

УДК 519.179.2, 519.876.5

ДИАМЕТР КОНФИГУРАЦИОННЫХ ГРАФОВ СО СТЕПЕННЫМ РАСПРЕДЕЛЕНИЕМ

М. М. Лери

*Институт прикладных математических исследований КарНЦ РАН,
ФИЦ «Карельский научный центр РАН», Петрозаводск, Россия*

Рассматриваются случайные конфигурационные графы с дискретным степенным распределением степеней вершин. Параметр распределения принимает фиксированные значения. Посредством имитационного моделирования найдены оценки диаметра графа. Предложены зависимости диаметра от объема графа и параметра распределения степеней вершин.

Ключевые слова: конфигурационный граф; степенное распределение; диаметр графа; имитационное моделирование.

M. M. Leri. DIAMETER OF POWER-LAW CONFIGURATION GRAPHS

We consider random configuration graphs with a power-law distribution of vertex degrees. The distribution parameter takes fixed values. Estimates of the diameter of the graph were found by simulations. Relationships are suggested for the dependence of the graph diameter on its volume and the vertex degree distribution parameter.

Keywords: configuration graph; power-law distribution; graph diameter; simulations.

ВВЕДЕНИЕ

За последние десятилетия исследование структуры случайных графов стало одним из наиболее востребованных и перспективных направлений в силу того, что такие модели находят широкое применение для описания структуры реальных сетей (см., например, [6, 7, 9, 10]). Так, случайные графы, степени вершин которых представляют собой независимые одинаково распределенные случайные величины с общим законом распределения, нашли применение как модели для описания различных видов телекоммуникационных сетей, включая глобальную сеть Интернет (см., например, [7, 12]).

Понятно, что с течением времени и со стремительным развитием как информационно-коммуникационных технологий, так и техни-

ческих средств связи топология больших сетей быстро меняется. Поэтому при построении математических моделей таких сетей с целью дальнейшего изучения их функционирования необходимо принимать во внимание не только распределение степеней вершин случайного графа, используемого для моделирования сети, но и другие важные структурные характеристики. Изучению таких характеристик посвящено множество работ (см., например, [6, 10, 11]), однако поле деятельности еще далеко не исчерпано.

В настоящей работе рассматривается одна из таких структурных характеристик – диаметр графа. Исследование проводится посредством имитационного моделирования на так называемых конфигурационных графах, впервые предложенных Б. Боллобашем [2].

КОНФИГУРАЦИОННЫЕ ГРАФЫ

Пусть N – число вершин графа. Степени вершин $\xi_1, \xi_2, \dots, \xi_N$ являются независимыми одинаково распределенными случайными величинами со степенным распределением, заданным следующим равенством [12]:

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

где $k = 1, 2, \dots$, и $\tau > 1$ – параметр распределения степеней вершин. Для построения конфигурационной модели вначале задаются степени вершин графа в соответствии с распределением (1) с заранее заданным параметром τ , определяя для каждой вершины число различных полуребер [12], занумерованных в произвольном порядке. Сумма степеней вершин рассматриваемого графа является случайной величиной, и если она оказывается нечетной, то степень равновероятно выбранной вершины увеличивается на 1, тем самым добавляя к ней одно полуребро. Далее все полуребра графа соединяются между собой попарно и равновероятно для образования ребер. Ясно, что такое построение допускает появление в графе петель, циклов и кратных ребер.

Известно (см., например, [6, 9, 12]), что структура конфигурационных графов со степенным распределением степеней вершин зависит от параметра распределения τ . Так, при $1 < \tau < 2$ в таких графах существует единственная так называемая гигантская компонента связности, число вершин в которой пропорционально N при $N \rightarrow \infty$, а объем любой другой из компонент этого графа бесконечно мал по сравнению с объемом гигантской компоненты. Более того, в [1] показано, что чем ближе значение параметра τ к 1, тем выше вероятность того, что все вершины графа образуют одну компоненту связности, содержащую в себе все N вершин графа. При $2 \leq \tau \leq 3$ в конфигурационном графе все еще можно выделить компоненту связности, число вершин в которой будет наибольшим по сравнению с другими компонентами графа, при этом общее число компонент возрастает с ростом N . Через η_1 обозначим объем (число вершин) гигантской (или наибольшей) компоненты связности.

