

УДК 519.179.2, 519.876.5

МИНИМУМ И МАКСИМУМ СРЕДНЕГО РАССТОЯНИЯ В КОНФИГУРАЦИОННЫХ ГРАФАХ СО СТЕПЕННЫМ РАСПРЕДЕЛЕНИЕМ

М. М. Лери

*Институт прикладных математических исследований КарНЦ РАН,
ФИЦ «Карельский научный центр РАН» (ул. Пушкинская, 11,
Петрозаводск, Республика Карелия, Россия, 185910)*

Рассматриваются конфигурационные графы, степени вершин которых имеют дискретное степенное распределение с фиксированным параметром. Посредством имитационного моделирования оценивается среднее расстояние (среднее арифметическое всех расстояний между всеми парами вершин графа) и рассматриваются две характеристики: минимальное среднее расстояние в графе и максимальное. В работе строятся зависимости этих характеристик от объема графа и параметра распределения степеней вершин и строится эмпирический доверительный интервал среднего расстояния в степенном конфигурационном графе.

Ключевые слова: конфигурационный граф; степенное распределение; среднее расстояние в графе; минимальное среднее расстояние в графе; максимальное среднее расстояние в графе; имитационное моделирование

Для цитирования: Лери М. М. Минимум и максимум среднего расстояния в конфигурационных графах со степенным распределением // Труды Карельского научного центра РАН. 2022. № 4. С. 45–50. doi: 10.17076/mat1573

Финансирование. Финансовое обеспечение исследований осуществлялось из средств федерального бюджета на выполнение государственного задания КарНЦ РАН (Институт прикладных математических исследований КарНЦ РАН).

M. M. Leri. MINIMUM AND MAXIMUM OF THE AVERAGE DISTANCE IN POWER-LAW CONFIGURATION GRAPHS

Institute of Applied Mathematical Research, Karelian Research Centre, Russian Academy of Sciences (11 Pushkinskaya St., 185910 Petrozavodsk, Karelia, Russia)

The article studies configuration graphs with a discrete fixed-parameter power-law distribution of node degrees. Simulation is applied to estimate the average distance (the arithmetic mean of all distances between all pairs of graph vertices), and two characteristics are considered: the minimum and the maximum average distance in the graph. The dependences of these characteristics on the graph size and the vertex degree distribution parameter are obtained and the empirical confidence interval of the average distance in the power-law configuration graph is constructed.

Key words: configuration graph; power-law distribution; average distance in a graph; minimum average distance in a graph; maximum average distance in a graph; simulations

For citation: Leri M. M. Minimum and maximum of the average distance in power-law configuration graphs. *Trudy Kareli'skogo nauchnogo tsentra RAN = Transactions of the Karelian Research Centre RAS*. 2022;4:45–50. doi: 10.17076/mat1573

Funding. The studies were funded from the federal budget through state assignment to the Karelian Research Centre RAS (Institute of Applied Mathematical Research, Karelian Research Centre RAS).

Сложные сети коммуникаций продолжают оставаться одним из объектов, вызывающих стойкий интерес исследователей (см., например, [4, 6, 8]). Очевидно, что топология таких сетей влияет на их функционирование, а также на сохранение работоспособности в условиях внешних воздействий. В качестве математических моделей для описания сложных сетей нашли широкое применение случайные графы [4–7], разнообразие которых состоит в выборе способа определения степеней вершин и задании связей между этими вершинами. Например, для моделирования таких объектов, как телекоммуникационные сети, одной из которых является глобальная сеть Интернет, применяют случайные графы с независимыми одинаково распределенными степенями вершин с общим дискретным законом распределения [4, 5, 10].

