Math-Net.Ru

Общероссийский математический портал

Ю. Л. Павлов, Предельное распределение объема гигантской компоненты в случайном графе Интернет-типа, *Дискрет. матем.*, 2007, том 19, выпуск 3, 22–34

DOI: <https://doi.org/10.4213/dm963>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением

<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 3.81.55.215

26 апреля 2023 г., 15:12:05

Предельное распределение объема гигантской компоненты в случайном графе Интернет-типа

© 2007 г. Ю. Л. Павлов

Изучается случайный граф, который может служить моделью больших систем передачи информации, включая Интернет. Приведены условия существования гигантской компоненты в таком графе и доказано, что предельное распределение этой компоненты является нормальным.

1. Введение

В последние годы появился ряд статей, посвященных изучению случайных графов, моделирующих функционирование больших сетей передачи данных (см., например, [1]–[3] и ссылки на другие источники в этих работах). Такие модели можно использовать для описания электрических и телефонных сетей, сетей телекоммуникаций, включая Интернет, социальных связей и даже взаимной цитируемости в научных публикациях. Несмотря на то, что подобные сети имеют огромные размеры, типичное расстояние между объектами сети достаточно мало, поэтому их иногда называют малыми мирами. Еще одной особенностью соответствующих графов является то, что число вершин степени, не меньшей k , пропорционально $k^{-\tau}$, где $\tau > 0$. В связи с этим, в таких моделях предполагается, что степени вершин графа являются независимыми случайными величинами с одинаковым распределением, являющимся дискретным аналогом распределения Парето. Обозначим ξ случайную величину, равную степени произвольной вершины, и положим

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

Пусть N означает число вершин в графе. Во многих статьях при $N \rightarrow \infty$ изучалось предельное поведение различных характеристик случайных графов, в которых число степеней каждой вершины имеет распределение (1), при этом оказалось, что свойства графов существенно зависят от значения параметра τ . Наблюдения за реальными сетями (см. [1], [2]) показали, что предложенные модели оказались адекватными, если $\tau \in (1, 2)$, а при других значениях τ поведение графов в большинстве случаев не соответствует действительности. Заметим, что при $\tau \in (1, 2)$ распределение (1) имеет конечное математическое ожидание и бесконечную дисперсию.

В [2] предполагалось, что $\tau \in (1, 2)$. Для удобства в [2] введено понятие полуребра, то есть ребра, инцидентного конкретной вершине, но для которой смежная вершина еще

не определена. Все полуребра графа, как и вершины, считаются различными (занумерованными). Процесс формирования графа состоит из двух этапов. На первом этапе для каждой вершины определяется ее степень, то есть число выходящих из нее полуребер, предполагается, что степени вершин являются реализациями независимых копий случайной величины ξ . На втором этапе полуребра, выходящие из всех вершин, соединяются между собой равновероятно; два полуребра, соединившись, образуют ребро. Легко видеть, что такая конструкция графа допускает петли и кратные ребра. Понятно, что для построения графа необходимо, чтобы общее число полуребер было четным. Для обеспечения этого условия искусственно вводят дополнительную вершину, степень которой равна 0 или 1 в зависимости от того, является суммарное число полуребер основных вершин четным или нет. Как отмечено в [2], введение этой искусственной вершины не влияет на предельное поведение графа.

Во многих работах (см., например, [1], [2]) указываются возможные варианты содержательной интерпретации рассмотренной модели случайного графа. Например, при моделировании Интернета вершинам графа могут соответствовать локальные сети, кратные ребра означают одновременную передачу данных несколькими компьютерами одной локальной сети в другую, а петля соответствует подключению одного компьютера к другому через общий маршрутизатор внутри одной локальной сети. В связи с этим такие случайные графы иногда называют графами Интернет-типа.