ДИАМЕТР ГРАФА

Одной из важных числовых характеристик сетевой структуры и, соответственно, случайного графа является диаметр. Исследованию этой характеристики случайных графов посвящены работы многих авторов (см., например, [3, 4, 8]), однако они рассматривали графы других видов, отличных от рассматриваемых в настоящей работе.

Обозначим $G = (V, E)$ неориентированный граф, где V – множество вершин, а E – множество ребер. Пусть $v, u \in V$ и $d(v, u)$ – длина простой цепи, соединяющей вершины v и u ($v \neq u$) графа G .

В конфигурационном графе:

- $d(v, v) = 1$, если при вершине v существует петля;
- $d(v, u) = \infty$, если вершины v и u принадлежат разным компонентам связности;
- если в графе есть кратные ребра, то может существовать более одной цепи, соединяющей вершину v с u , но длины этих цепей будут равны.

Расстоянием от вершины v до вершины u называют длину цепи наименьшей длины:

$$\nu(v, u) = \min_{u \in V} d(v, u)$$

среди всех $d(v, u) \neq \infty$. Через $\rho(v)$ обозначим максимальную длину кратчайшего пути до вершины v графа G :

$$\rho(v) = \max_{u \in V} \nu(v, u).$$

Для нахождения диаметра графа достаточно найти кратчайшие пути между всеми парами вершин, и наибольшая длина среди всех кратчайших путей и будет диаметром графа. В конфигурационных графах рассматриваются пути только внутри компонент связности. В настоящей работе для нахождения кратчайших путей был применен алгоритм Дейкстры [5]. Для фиксированной вершины v графа этот алгоритм находит $\rho(v)$ – максимум среди всех кратчайших путей до этой вершины от всех остальных вершин графа. Тогда диаметр графа $D(G)$ будет равен

$$D(G) = \max_{v \in V} \rho(v),$$

то есть наибольшему расстоянию между двумя любыми вершинами графа G .

РЕЗУЛЬТАТЫ

В работе было поставлено несколько целей. Во-первых, оценить диаметр конфигурационного графа D и найти зависимость D от числа вершин графа N и параметра распределения степеней вершин τ . Во-вторых, найти значения параметра τ (или интервалы значений), при которых диаметр конфигурационного графа принимает 1) наибольшие значения D_{\max} и 2) наименьшие значения D_{\min} . И в-третьих, найти зависимости объема наибольшей компоненты при максимуме диаметра $\eta_1(D_{\max})$

от N , а также объема наибольшей компоненты при минимуме диаметра $\eta_1(D_{\min})$ от N . Исследование проводилось с помощью методов имитационного моделирования с последующей статистической обработкой данных посредством ПО Wolfram Mathematica 9.0.

Рассматривались конфигурационные графы размерностей $1000 \leq N \leq 9500$ с шагом 500 (ограничение сверху продиктовано непредвиденным выходом из строя вычислительного кластера ЦКП). Значения параметра τ изменялись от 1,01 до 3,0 с шагами: 0,01 для $1,01 \leq \tau \leq 1,1$ и 0,1 для $1,1 \leq \tau \leq 3,0$. Для каждой пары значений (N, τ) генерировалось 100 графов, то есть всего 52 200 графов.

В результате проведенного исследования получена следующая регрессионная зависимость диаметра конфигурационного графа D от объема графа N и параметра распределения степеней вершин τ :

$$D = -5,27 + (1,48\tau + 0,05\tau^5 - 0,02\tau^6) \ln N \quad (2)$$

с коэффициентом детерминации $R_1^2 = 0,93$. Здесь и далее все коэффициенты регрессионных уравнений значимы на уровне значимости 0,05.

