Труды Карельского научного центра РАН
№ 6. 2021. С. 59–63
DOI: 10.17076/mat1423

УДК 519.179.4

АСИМПТОТИКА ЧИСЛА РЕБЕР ИНТЕРНЕТ-ГРАФА

Ю. Л. Павлов

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

Рассматриваются конфигурационные графы с N вершинами. Степени вершин независимы и одинаково распределены по закону, зависящему от неизвестной медленно меняющейся функции. Степень каждой вершины имеет конечное математическое ожидание и бесконечную дисперсию. Такие модели можно использовать для описания топологии различных сетей коммуникаций и сети Интернет. В статье доказана локальная предельная теорема для числа ребер графа при $N \to \infty$.

Ключевые слова: конфигурационный граф; степень вершины; предельное распределение; медленно меняющаяся функция; число ребер.

Yu. L. Pavlov. ASYMPTOTICS OF THE NUMBER OF EDGES OF AN INTERNET GRAPH

We consider configuration graphs with N vertices. The degrees of the vertices are independent and identically distributed according to a distribution law that depends on an unknown slowly varying function. The degree of a vertex has a finite expectation and an infinite variance. Such models can be used to describe various communication networks and Internet topologies. The paper proves the local limit theorem for the number of edges in a graph as $N \to \infty$.

Keywords: configuration graph; vertex degree; limit distribution; slowly varying function; number of edges.

Введение

Для моделирования сложных сетей коммуникаций широко используются методы теории случайных графов (см., например, [5]). Наблюдения за реальными сетями показали, что большинство из них обладают схожими свойствами. К таким сетям относятся транспортные, электрические, телефонные, социальные сети, сети сотрудничества ученых и, конечно, сеть Интернет. Оказалось, что число узлов сети, имеющих степень не меньше, чем k, при достаточно больших k обычно пропорционально $k^{-\tau}$, где $\tau > 0$. Более того, было показано, что в случайных графах, служащих моделями таких сетей, степени вершин можно считать независимыми одинаково распределенными целочисленными случайными величинами. Обозначим ξ случайную величину, равную степени любой вершины графа. Результаты наблюдений позволили установить, что распределение ξ можно задать следующим образом:

$$\mathbf{P}\{\xi \geqslant k\} = \frac{h(k)}{k^{\tau}}, \quad k = 1, 2, \ldots, \quad \tau > 0, \quad (1)$$

где $h(x)$ – медленно меняющаяся функция. По определению (см., например, [1]), медленно меняющаяся функция определена для всех $x \geqslant 0$ и положительна. Следовательно, как видно из (1), $\mathbf{P}\{\xi = k\} > 0$ для всех натуральных k. Обозначим

$$p_k = \mathbf{P}\{\xi = k\}$$

$$= \frac{h(k)}{k^\tau} - \frac{h(k+1)}{(k+1)^\tau}, \quad k = 1, 2, \ldots \qquad (2)$$

Было обнаружено, что для большинства сетей значения параметра τ принадлежат интервалу $(1, 2)$. Из (1) и (2) следует, что

$$m = \mathbf{E}\xi = \sum_{k=1}^{\infty} kp_k$$

$$= \sum_{k=1}^{\infty} \left(\frac{h(k)}{k^{\tau-1}} - \frac{h(k+1)}{(k+1)^{\tau-1}} + \frac{h(k+1)}{(k+1)^\tau} \right).$$

Это значит, что при $\tau \in (1, 2)$ распределение (1) имеет конечное математическое ожидание m. Аналогично можно показать, что дисперсия ξ бесконечна.

Одним из наиболее часто используемых для моделирования сетей видов случайных графов является так называемый конфигурационный граф. Конструкция конфигурационных графов была предложена в [4]. Пусть граф содержит N вершин. Степень каждой вершины равна числу выходящих из нее полуребер, т. е. ребер, для которых смежные вершины еще не определены. Все полуребра различимы. Ребра графа образуются путем попарного равновероятного соединения полуребер друг с другом. Для обеспечения четности суммы степеней вершин в случае необходимости в граф вводится вспомогательная вершина единичной степени или дополнительное полуребро добавляется к какой-нибудь равновероятно выбранной вершине. Известно [6], что появление такой вершины вместе с инцидентным ей ребром не влияет на асимптотические свойства графа при $N \to \infty$. Нетрудно видеть, что такая конструкция графа допускает появление петель и кратных ребер. Конфигурационные графы с распределением (1) независимых степеней вершин при $\tau \in (1, 2)$ нередко используются для моделирования сети Интернет, поэтому иногда их называют Интернет-графами.