В статье [2] при $N \rightarrow \infty$ изучалось предельное поведение ряда характеристик случайных графов Интернет-типа. Заметим, что все полученные в [2] результаты справедливы с вероятностью, стремящейся к единице. Перечислим некоторые из них. Найден интервал, которому принадлежит сумма степеней вершин. Получена оценка снизу для максимальной степени вершины графа. Показано, что если степень вершины меньше, чем N^γ , где $\gamma < 1/2$, то у такой вершины нет петель, а если степень больше, чем N^γ , где $\gamma > 1/2$, то есть хотя бы одна петля. Компонента связности графа, содержащая вершину с максимальной степенью, названа гигантской, поскольку ее объем (то есть число входящих в нее вершин) имеет математическое ожидание cN , где c — некоторая положительная постоянная, $c < 1$. Найдена верхняя граница расстояния между двумя произвольно выбранными вершинами гигантской компоненты. Под расстоянием между вершинами понимается, как обычно, минимальное число ребер, образующих простую цепь, соединяющую эти вершины. Оказалось, что такое расстояние имеет порядок $\ln \ln N$ и, следовательно, пренебрежимо мало по сравнению с объемом гигантской компоненты. Для получения этой оценки и определения величины c была использована известная (см. [2], [4]) связь между случайными графами и ветвящимися процессами. Доказана локальная предельная теорема для расстояния между двумя произвольно выбранными вершинами графа.

В [5], [6] получена локальная предельная теорема для суммы степеней вершин графа и найдены предельные распределения крайних членов вариационного ряда степеней вершин, расположенных в неубывающем порядке. В [7] приведена локальная предельная теорема для числа петель вершины при условии, что степень этой вершины пропорциональна $\sqrt{N}$.

Основным результатом настоящей статьи является теорема 1, в которой показано, что предельное распределение объема гигантской компоненты является нормальным. Статья написана под влиянием работы [2] и ее целью, кроме доказательства теоремы 1, было развитие идей и уточнение результатов статьи [2]. Такое развитие заключалось, в основном, в получении предельных распределений рассмотренных характеристик. Новыми результатами такого типа можно считать леммы 1 и 2 о предельном поведении членов вариационного ряда степеней вершин, лемму 3 о сумме степеней вершин, лемму 4 о числе

петель вершин, теорему 2 об объеме центральной части (ядра) графа и теорему 3 об объемах подграфов, входящих в состав ядра, а также лемму 6, уточняющую степень близости рассмотренных ветвящихся процессов. Леммы 5 и 7 только уточняют известные результаты (см. [2]), указывая скорость сближения рассмотренных там вероятностей. Описанный ниже алгоритм построения графа отличается от соответствующих процедур из [2] только формой изложения. Леммы 1–3 получены автором совместно с М. М. Степановым [5]. Второе утверждение леммы 4 принадлежит М. М. Степанову [7].

Статья состоит из семи разделов. В разделе 2 дано описание рассматриваемых графов и приведена формулировка основной теоремы. В разделе 3 доказываются вспомогательные утверждения, сформулированные в виде лемм 1–4. В разделе 4 описывается структура так называемого ядра графа — главной части гигантской компоненты. В разделе 5 приводится алгоритм построения графа. Связь между случайным графом и ветвящимися процессами рассматривается в разделе 6. Раздел 7 содержит доказательство теоремы 1.

2. Основной результат

Пусть случайный граф G содержит N вершин, имеющих номера $1, 2, \dots, N$. Обозначим $\xi_1, \xi_2, \dots, \xi_N$ независимые случайные величины, равные степеням вершин $1, 2, \dots, N$ соответственно и имеющие общее распределение (1), где $\tau \in (1, 2)$. Из (1) ясно, что

$$p_k = \mathbf{P}\{\xi_1 = k\} = k^{-\tau} - (k+1)^{-\tau}, \quad k = 1, 2, \dots \quad (2)$$

Рассмотрим множество графов $G(k_1, k_2, \dots, k_N)$, для которых

$$\xi_1 = k_1, \xi_2 = k_2, \dots, \xi_N = k_N.$$

Обозначим g некоторую реализацию графа G . Ясно, что

$$\mathbf{P}\{g \in G(k_1, k_2, \dots, k_N)\} = p_{k_1} p_{k_2} \dots p_{k_N}.$$

Поскольку все соединения полуребер равновероятны, естественно считать, что и все графы в $G(k_1, k_2, \dots, k_N)$ равновероятны, следовательно,

$$\mathbf{P}\{g \mid g \in G(k_1, k_2, \dots, k_N)\} = \frac{1}{(l-1)(l-3) \dots 3 \cdot 1},$$

где $l = k_1 + \dots + k_N$. Таким образом, вероятность появления конкретного графа $g \in G(k_1, k_2, \dots, k_N)$ равна

$$\mathbf{P}\{g\} = \frac{p_{k_1} p_{k_2} \dots p_{k_N}}{(l-1)(l-3) \dots 3 \cdot 1}. \quad (3)$$

Обозначим $\varphi(z)$ производящую функцию распределения (2), то есть

$$\varphi(z) = \sum_{k=1}^{\infty} p_k z^k. \quad (4)$$

Используя (2), нетрудно получить, что

$$\mathbf{M}\xi_1 = \zeta(\tau), \quad (5)$$

где

$$\zeta(\tau) = \sum_{k=1}^{\infty} k^{-\tau} \quad (6)$$

— значение дзета-функции Римана в точке τ .