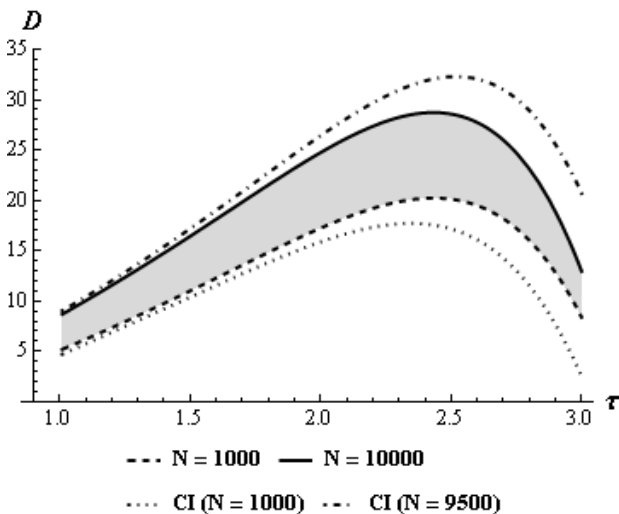


Рис. 1. Регрессионная зависимость D от N и τ
Fig. 1. Regression dependence of D on N and τ

На рис. 1 обозначены нижняя граница доверительного интервала зависимости (2) при $N = 1000$ ($CI(N = 1000)$) и верхняя граница при $N = 9500$ ($CI(N = 9500)$). Интервалы были построены с целью рассмотрения возможности экстраполяции прогнозных значений диаметра графа до объема $N = 10000$. Для большей наглядности на рис. 2 представлены усредненные (по 100 экспериментам) экспериментальные значения D в зависимости от N и τ .

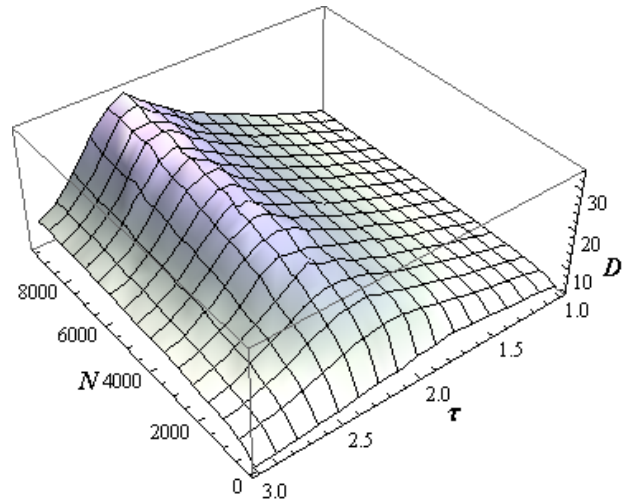


Рис. 2. График усредненных (по 100 экспериментам) экспериментальных значений D от N и τ
Fig. 2. Dependence of averaged experimental D on N and τ

Конфигурационные графы с параметром τ распределения (1) из интервала (1,2) представляют особый интерес вследствие того, что они подходят для моделирования сети Интернет в случае, когда в качестве узлов сети рассматриваются автономные системы [7, 12]. Поэтому была построена регрессионная зависимость D от N и τ для $1 < \tau < 2$:

$$D = 2,52 + 0,63\tau^{1,83} \ln N \quad (R_2^2 = 0,97). \quad (3)$$

Как и ранее, на рис. 3 обозначены нижняя граница доверительного интервала зависимости (3) при $N = 1000$ ($CI(N = 1000)$) и верхняя граница при $N = 9500$ ($CI(N = 9500)$). На рис. 4 представлена зависимость усредненных (по 100 экспериментам) экспериментальных значений D от N и $1 < \tau < 2$.