Во многих публикациях рассматриваются графы, в которых известен явный вид медленно меняющейся функции $h(x)$ распределения (1). В частности, в [6] предполагалось, что

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

откуда следует, как нетрудно проверить, что $h(x) = 1$. В ряде работ рассматривались условные конфигурационные графы при условии, что известны не только явный вид медленно меняющейся функции, но и число ребер. Кроме того, в некоторых публикациях предполагалось, что параметр распределения степеней вершин τ является случайной величиной с заданным распределением. В этом случае получались конфигурационные графы со случайными распределениями случайных степеней вершин. В статье [3] проведено обобщение таких работ за счет рассмотрения моделей, в которых при $k \to \infty$

$$p_k \sim \frac{d}{k^g (lnk)^q},$$

где $d > 0$, $g \geqslant 1$, $q \geqslant 0$, $g + q > 1$. Это значит, что медленно меняющаяся функция сводилась к степени логарифма. Подробное исследование конфигурационных графов с неизвестной функцией $h(x)$ только начинается (см., например, [2]). Один из успешно применяемых методов исследования свойств условных конфигурационных графов основан на использовании обобщенной схемы размещения частиц по ячейкам. Рассматривались также условные графы при условии, что число ребер точно не известно, но ограничено сверху. Такие модели соответствуют сетям, в которых число возможных связей имеет технические ограничения. В таких случаях используется аналог обобщенной схемы, отличающийся тем, что сумма участвующих в схеме независимых случайных величин не равна точно известной величине, а ограничена. Применительно к конфигурационным графам эти суммы являются суммами степеней вершин графа, т. е. удвоенному числу ребер. Таким образом для получения асимптотических результатов с помощью обобщенной схемы размещения необходимо исследовать предельное поведение числа ребер.

В статье рассматриваются конфигурационные графы с N вершинами и распределением (1) случайных независимых степеней вершин при $\tau \in (1, 2)$. В следующем разделе найдены предельные распределения наибольших членов вариационного ряда степеней вершин и числа вершин заданной степени. Главным результатом статьи является доказанная в последнем разделе локальная предельная теорема для числа ребер такого графа.

Структура степеней вершин

Обозначим $\xi_1, \ldots, \xi_N$ степени вершин $1, \ldots, N$ соответственно. Все эти случайные величины независимы и одинаково распределены по закону (1). Рассмотрим вариационный ряд

$$\xi_{(1)} \leqslant \xi_{(2)} \leqslant \ldots \leqslant \xi_{(N)},$$

полученный для случайных величин $\xi_1, \ldots, \xi_N$ расположением их в неубывающем порядке.

Теорема 1. *Пусть $N \to \infty$ и последовательность $r = r(N)$ выбрана так, что*

$$\frac{Nh(r)}{r^\tau} \to \gamma,$$

где γ – некоторая положительная постоянная. Тогда для любого фиксированного целого неотрицательного s

$$\mathbf{P}\{\xi_{(N-s)} \leqslant r\} \to e^{-\gamma} \sum_{i=0}^{s} \frac{\gamma^i}{i!}.$$

Доказательство. Из (1) находим, что

$$\mathbf{P}\{\xi_{(N-s)} \leqslant r\}$$

$$= \sum_{i=0}^{s} \binom{N}{i} \mathbf{P}^{N-i}\{\xi \leqslant r\} \mathbf{P}^{i}\{\xi > r\}$$

$$= \sum_{i=0}^{s} \frac{N!}{i!(N-i)!} \left(1 - \frac{h(r)}{r^\tau}\right)^{N-i} \frac{h^i(r)}{r^{i\tau}}. \qquad (3)$$

Из условий теоремы следует, что $r \to \infty$ и для любого фиксированного i

$$\frac{N!}{i!(N-i)!} \left(1 - \frac{h(r)}{r^\tau}\right)^{N-i} \frac{h^i(r)}{r^{i\tau}}$$

$$= \frac{1}{i!} \left(\frac{Nh(r)}{r^\tau}\right)^{i} \left(1 - \frac{h(r)}{r^\tau}\right)^{N} (1 + o(1))$$

$$\to \frac{\gamma^i e^{-\gamma}}{i!}.$$

Отсюда и из (3) получаем утверждение теоремы 1. □

Теорема 2. *Пусть $N \to \infty$. Справедливы следующие утверждения.*

1. *Если r фиксировано, то равномерно относительно $u_r = (k - Np_r)/\sqrt{Np_r(1-p_r)}$ в любом фиксированном конечном интервале*

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

2. *Если $r \to \infty$, то равномерно относительно $(k - Np_r)/\sqrt{Np_r}$ в любом фиксированном конечном интервале*