В работе рассматриваются конфигурационные графы [1], состоящие из N вершин, степени $\xi_1, \xi_2, \dots, \xi_N$ которых представляют собой независимые одинаково распределенные случайные величины с дискретным степенным распределением [10]:

$$\mathbf{P}\{\xi = k\} = k^{-\tau} - (k+1)^{-\tau}, \quad k = 1, 2, \dots, \quad (1)$$

где параметр $\tau > 1$ является фиксированным для всех вершин графа. При построении конфигурационной модели сначала определяют степени каждой из N вершин в соответствии с распределением (1) с заранее заданным параметром τ . Степень вершины задает число различных полуредер [10]. Все полуредера графа нумеруются в произвольном порядке. В настоящей работе сумма степеней вершин конфигурационного графа является случайной величиной. Если она нечетна, то одно полуредеро добавляется к равновероятно выбранной вершине. Далее редера графа строятся посредством попарного, равновероятного соединения всех полуредер между собой. Очевидно, что построенный таким образом граф может иметь петли, кратные редера и циклы.

Также о конфигурационных графах известно (см., например, [4, 6, 10]), что при $\tau > 1$ число компонент связности такого графа асимптотически почти наверное больше 1.

Изучение структурных характеристик сложных сетей представляет определенный интерес в силу их значимости как при моделировании сетевой топологии, так и при изучении процессов, происходящих в таких сетях [4, 6]. Одной из важных числовых характеристик сети является расстояние между ее узлами, чему в случайном графе, как сетевой модели, соответствует расстояние между вершинами графа (см., например, [2, 7]).

В настоящей работе рассматривались две структурные характеристики: минимальное среднее расстояние в графе и максимальное среднее расстояние в графе. Под расстоянием $d(v, u)$ между двумя произвольно выбранными вершинами v и u графа ($v \neq u$) будем понимать длину цепи (число редер, соединяющих вершины v и u) от вершины v до вершины u наименьшей длины. Если вершины v и u принадлежат разным компонентам связности, то расстояние между ними полагают равным ∞ . Среднее расстояние вычисляется как среднее арифметическое всех расстояний графа:

$$dist = \frac{\sum_{v \neq u} d(v, u)}{k}, \quad (2)$$

где k – число расстояний $d(v, u) \neq \infty$ между всеми парами вершин графа. Исследование проводилось с помощью методов имитационного моделирования и рассматривались конфигурационные графы, число вершин которых принимало следующие значения: $10 \leq N \leq 100$ с шагом 10, $100 \leq N \leq 1000$ с шагом 50, $1000 \leq N \leq 7000$ с шагом 500. Значения параметра τ изменялись от 1,1 до 3,0 с шагом 0,1. Для каждой пары значений (N, τ) генерировалось по 100 графов. Расстояния $d(v, u)$ между всеми парами вершин каждого графа находились посредством алгоритма Дейкстры [3]. Затем для каждого графа рассчитывалось

среднее расстояние $dist$ по формуле (2), после чего для каждой пары значений (N, τ) вычислялось минимальное среднее расстояние $dist_{\min}$ как минимум из 100 средних расстояний, полученных в 100 экспериментах, и максимальное среднее расстояние $dist_{\max}$ – максимум по 100 имитационным графам. Цель работы состояла в нахождении зависимостей этих характеристик от числа вершин графа N и параметра распределения степеней вершин τ и построении эмпирического доверительного интервала среднего расстояния в степенном конфигурационном графе.

Для наглядности на рис. 1 графически показана зависимость экспериментальных значений среднего расстояния $dist$ в графе от числа его вершин N и параметра распределения (1) степеней вершин τ .