Введем случайную величину ξ^* , для которой

$$p_k^* = \mathbf{P}\{\xi^* = k\} = (k+1)p_{k+1}/\zeta(\tau), \quad k = 0, 1, 2, \dots, \quad (7)$$

и пусть

$$F(z) = \sum_{k=0}^{\infty} p_k^* z^k. \quad (8)$$

Рассмотрим вспомогательный, начинающийся с одной частицы, ветвящийся процесс Гальтона–Ватсона $\mu(t)$ с распределением (7) числа прямых потомков каждой частицы. Нетрудно видеть, что при $\tau \in (1, 2)$ случайная величина ξ^* имеет бесконечное математическое ожидание, следовательно, процесс $\mu(t)$ является надкритическим. Обозначим q вероятность вырождения этого процесса. Хорошо известно [8], что q удовлетворяет уравнению

$$q = F(q) \quad (9)$$

и $0 < q < 1$.

Обозначим $D^{(N)}$ событие, состоящее в том, что произвольная (но фиксированная) вершина графа принадлежит гигантской компоненте. Известно (см. [2], [4], а также раздел 7), что при $N \rightarrow \infty$

$$\mathbf{P}\{D^{(N)}\} \rightarrow 1 - \varphi(q). \quad (10)$$

Для случайной величины ν , равной объему гигантской компоненты, справедливо следующее утверждение.

Теорема 1. Пусть $N \rightarrow \infty$. Тогда

$$\mathbf{P}\left\{\frac{N\mathbf{P}\{D^{(N)}\} - \nu}{\sqrt{N\varphi(q)(1-\varphi(q))}} < x\right\} \rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-y^2/2} dy.$$

3. Вспомогательные утверждения

Построим вариационный ряд

$$\xi_{(1)} \leq \xi_{(2)} \leq \dots \leq \xi_{(N)}, \quad (11)$$

расположив $\xi_1, \xi_2, \dots, \xi_N$ в неубывающем порядке. Рассмотрим m -е экстремумы $\xi_{(m+1)}, \xi_{(N-m)}$ этого ряда при фиксированных целых неотрицательных m . Асимптотическое поведение порядковых статистик хорошо изучено (см., например, [9], [10]). Приведенные ниже леммы 1 и 2 фактически уточняют результаты [9], [10] для ряда (11).

Лемма 1. Пусть $N \rightarrow \infty$ и m — фиксированное натуральное число. Тогда

$$\mathbf{P}\{\xi_{(m)} = 1\} \rightarrow 1.$$

Доказательство. Из (1) следует, что

$$\begin{aligned} \mathbf{P}\{\xi_{(m)} = 1\} &= 1 - \mathbf{P}\{\xi_{(m)} \geq 2\} \\ &= 1 - \sum_{k=0}^{m-1} \binom{N}{k} \mathbf{P}\{\xi_1 < 2, \dots, \xi_k < 2, \xi_{k+1} \geq 2, \dots, \xi_N \geq 2\} \\ &= 1 - \sum_{k=0}^{m-1} \frac{N!}{k!(N-k)!} (1 - 2^{-\tau})^k 2^{-(N-k)\tau} \rightarrow 1. \end{aligned}$$

Лемма 2. Пусть $N \rightarrow \infty$ и m — фиксированное целое неотрицательное число. Тогда

$$\mathbf{P}\{\tau \ln \xi_{(N-m)} - \ln N < x\} = \frac{\Gamma(m+1, e^{-x})}{m!} + o(1),$$

