25), (26) and from Lemma 2.

Theorem 3 is established in the same way as Theorem 2. It is clear that $p_r \to \infty$. Note that the relation (24) remains valid for the values of k considered in Theorem 3 if we replace u_r with $(k-Np_r)/\sqrt{Np_r}$. Taking (5), (17), (22), (23) into account, we obtain

$$\tilde{z}^{(r)}_{n-k}(n-kr) \sim z_N(n).$$

From this relation and from Lemmas 4, 6 we again come to (25). Theorem 3 follows from (25) and Lemma 2 considering that as for $p_r \to \infty$ binomial probabilities admit the Poisson approximation:

$$\binom{N}{k} p_r^k (1-p_r)^{N-k} \sim \frac{(Np_r)^k}{k!} e^{-Np_r}.$$

Acknowledgements The study was carried out under state order to the Karelian Research Centre of the Russian Academy of Sciences (Institute of Applied Mathematical Research KarRC RAS).

References

1. R. Hofstad, *Random Graphs and Complex Networks. Vol. 1*, Cambridge University Press, Cambridge (2017).
2. E. Seneta, *Regularly varying functions*, Springer-Verlag, Berlin (1976).
3. B. Bollobas, "A probabilistic proof of an asymptitic formula of the number of labelled regular graphs", *European J. Combin.*, **1**, No. 4, 311–316 (1980).
4. H. Reittu and I. Norros, "On the power-law random graph model of massive data networks", *Perform. Evaluation*, **55**, No. 1-2, 3–23 (2004).
5. Yu. L. Pavlov, "On conditional Internet graphs whose vertex degrees have

no mathematical expectation", *Discrete Mathematics and Applications*, **20**, No. 5-6, 509–524 (2010).

6. Yu. L. Pavlov and I. A. Cheplyukova, "Random Internet type graphs and the generalized allocation scheme", *Discrete Mathematics and Applications*, **18**, No. 5, 447–464 (2008).
7. Yu. L. Pavlov, "Conditional configuration graphs with discrete power-law distribution of vertex degrees", *Sbornic: Mathematics*, **209**, No. 2, 258–275 (2018).
8. Yu. L. Pavlov and E. V. Khvorostyanskaya, "On the limit distributions of the degrees of vertices in configuration graphs with a bounded number of edges", *Sbornik: Mathematics*, **207**, No. 3, 400–417 (2016).
9. S. Aljanĉić, R. Bojanić and M. Tomić, *Slowly varying functions with remainder term and their applications in analysis*, Serb. Acad. Sci. and Arts Monographs, vol. 462, sect. nat. math., No. 41, Belgrad (1974).
10. V. F. Kolchin, *Random Mappings*, Optimisation Softwere Inc., New York (1986).
11. A. N. Chuprunov and I. Fasekas, "An analogue of the generalized allocation scheme: limit theorems for the number of cells containing a given number of particles", *Discrete Mathematics and Applications*, **22**, No. 1, 101–122 (2012).
12. A. N. Chuprunov and I. Fasekas, "An analogue of the generalized allocation scheme: limit theorems for the maximum cell load", *Discrete Mathematics and Applications*, **22**, No. 3, 307–314 (2012).
13. V. M. Zolotarev and I. S. Shiganov, "Additions", in book: E. Seneta, *Pravil'no menyayushiesya funcii*, Nauka, Moscow, 100–131 (1985). In Russian.