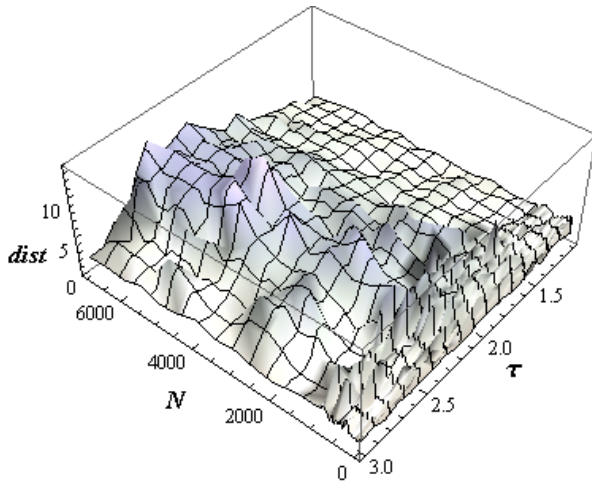


Рис. 1. Зависимость экспериментальных значений от N и τ
Fig. 1. Dependence of experimental values of $dist$ on N and τ

Были получены следующие регрессионные уравнения зависимостей среднего расстояния $dist$, а также минимума $dist_{\min}$ и максимума $dist_{\max}$ среднего расстояния в графе от N и τ :

$$dist = -0,361 - (0,075 - 0,658\tau^2 + 0,205\tau^3) \ln N, \quad (3)$$

$$dist_{\min} = -0,226 - (0,567 - 1,091\tau + 0,277\tau^2) \ln N, \quad (4)$$

$$dist_{\max} = 0,432 - (0,547 - 1,103\tau^2 + 0,323\tau^3) \ln N \quad (5)$$

с коэффициентами детерминации R^2 этих уравнений, равными 0,82, 0,71 и 0,74 соответственно. Здесь и далее все коэффициенты регрессионных уравнений значимы на уровне значимости 0,05, а гипотеза $H_0 : R^2 = 0$ отвергается на 5%-м уровне значимости.

В работе [7] показано, что «типичное расстояние» (математическое ожидание расстояния) в конфигурационном графе при $N \rightarrow \infty$ растет как $\ln \ln N$ при $1 < \tau < 2$ и как $\ln N$ при $\tau > 2$. В настоящей работе рассматривается среднее расстояние в графах в доасимптотической области ($N \leq 7000$) на интервале изменения параметра $\tau \in (1, 3]$, и заметим, что согласно (3)–(5) все три рассматриваемые характеристики растут логарифмически с ростом числа вершин графа N .

На рис. 2 графически представлены зависимости (3)–(5) при четырех объемах графа: 10, 100, 1000 и 7000. Сплошные линии на графиках соответствуют зависимости $dist$ от τ , а точечная и пунктирная – зависимостям $dist_{\min}$ и $dist_{\max}$ соответственно.

Минимальное и максимальное средние расстояния в графе можно рассматривать в качестве граничных значений эмпирического доверительного интервала для среднего расстояния $dist$. Отметим (см. рис. 2), что этот интервал расширяется примерно к середине отрезка $1, 1 \leq \tau \leq 3$, сужаясь к его границам.

Исследование показало, что все три характеристики: среднее расстояние, минимальное и максимальное среднее расстояние в конфигурационном графе имеют максимумы на рассмотренном интервале изменения параметра $\tau \in (1, 3]$.

Введем следующие обозначения:

$M_{\text{mean}} = \max dist_{\text{mean}}$ – максимум среднего расстояния в графе;

$M_{\text{min}} = \max dist_{\text{min}}$ – максимум минимального среднего расстояния;

$M_{\text{max}} = \max dist_{\text{max}}$ – максимум максимального среднего расстояния.

По экспериментальным данным для каждого значения N были вычислены M_{mean} , M_{min} , M_{max} и соответствующие им значения τ , при которых эти максимумы достигаются:

$$\begin{aligned} \hat{\tau}_{\text{mean}} &= \tau(\max dist_{\text{mean}}), \\ \hat{\tau}_{\text{min}} &= \tau(\max dist_{\text{min}}), \\ \hat{\tau}_{\text{max}} &= \tau(\max dist_{\text{max}}). \end{aligned}$$

Полученные результаты представлены в таблице.