где

$$\Gamma(w, v) = \int_v^\infty y^{w-1} e^{-y} dy.$$

Доказательство. С учетом (1) нетрудно видеть, что

$$\begin{aligned} \mathbf{P}\{\tau \ln \xi_{(N-m)} - \ln N < x\} &= \mathbf{P}\{\xi_{(N-m)} < (Ne^x)^{1/\tau}\} \\ &= \sum_{k=0}^m \binom{N}{k} \mathbf{P}\{\xi_1 < (Ne^x)^{1/\tau}, \dots, \xi_{N-k} < (Ne^x)^{1/\tau}, \xi_{N-k+1} \geq (Ne^x)^{1/\tau}, \dots, \xi_N \geq (Ne^x)^{1/\tau}\} \\ &= \sum_{k=0}^m \frac{N!}{(N-k)!k!} \left(1 - \frac{1}{Ne^x}\right)^{N-k} \frac{1}{(Ne^x)^k}. \end{aligned}$$

Применяя формулу Стирлинга и известные свойства дополнительной неполной гамма-функции (см. [11]), отсюда получаем, что

$$\mathbf{P}\{\tau \ln \xi_{(N-m)} - \ln N < x\} = e^{-e^{-x}} \sum_{k=0}^m \frac{e^{-kx}}{k!} + o(1) = \Gamma(m+1, e^{-x})/m! + o(1). \quad (12)$$

Полагая $m = 0$, из леммы 2 и (12) получаем следующее утверждение.

Следствие 1. Справедливо равенство

$$\mathbf{P}\{\xi_{(N)} < (Ne^x)^{1/\tau}\} = e^{-e^{-x}} + o(1).$$

Рассмотрим предельное поведение суммы $\eta_N = \xi_1 + \dots + \xi_N$.

Лемма 3. При $N \rightarrow \infty$

$$\sup_k \left| N^{1/\tau} \mathbf{P}\{\eta_N = k\} - g\left(\frac{k - N\zeta(\tau)}{N^{1/\tau}}\right) \right| \rightarrow 0,$$

где $g(x)$ — плотность устойчивого распределения с показателем τ и характеристической функцией

$$\exp \left\{ \frac{\Gamma(2-\tau)}{\tau-1} \cos \frac{\pi\tau}{2} |t|^\tau \left(1 + i \frac{t}{|t|} \operatorname{tg} \frac{\pi\tau}{2} \right) \right\}.$$

Доказательство. Обозначим $H(x)$ функцию распределения случайной величины ξ_1 . Из теоремы 2.1.1 книги [12] и из (5) следует, что если $H(x)$ принадлежит области притяжения устойчивого закона с показателем τ , то такой закон является предельным для случайной величины $(\eta_N - N\zeta(\tau))/N^{1/\tau}$. Из (1) следует, что при $|x| \rightarrow \infty$

$$H(x) = \begin{cases} 0, & x \leq 0, \\ 1 - x^{-\tau} + o(x^{-\tau}), & x > 0. \end{cases} \quad (13)$$

Согласно теореме 2.6.1 из [12], условия (13) достаточно для того, чтобы $H(x)$ принадлежала области притяжения указанного распределения. Тогда утверждение леммы следует из теоремы 4.2.1 из [12], а вид характеристической функции ясен из теоремы 2.2.2 из [12].

Выберем произвольную вершину графа, например, вершину 1 и обозначим λ число ее петель.

Лемма 4. При $N \rightarrow \infty$ справедливы следующие утверждения:

1. $\mathbf{P}\{\lambda = 0 \mid \xi_1 = o(\sqrt{N})\} \rightarrow 1$.
2. $\mathbf{P}\{\lambda = m \mid \xi_1 = y\sqrt{N}\} \rightarrow \alpha^m e^{-\alpha} / m!$, $m = 0, 1, 2, \dots$, $0 < y < \infty$, $\alpha = y^2 / (2\zeta(\tau))$.
3. $\mathbf{P}\{\lambda \geq 1 \mid \xi_1 / \sqrt{N} \rightarrow \infty\} \rightarrow 1$.

Доказательство. Ясно, что

$$\mathbf{P}\{\lambda = m \mid \xi_1 = u\} = \frac{\mathbf{P}\{\lambda = m, \xi_1 = u\}}{\mathbf{P}\{\xi_1 = u\}}, \quad (14)$$

$$\mathbf{P}\{\lambda = m, \xi_1 = u\} = \sum_{l=N}^{\infty} \mathbf{P}\{\lambda = m, \xi_1 = u, \eta_N = l\}. \quad (15)$$

Из (3), (14) и (15) следует, как нетрудно видеть, что

$$\mathbf{P}\{\lambda = m \mid \xi_1 = u\} = \sum_{l=N}^{\infty} \mathbf{P}\{\eta_{N-1} = l - u\} R_m, \quad (16)$$