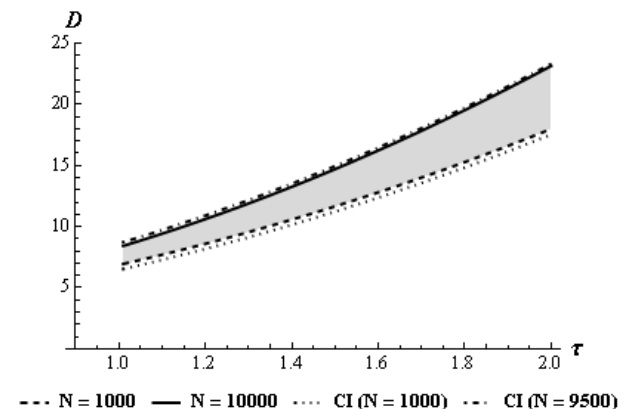


Рис. 3. Регрессионная зависимость D от N и $1 < \tau < 2$
Fig. 3. Regression dependence of D on N and $1 < \tau < 2$

Далее при оценке значимости различия между коэффициентами корреляции $r_1 = \sqrt{R_1^2}$ и $r_2 = \sqrt{R_2^2}$ регрессионных моделей (2) и (3) соответственно, получаем, что при уровне значимости 0,05 нулевая гипотеза $H_0 : r_1 = r_2$ отвергается, следовательно, различие между коэффициентами корреляции значимо.

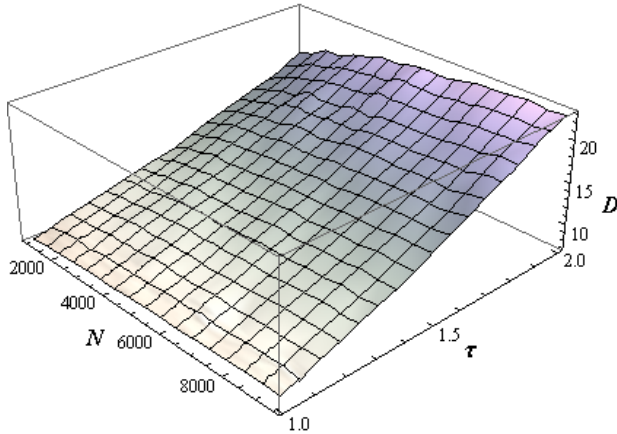


Рис. 4. График усредненных (по 100 экспериментам) экспериментальных значений D от N и $1 < \tau < 2$
 Fig. 4. Dependence of averaged experimental D on N and $1 < \tau < 2$

По экспериментальным данным для каждой пары (N, τ) вычислены средние значения диаметра D (из 100 экспериментов). Далее для каждого N были найдены D_{\max} и соответствующие им значения τ , при которых достигается этот максимум D , а также значения D_{\min} и соответствующие им значения τ , при которых этот минимум достигается (см. табл. 1 и 2 для D_{\max} и D_{\min} соответственно).

Таблица 1. Средние максимального значения диаметра D_{\max} для каждого значения N , соответствующие им значения τ и η_1
 Tab. 1. Average maximum diameter values D_{\max} for each N and corresponding values of τ and η_1

N	D_{\max}	$\tau(D_{\max})$	$\eta_1(D_{\max})$
1000	18,11	2,1	395,68
1500	20,48	2,2	486,00
2000	22,06	2,3	505,82
2500	24,84	2,3	656,65
3000	24,42	2,4	507,67
3500	26,17	2,3	971,09
4000	26,32	2,4	770,50
4500	28,25	2,3	1100,54
5000	29,24	2,4	829,89
5500	28,82	2,3	1424,23
6000	29,80	2,3	1537,77
6500	31,04	2,4	1261,07
7000	30,56	2,4	1349,24
7500	32,44	2,4	1396,67
8000	31,59	2,4	1469,00
8500	31,92	2,4	1698,38
9000	33,65	2,4	1784,50
9500	34,12	2,4	1733,56

Таблица 2. Средние минимального значения диаметра D_{\min} для каждого значения N , соответствующие им значения τ и η_1