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

Доказательство. Нетрудно видеть, что

$$\mathbf{P}\{\mu_r = k\} = \binom{N}{k}$$

$$\times \mathbf{P}\{\xi_1 = r, \ldots, \xi_k = r, \xi_{k+1} \neq r, \ldots, \xi_N \neq r\}$$

$$= \binom{N}{k} p_r^k (1 - p_r)^{N-k}. \qquad (4)$$

Если r фиксировано, то $Np_r(1-p_r) \to \infty$. Следовательно, для оценки вероятности (4) можно использовать нормальное приближение биномиальной вероятности

$$\binom{N}{k} p_r^k (1 - p_r)^{N-k}$$

$$= \frac{1 + o(1)}{\sqrt{2\pi Np_r(1-p_r)}} \exp\left\{-\frac{(k - Np_r)^2}{2Np_r(1-p_r)}\right\},$$

справедливое равномерно относительно $(k - Np_r)/\sqrt{Np_r(1-p_r)}$ в любом фиксированном конечном интервале. Отсюда и из (4) следует первое утверждение теоремы 2.

Если $r \to \infty$, то для доказательства второго утверждения теоремы 2 достаточно к вероятности (4) применить пуассоновское приближение биномиальной вероятности, которое при $p_r \to 0$ также выполняется равномерно относительно $(k - Np_r)/\sqrt{Np_r}$ в любом фиксированном конечном интервале. □

Предельное распределение числа ребер

Обозначим ζ_N сумму степеней вершин графа: $\zeta_N = \xi_1 + \ldots + \xi_N$. Разумеется, число ребер графа равно половине суммы степеней, поэтому для оценки предельного поведения числа ребер достаточно найти асимптотику ζ_N.

Введем стремящуюся к бесконечности при $N \to \infty$ последовательность B_N, удовлетворяющую условию

$$B_N \sim (Nh(B_N))^{1/\tau}. \qquad (5)$$

Пусть $g(x)$ означает плотность распределения устойчивого закона с показателем τ и характеристической функцией

$$\varphi(t) = \exp\left\{-c|t|^\tau \left(1 - i\frac{t}{|t|} \tan\frac{\pi\tau}{2}\right)\right\}, \qquad (6)$$

где

$$c = -\frac{\Gamma(2-\tau)}{\tau - 1} \cos\frac{\pi\tau}{2}, \qquad (7)$$

$\Gamma(x)$ – гамма-функция.

Теорема 3. *Пусть $N \to \infty$. Тогда*

$$\sup_k \left| B_N \mathbf{P}\{\zeta_N = k\} - g\left(\frac{k - Nm}{B_N}\right) \right| \to 0.$$

Доказательство. Обозначим $F_\xi(x)$ функцию распределения случайной величины ξ. Из (1) видно, что при $x < 0$

$$F_\xi(x) = 0, \tag{8}$$

а если $x \geqslant 0$, то

$$F_\xi(x) = 1 - \sum_{k \geqslant x} p_k. \tag{9}$$

Из (1) и (2) следует, что

$$\sum_{k \geqslant x} p_k = \frac{h(\lceil x \rceil)}{(\lceil x \rceil)^\tau}.$$

Отсюда и из определения медленно меняющейся функции получаем, что при $x \to \infty$

$$\sum_{k \geqslant x} p_k = \frac{h(x)}{x^\tau}(1 + o(1)),$$

поэтому из (9) находим, что

$$F_\xi(x) = 1 - \frac{h(x)}{x^\tau}(1 + o(1)). \tag{10}$$

Согласно теореме 2.6.1 книги [1], для того чтобы закон распределения $F(x)$ принадлежал области притяжения устойчивого закона с показателем τ, $0 < \tau < 2$, необходимо и достаточно, чтобы при $|x| \to \infty$

$$F(x) = \frac{c_1 + o(1)}{|x|^\tau} w(|x|), \quad x < 0,$$

$$F(x) = 1 - \frac{c_2 + o(1)}{x^\tau} w(x), \quad x > 0,$$

где $w(x)$ – медленно меняющаяся функция, а c_1, c_2 – константы такие, что $c_1 \geqslant 0$, $c_2 \geqslant 0$, $c_1 + c_2 > 0$. Соотношения (8) и (10) означают, что функция $F_\xi(x)$ принадлежит области притяжения некоторого устойчивого закона $G(x)$ с показателем τ, при этом $c_1 = 0$, $c_2 = 1$, $w(x) = h(x)$. Найдем явный вид характеристической функции $\varphi_G(t)$ функции распределения $G(x)$. Для этого используем теорему 2.2.2 [1]. В указанной теореме установлены необходимые и достаточные условия того, чтобы функция распределения была устойчивой. Эти условия представляют собой общий вид логарифма характеристической функции распределения. Зависимость параметров данной функции от τ, c_1, c_2 также найдена в ходе доказательства теоремы 2.2.2. Используя эти результаты, получаем, что логарифм характеристической функции $\varphi_G(t)$ устойчивого закона $G(x)$ имеет вид:

$$\ln \varphi_G(t) = idt - c|t|^\tau \left(1 - i\frac{t}{|t} \tan \frac{\pi\tau}{2}\right), \tag{11}$$