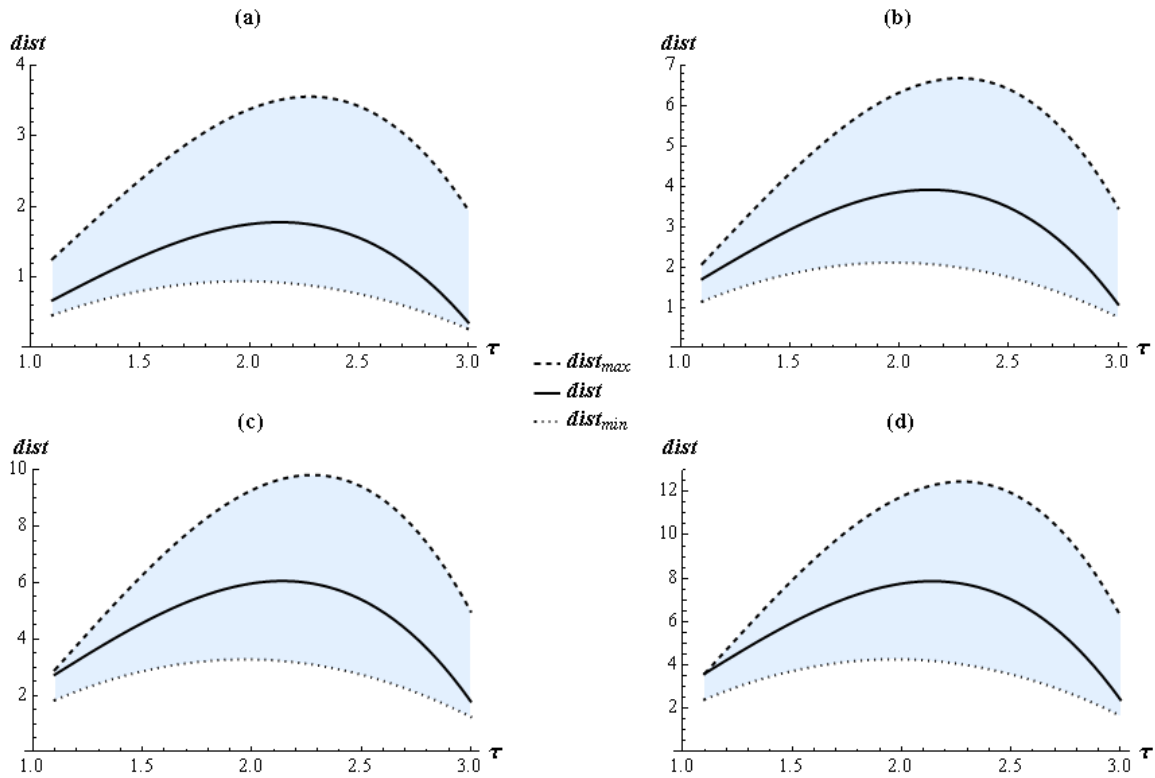


Рис. 2. Регрессионные зависимости $dist$, $dist_{\min}$ и $dist_{\max}$ от τ при: (a) $N = 10$, (b) $N = 100$, (c) $N = 1000$ и (d) $N = 7000$

Fig. 2. Regression dependencies of $dist$, $dist_{\min}$ and $dist_{\max}$ on τ for (a) $N = 10$, (b) $N = 100$, (c) $N = 1000$ and (d) $N = 7000$

Максимум среднего расстояния M_{mean} и соответствующие значения $\hat{\tau}_{\text{mean}}$, максимум минимального среднего расстояния M_{min} с соответствующими значениями $\hat{\tau}_{\text{min}}$ и максимум максимального среднего расстояния M_{max} с соответствующими значениями $\hat{\tau}_{\text{max}}$ для некоторых объемов графа N