Tab. 2. Average minimum diameter values D_{\min} for each N and corresponding values of τ and η_1

N	D_{\min}	$\tau(D_{\min})$	$\eta_1(D_{\min})$
1000	7,11	1,01	967,22
1500	7,29	1,02	1450,55
2000	7,44	1,01	1938,81
2500	7,52	1,01	2423,66
3000	7,70	1,01	2912,20
3500	7,76	1,01	3400,05
4000	7,65	1,01	3889,38
4500	7,83	1,01	4376,89
5000	7,90	1,01	4864,05
5500	7,99	1,01	5351,65
6000	7,96	1,01	5842,73
6500	7,91	1,01	6327,08
7000	8,05	1,02	6815,16
7500	8,10	1,01	7314,74
8000	8,17	1,01	7792,54
8500	8,15	1,02	8275,15
9000	8,10	1,02	8774,20
9500	8,14	1,01	9254,34

Построены регрессионные зависимости: D_{\max} от N (см. табл. 1 и рис. 5):

$$D_{\max} = 6,97 \ln N - 30,59 \quad (R^2 = 0,99)$$

и D_{\min} от N (см. табл. 2 и рис. 6):

$$D_{\min} = 0,47 \ln N + 3,83 \quad (R^2 = 0,99).$$

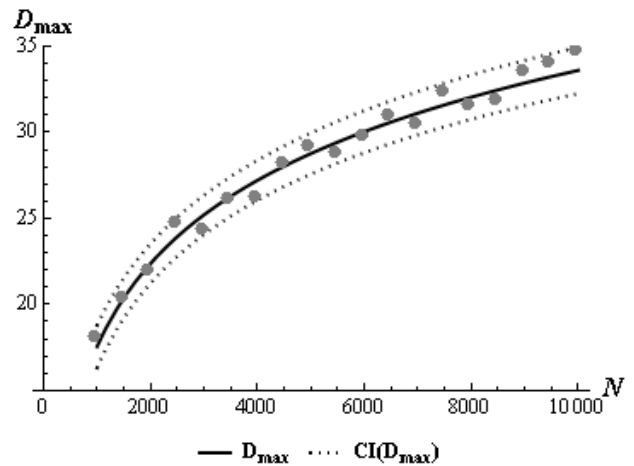


Рис. 5. Зависимость D_{\max} от N . Сплошной линией показана регрессионная зависимость, точечными линиями – доверительный интервал этой зависимости, серыми точками обозначены выборочные значения D_{\max}

Fig. 5. Dependence of D_{\max} on N . Solid line shows regression dependence, dotted lines – confidence interval of this dependence, gray dots mark sample values D_{\max}

Таким образом, в степенных конфигурационных графах с ростом объема графа N диа-

метр достигает наибольших размеров при значениях параметра $2.3 \leq \tau \leq 2.4$. При малых объемах ($N < 2000$) это значение τ может уменьшиться до 2.1 (см. значения $\tau(D_{\max})$ в табл. 1). Кроме того, вне зависимости от размера графа, чем ближе значение параметра τ к 1, тем меньше будет диаметр такого графа (см. значения $\tau(D_{\min})$ в табл. 2).

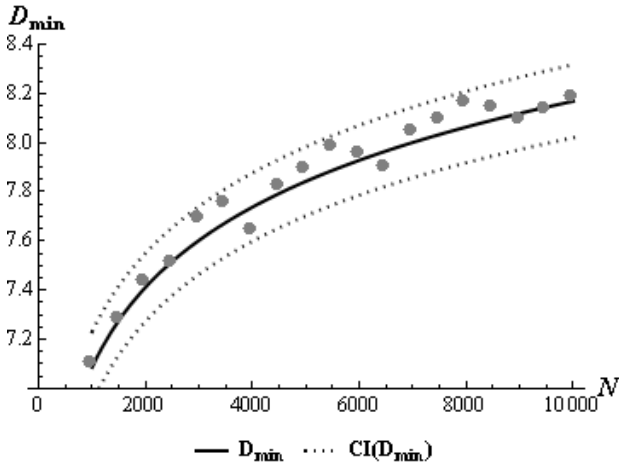


Рис. 6. Зависимость D_{\min} от N . Сплошной линией показана регрессионная зависимость, точечными линиями – доверительный интервал этой зависимости, серыми точками обозначены выборочные значения D_{\min}