где d – некоторая константа, а c определено в (7). В (11) параметр d можно сделать равным нулю, подобрав для $G(x)$ нужным образом нормирующие множители в определении области притяжения. Обозначим $\varphi_\xi(t)$ характеристическую функцию случайной величины ξ. Поскольку математическое ожидание m случайной величины ξ конечно, из теоремы 2.6.5 [1] следует, что в достаточно малой окрестности нуля

$$\ln \varphi_G(t)$$

$$= itm - c|t|^\tau q(t) \left(1 - i\frac{t}{|t} \tan \frac{\pi\tau}{2}\right), \tag{12}$$

где функция $q(t)$ при $t \to 0$ является медленно меняющейся. В ходе доказательства теоремы 2.6.5 [1] показано, что при $t \to 0$

$$q(t) = h\left(\frac{1}{|t|}\right). \tag{13}$$

Очевидно, что $\varphi_\xi(0) = 1$. Пусть $t \neq 0$ и фиксировано. Из (5), (12), (13) находим, используя определение медленно меняющейся функции, что при $N \to \infty$

$$\varphi_\xi^N\left(\frac{t}{B_N}\right) \exp\left\{-\frac{itNm}{B_N}\right\}$$

$$= \exp\left\{-c\frac{|t|^\tau}{h(B_N)} h\left(\frac{B_N}{|t|}\right)\left(1 - \frac{t}{|t|} \tan \frac{\pi\tau}{2}\right)\right\}$$

$$\to \exp\left\{-c|t|^\tau \left(1 - i\frac{t}{|t|} \tan \frac{\pi\tau}{2}\right)\right\}.$$

Это соотношение показывает, что распределение суммы $(\zeta_N - Nm)/B_N$ слабо сходится к устойчивому закону с показателем τ, плотностью распределения $g(x)$ и характеристической функцией (6). Для того чтобы доказать локальную сходимость к этому закону и, следовательно, завершить доказательство теоремы 3, можно воспользоваться теоремой 4.2.1 [1]. Согласно этой теореме, для того, чтобы локальная сходимость имела место, необходимо и достаточно, чтобы функция распределения $F_\xi(x)$ принадлежала области притяжения устойчивого закона $G(x)$ и шаг распределения ξ был максимальным. Принадлежность области притяжения была доказана выше. В распределении (1) медленно меняющаяся функция $h(x)$ положительна при всех $x \geqslant 0$, поэтому шаг распределения ξ равен единице и максимален. $\square$

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

Литература

1. *Ибрагимов И. А., Линник Ю. В.* Независимые и стационарно связанные величины. М.: Наука, 1965. 524 с.

2. *Павлов Ю. Л.* Максимальное дерево случайного леса в конфигурационном графе // Математический сборник. 2021. Т. 212 (принято к публикации).

3. *Павлов Ю. Л.* Условные конфигурационные графы со случайным параметром распределения степеней // Математический сборник. 2018. Т. 209, вып. 2. С. 120–137. doi: 10.4213/sm8832

4. *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

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

6. *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/S0166-53/6(3)00097-x

Поступила в редакцию 28.03.2021

References

1. *Ibragimov I. A., Linnik Yu. V.* Independent and stationary sequences of random variables. Groningen: Wolters-Noordhoff, 1971. 443 p.

2. *Pavlov Yu. L.* Maksimal'noe derevo sluchainogo lesa v konfiguratsionnom grafe [Maximum random forest tree in a configuration graph] *Matematicheskii sbornik* [Math. Coll. Articles]. 2021. Vol. 212 (accepted for publication).

3. *Pavlov Yu. L.* Conditional configuration graphs with discrete power-law distribution of vertex degrees. *SB MATH.* 2018. Vol. 209, iss. 2. P. 258–275. doi: 10.1070/SM8832

4. *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

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

6. *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/S0166-53/6(3)00097-x

Received March 28, 2021

СВЕДЕНИЯ ОБ АВТОРЕ:

Павлов Юрий Леонидович
главный научный сотрудник, д. ф.-м. н., профессор
Институт прикладных математических исследований КарНЦ РАН, Федеральный исследовательский центр «Карельский научный центр РАН»
ул. Пушкинская, 11, Петрозаводск,
Республика Карелия, Россия, 185910
эл. почта: pavlov@krc.karelia.ru
тел.: (8142) 781218

CONTRIBUTOR:

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