Maximum of the average distance M_{mean} and corresponding values of $\hat{\tau}_{\text{mean}}$, maximum of minimal average distance M_{min} with corresponding values of $\hat{\tau}_{\text{min}}$ and maximum of maximal average distance M_{max} with corresponding values of $\hat{\tau}_{\text{max}}$ for some graph sizes N

N	M_{mean}	$\hat{\tau}_{\text{mean}}$	M_{min}	$\hat{\tau}_{\text{min}}$	M_{max}	$\hat{\tau}_{\text{max}}$
10	1,97	1,3	3,51	1,3	1,25	1,1
100	3,81	1,8	7,53	2,1	2,41	1,3
500	5,76	2,0	9,15	2,3	3,35	1,7
1000	6,65	2,3	11,81	2,3	3,45	1,9
2000	8,20	2,3	14,08	2,2	4,78	2,1
3000	8,70	2,4	13,61	2,4	5,01	2,2
4000	9,49	2,3	18,06	2,3	6,03	2,3
5000	9,81	2,3	16,59	2,6	6,03	2,1
6000	10,44	2,4	17,72	2,5	6,61	2,4
7000	10,57	2,4	16,91	2,4	7,06	2,3

Построены регрессионные зависимости максимумов среднего расстояния, минимального среднего расстояния и максимального

среднего расстояния от объема графа N :

$$M_{\text{mean}} = 1,405 \ln N - 2,519 \quad (R^2 = 0,96); \quad (6)$$

$$M_{\text{min}} = 0,807 \ln N - 1,231 \quad (R^2 = 0,91); \quad (7)$$

$$M_{\text{max}} = 2,394 \ln N - 3,806 \quad (R^2 = 0,91). \quad (8)$$

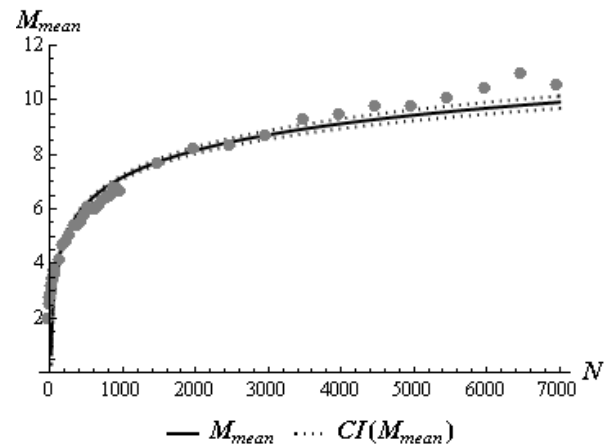


Рис. 3. Регрессионная зависимость M_{mean} от N

Fig. 3. Regression dependence of M_{mean} on N

На рис. 3, 4 и 5 сплошными линиями показаны регрессионные зависимости (6), (7) и

(8), точечными – доверительные интервалы этих зависимостей, серыми точками обозначены выборочные значения M_{mean} , M_{min} и M_{max} соответственно.

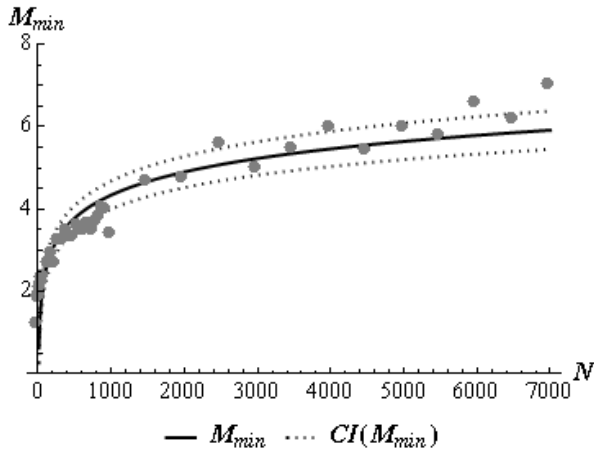


Рис. 4. Регрессионная зависимость M_{min} от N
Fig. 4. Regression dependence of M_{min} on N