Fig. 6. Dependence of D_{\min} on N . Solid line shows regression dependence, dotted lines – confidence interval of this dependence, gray dots mark sample values D_{\min}

Рассмотрим объемы наибольшей компоненты графа при максимуме и минимуме его диаметра. Средние значения этих величин приведены в последних столбцах таблиц 1 и 2: $\eta_1(D_{\max})$ и $\eta_1(D_{\min})$ соответственно. По этим данным легко вычислить значения этих же характеристик в %. Зависимости экспериментальных средних $\eta_1(D_{\max})$ и $\eta_1(D_{\max})$ в % от объема графа N показаны на рис. 7, а зависимости $\eta_1(D_{\min})$ и $\eta_1(D_{\min})$ в % от N – на рис. 8. Выборочные коэффициенты корреляции между значениями η_1 и размером графа N равны следующим значениям:

$$\begin{aligned} r(\eta_1(D_{\max}), N) &= 0,95; \\ r(\eta_1(D_{\max})\%, N) &= -0,67; \\ r(\eta_1(D_{\min}), N) &= 0,99; \\ r(\eta_1(D_{\min})\%, N) &= 0,91. \end{aligned}$$

Были построены следующие зависимости объема наибольшей компоненты графа от числа вершин N при максимуме и минимуме диаметра графа:

- 1) зависимость между $\eta_1(D_{\max})$ и N имеет вид линейной, размер наибольшей компоненты варьируется примерно от 390 до 2000 вершин (рис. 7, а), и регрессионная зависимость имеет вид:

$$\eta_1(D_{\max}) = 219,41 + 0,17N;$$

- 2) между $\eta_1(D_{\max})$ в % и объемом графа N зависимость обратно пропорциональна, значения варьируются примерно от 16 до 40 % (рис. 7, б), и можно построить регрессионную зависимость следующего вида:

$$\eta_1(D_{\max})\% = 83,79 - 7,24 \ln N;$$

- 3) получена выраженная линейная зависимость $\eta_1(D_{\min})$ от N (рис. 8, а):

$$\eta_1(D_{\min}) = 0,98N - 13,41;$$

- 4) значения объема наибольшей компоненты в % $\eta_1(D_{\min})$ варьируются в небольшом интервале от 96,6 до 97,6%, однако имеют тенденцию к увеличению (рис. 8, б), и можно построить регрессионную зависимость следующего вида:

$$\eta_1(D_{\min})\% = 94,14 + 0,37 \ln N.$$

Таким образом, анализируя полученные в настоящей работе результаты совместно с выводами из [1], можно сказать, что в конфигурационных графах со степенным распределением степеней вершин чем ближе значение параметра τ этого распределения к 1, тем меньше диаметр соответствующего графа. В этом случае графы являются более «связными», то есть в них существует гигантская компонента связности, в которую входит более 95 % вершин, однако кратчайшие расстояния между вершинами этой компоненты невелики, откуда и следует малый размер диаметра таких графов. Это также соответствует и так называемой «теории шести рукопожатий» (см., например, [6]). С ростом значения параметра τ объем наибольшей компоненты графа уменьшается, однако его диаметр увеличивается, что связано, по-видимому, с появлением более длинных простых цепочек вершин. Максимального значения длина таких цепей, а соответственно, и диаметр графа, достигает примерно в середине следующего интервала изменения параметра: $\tau \in [2, 3]$, а затем начинает уменьшаться с приближением к его правой границе. Размер наибольшей компоненты в этом случае еще больше уменьшается, а число компонент связности значительно увеличивается.