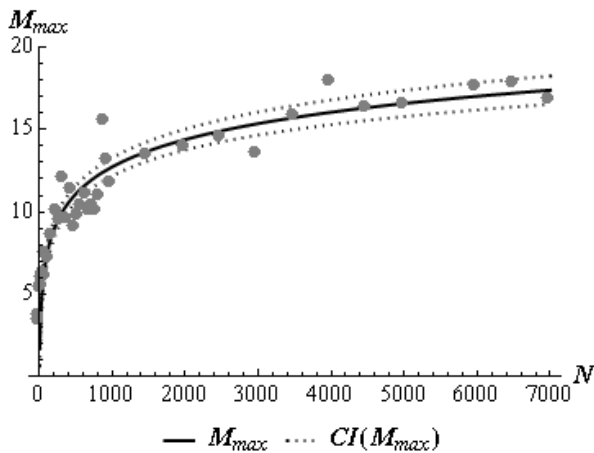


Рис. 5. Регрессионная зависимость M_{max} от N
Fig. 5. Regression dependence of M_{max} on N

С увеличением размера графа максимум среднего расстояния M_{mean} , а также максимумы минимального M_{min} и максимального M_{max} среднего расстояния растут логарифмически. Что касается соответствующих значений параметра τ , при которых эти максимумы достигаются $\hat{\tau}_{\text{mean}}$, $\hat{\tau}_{\text{min}}$ и $\hat{\tau}_{\text{max}}$, то при малых объемах графа они возрастают с ростом N , однако уже начиная с $N > 1000$ $\hat{\tau}_{\text{mean}}$, $\hat{\tau}_{\text{min}}$ и $\hat{\tau}_{\text{max}}$ варьируются в пределах $2 \leq \tau \leq 2,5$. По уравнениям (3)–(5) можно вычислить $\hat{\tau}_{\text{mean}} = 2,14$, $\hat{\tau}_{\text{min}} = 1,97$ и $\hat{\tau}_{\text{max}} = 2,28$, что также укладывается в данные пределы.

Исследование показало, что в степенных конфигурационных графах среднее расстоя-

ние принимает наименьшие значения на концах интервала изменения параметра распределения степеней вершин $\tau \in (1, 3]$, то есть чем ближе значение параметра τ к 1 или к 3, тем меньше среднее расстояние. Этот вывод также относится и к минимальному и максимальному среднему расстоянию в графе, причем разница между минимальным и максимальным средним расстоянием на концах интервала $\tau \in (1, 3]$ меньше, чем в его середине, т. е. эмпирический доверительный интервал среднего расстояния $dist$ в конфигурационном графе расширяется примерно в середине отрезка $1 < \tau \leq 3$ и сужается к его граничным значениям. Полученные результаты также показывают, что при $\tau < 1,5$ максимальное среднее расстояние (верхняя граница эмпирического доверительного интервала) не превышает 6 ребер, что соответствует так называемой «теории шести рукопожатий» (см., например, [4]). Чем ближе значение параметра τ к 2,4, тем больше будет среднее расстояние в конфигурационном графе. Так, наибольших значений среднее расстояние достигает при $2,3 \leq \tau \leq 2,4$, не превышая при этом 10–11 ребер, минимальное среднее расстояние достигает максимума при $\tau \approx 2$ и не превышает 7–8 ребер, а максимальное – при $2,3 \leq \tau \leq 2,5$ (для графов с числом вершин $N > 1000$), достигая длины в 18–19 ребер.

ЛИТЕРАТУРА

1. Bollobas B. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs // European J. Comb. 1980. Vol. 1, iss. 4. P. 311–316. doi: 10.1016/S0195-6698(80)80030-8
2. Chung F., Lu L. The average distances in random graphs with given expected degrees // Proc. of the National Acad. Sci. of the USA. 2002. Vol. 99, iss. 25. P. 15879–15882. doi: 10.1073/pnas.252631999
3. Dijkstra E. W. A note on two problems in connexion with graphs // Numer. Math. 1959. Vol. 1, iss. 1. P. 269–271. doi: 10.1007/BF01386390
4. Durrett R. Random graph dynamics. Cambridge: Cambridge Univ. Press, 2007. 221 p.
5. Faloutsos C., Faloutsos P., Faloutsos M. On power-law relationships of the internet topology // Comp. Comm. Rev. 1999. Vol. 29. P. 251–262. doi: 10.1145/316194.316229
6. Hofstad R. Random graphs and complex networks. Vol. One. Cambridge: Cambridge Univ. Press, 2017. 337 p. doi: 10.1017/9781316779422
7. Hofstad R. Random graphs and complex networks. Vol. 2. 2018. URL: <https://www.win.tue.nl/~rhofstad/NotesRGCNII.pdf> (дата обращения: 20.04.2022).

8. *Newman M. E. J.* Networks. An Introduction. Oxford: Oxford Univ. Press, 2010. 772 p. doi: 10.1093/acprof:oso/9780199206650.001.0001

9. *Newman M. E. J.* Networks. Second edition. Oxford: Oxford Univ. Press, 2018. 800 p. doi: 10.1093/oso/9780198805090.001.0001

10. *Reittu H., Norros I.* On the power-law random graph model of massive data networks // Performance Evaluation. 2004. Vol. 55, iss. 1-2. P. 3–23. doi: 10.1016/S0166-5316(03)00097-X

REFERENCES

1. *Bollobas B.* A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European J. Comb.* 1980;1(4):311–316. doi: 10.1016/S0195-6698(80)80030-8

2. *Chung F., Lu L.* The average distances in random graphs with given expected degrees. *Proc. of the National Acad. Sci. of the USA.* 2002;99(25):15879–15882. doi: 10.1073/pnas.252631999

3. *Dijkstra E. W.* A note on two problems in connexion with graphs. *Numer. Math.* 1959;1(1):269–271. doi: 10.1007/BF01386390

4. *Durrett R.* Random graph dynamics. Cambridge: Cambridge Univ. Press; 2007. 221 p.

5. *Faloutsos C., Faloutsos P., Faloutsos M.* On power-law relationships of the internet topology. *Comp. Comm. Rev.* 1999;29:251–262. doi: 10.1145/316194.316229

6. *Hofstad R.* Random graphs and complex networks. Vol. One. Cambridge: Cambridge Univ. Press; 2017. 337 p. doi: 10.1017/9781316779422

7. *Hofstad R.* Random graphs and complex networks. Vol. 2. 2018. URL: <https://www.win.tue.nl/~rhofstad/NotesRGCNII.pdf> (accessed: 20.04.2022).

8. *Newman M. E. J.* Networks. An Introduction. Oxford: Oxford Univ. Press; 2010. 772 p. doi: 10.1093/acprof:oso/9780199206650.001.0001

9. *Newman M. E. J.* Networks. Second edition. Oxford: Oxford Univ. Press; 2018. 800 p. doi: 10.1093/oso/9780198805090.001.0001

10. *Reittu H., Norros I.* On the power-law random graph model of massive data networks. *Performance Evaluation.* 2004;55(1-2):3–23. doi: 10.1016/S0166-5316(03)00097-X

Поступила в редакцию / received: 22.04.2022; принята к публикации / accepted: 11.05.2022.

Автор заявляет об отсутствии конфликта интересов / The author declares no conflict of interest.

СВЕДЕНИЯ ОБ АВТОРЕ:

Лери Марина Муксумовна
канд. техн. наук, научный сотрудник
e-mail: leri@krc.karelia.ru

CONTRIBUTOR:

Leri, Marina
Cand. Sci. (Tech.), Researcher