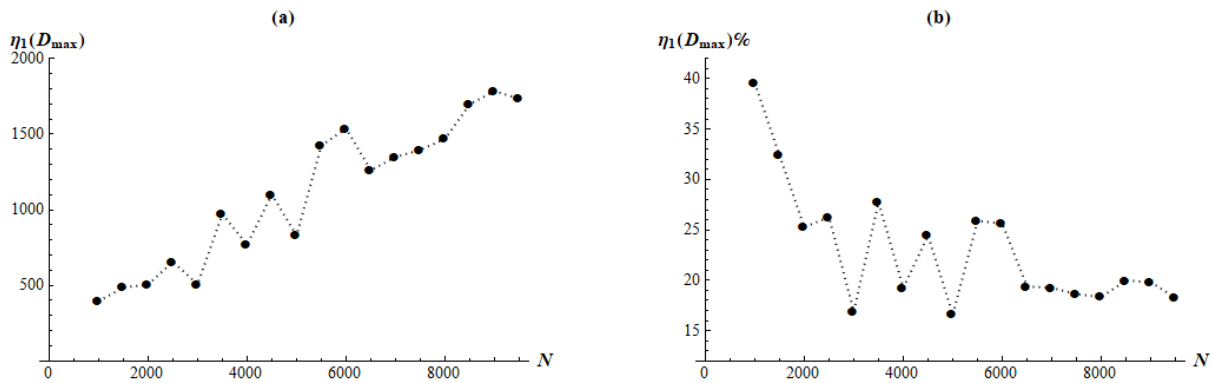


Рис. 7. Зависимости: (а) $\eta_1(D_{\max})$ и (б) $\eta_1(D_{\max})$ в % от N
 Fig. 7. Dependences of: (а) $\eta_1(D_{\max})$ and (б) $\eta_1(D_{\max})$ in % on N

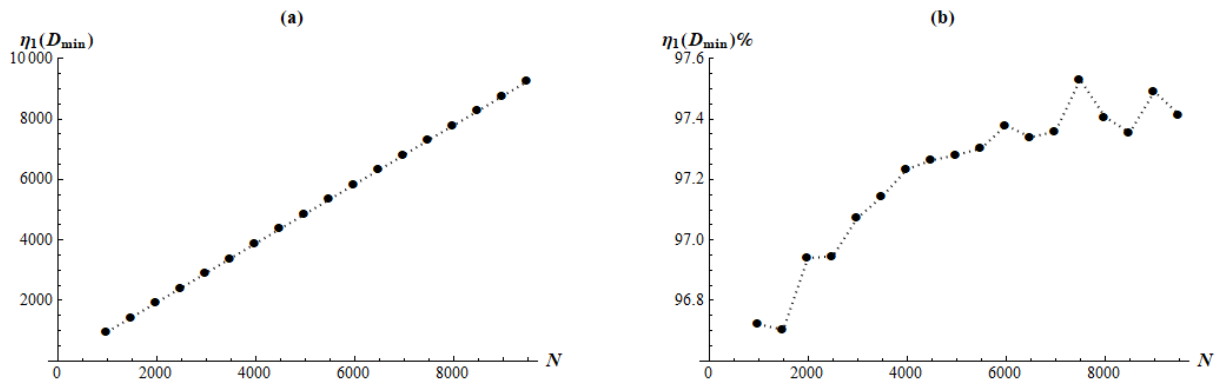


Рис. 8. Зависимости: (а) $\eta_1(D_{\min})$ и (б) $\eta_1(D_{\min})$ в % от N
 Fig. 8. Dependences of: (а) $\eta_1(D_{\min})$ and (б) $\eta_1(D_{\min})$ in % on N

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

Исследования выполнены на научном оборудовании Центра коллективного пользования Федерального исследовательского центра «Карельский научный центр Российской академии наук».

ЛИТЕРАТУРА

1. Лери М. М. Об условиях связности конфигурационных графов // Труды КарНЦ РАН. 2020. № 7. С. 71–75. doi: 10.17076/mat1227
2. Bollobas B. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs // Eur. J. Combin. 1980. Vol. 1, iss. 4. P. 311–316. doi: 10.1016/S0195-6698(80)80030-8

3. Bollobas B., Riordan O. M. The diameter of a scale-free random graph // Combinatorica. 2004. Vol. 24, no. 1. P. 5–34. doi: 10.1007/s00493-004-0002-2
4. 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
5. 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
6. Durrett R. Random graph dynamics. Cambridge: Cambridge Univ. Press, 2007. 221 p.
7. 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
8. Fernholz D., Ramachandran V. The diameter of sparse random graphs // Random Struct.

Alg. 2007. Vol. 31, iss. 4. P. 482–516.
doi: 10.1002/rsa.20197

9. *Hofstad R.* Random graphs and complex networks. Vol. 1. Cambridge: Cambridge Univ. Press, 2017. 337 p. doi: 10.1017/9781316779422

10. *Hofstad R.* Random graphs and complex networks. 2018. Vol. 2. 314 p. URL: <https://www.win.tue.nl/~rhofstad/NotesRGCNII.pdf> (дата обращения: 25.04.2021).

11. *Newman M. E. J.* The structure and function of complex networks // *SIAM Rev.* 2003. Vol. 45, iss. 2. P. 167–256. doi: 10.1137/S003614450342480

12. *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

Поступила в редакцию 26.04.2021

REFERENCES

1. *Leri M. M.* Ob usloviyakh svyaznosti konfiguratsionnykh grafov [On conditions of configuration graphs' connectivity]. *Trudy KarNTs RAN* [Trans. KarRC RAS]. 2020. No. 7. P. 71–75. doi: 10.17076/mat1227

2. *Bollobas B.* A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *Eur. J. Combin.* 1980. Vol. 1, iss. 4. P. 311–316. doi: 10.1016/S0195-6698(80)80030-8

3. *Bollobas B., Riordan O. M.* The diameter of a scale-free random graph. *Combinatorica.* 2004. Vol. 24, no. 1. P. 5–34. doi: 10.1007/s00493-004-0002-2

4. *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

5. *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

6. *Durrett R.* Random graph dynamics. Cambridge: Cambridge Univ. Press, 2007. 221 p.

7. *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

8. *Fernholz D., Ramachandran V.* The diameter of sparse random graphs. *Random Struct. Alg.* 2007. Vol. 31, iss. 4. P. 482–516. doi: 10.1002/rsa.20197

9. *Hofstad R.* Random graphs and complex networks. Vol. 1. Cambridge: Cambridge Univ. Press, 2017. 337 p. doi: 10.1017/9781316779422

10. *Hofstad R.* Random graphs and complex networks. 2018. Vol. 2. 314 p. URL: <https://www.win.tue.nl/~rhofstad/NotesRGCNII.pdf> (accessed: 25.04.2021).

11. *Newman M. E. J.* The structure and function of complex networks. *SIAM Rev.* 2003. Vol. 45, iss. 2. P. 167–256. doi: 10.1137/S003614450342480

12. *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

Received April 26, 2021

СВЕДЕНИЯ ОБ АВТОРЕ:

Лери Марина Муксумовна

научный сотрудник, к. т. н.
Институт прикладных математических исследований
КарНЦ РАН, Федеральный исследовательский центр
«Карельский научный центр РАН»
ул. Пушкинская, 11, Петрозаводск,
Республика Карелия, Россия, 185910
эл. почта: leri@krc.karelia.ru
тел.: (8142) 781218

CONTRIBUTOR:

Leri, Marina

Institute of Applied Mathematical Research,
Karelian Research Centre,
Russian Academy of Sciences
11 Pushkinskaya St., 185910 Petrozavodsk,
Karelia, Russia
e-mail: leri@krc.karelia.ru
tel.: (8142) 781218