где

$$R_m = \frac{1}{m!} \binom{u}{2} \binom{u-2}{2} \dots \binom{u-2m+2}{2} \frac{(l-u)(l-u-1)\dots(l-2u+2m+1)}{(l-1)(l-3)\dots(l-2u+2m+1)}. \quad (17)$$

В силу очевидного неравенства $R_m \leq 1$, из (16) и леммы 3 находим, что

$$\mathbf{P}\{\lambda = m \mid \xi_1 = u\} = \sum \mathbf{P}\{\eta_{N-1} = l - u\} R_m + \delta, \quad (18)$$

где суммирование проводится по всем l таким, что

$$N\zeta(\tau) - AN^{1/\tau} \leq l - u \leq N\zeta(\tau) + AN^{1/\tau},$$

а величина δ может быть сделана сколь угодно малой выбором достаточно больших N , A . Положим $m = 0$, $u = y\sqrt{N}$. Если $y \rightarrow 0$, то из (17) прямыми вычислениями нетрудно получить, что $R_0 \rightarrow 1$, а если $y \rightarrow \infty$, то $R_0 \rightarrow 0$. В силу леммы 3 и (18), это завершает доказательство утверждений 1 и 3. Для того чтобы убедиться в справедливости утверждения 2, достаточно с помощью формулы Стирлинга преобразовать (17) к виду $R_m = \alpha^m / (m! e^\alpha) + o(1)$ и опять применить лемму 3 и (18).

4. Ядро

Обозначим i^* вершину графа, имеющую максимальную степень $\xi_{(N)}$. Если таких вершин несколько, то в качестве i^* можно взять любую из них. В [2] показано, что вершины, имеющие достаточно большие степени, концентрируются вокруг i^* , соединены с ней и находятся на небольшом (порядка $\ln \ln N$) расстоянии от i^* . Множество таких вершин названо ядром графа. В этом разделе, следуя (2), мы кратко опишем структуру ядра. Прямыми следствиями полученных в [2] результатов являются доказанные ниже теоремы 2 и 3, в которых указаны предельные распределения объемов ядра и составляющих его подмножеств.

Введем возрастающую функцию $l(N)$ такую, что $l(1) = 1$ и при $N \rightarrow \infty$

$$\frac{l(N)}{\ln \ln \ln N} \rightarrow 0, \quad \frac{l(N)}{\ln \ln \ln \ln N} \rightarrow \infty. \quad (19)$$

Положим также

$$\varepsilon(N) = l(N) / \ln N. \quad (20)$$

Очевидно, что при $N \rightarrow \infty$

$$\varepsilon(N) \rightarrow 0, \quad N^{\varepsilon(N)} \rightarrow \infty. \quad (21)$$

Пусть

$$s = \left\lceil \frac{\ln \ln N}{-\ln(\tau - 1)} \right\rceil, \quad (22)$$

здесь $\lceil x \rceil$ означает наименьшее натуральное число, большее или равное x .

Далее через i будем обозначать произвольную вершину графа, ясно, что вершина i может иметь любой номер от 1 до N .

Ядром называется множество

$$C = \{i: \xi_i \geq N^Q\}, \quad (23)$$

где

$$Q = \frac{(\tau - 1)^s}{\tau} + \varepsilon(N) \sum_{k=0}^{s-1} (\tau - 1)^k + \varepsilon(N). \quad (24)$$

Введем убывающую последовательность

$$b_j = \frac{(\tau - 1)^j}{\tau} + \varepsilon(N) \sum_{k=0}^{j-1} (\tau - 1)^k, \quad j = 1, 2, \dots \quad (25)$$

Введем обозначения

$$\begin{aligned} U_1 &= \{i: \xi_i \geq N^{b_1 + \varepsilon(N)}\}, \\ U_j &= \{i: N^{b_j + \varepsilon(N)} \leq \xi_i < N^{b_{j-1} + \varepsilon(N)}\}, \quad j = 2, 3, \dots, s. \end{aligned} \quad (26)$$

Из следствия 1 и (25), (26) ясно, что $i^* \in U_1$. Кроме того, очевидно, что $C = \bigcup_{j=1}^s U_j$. В [2] показано, что при $N \rightarrow \infty$ с вероятностью, стремящейся к единице, ядро является связным подграфом и расстояние между любыми двумя вершинами ядра не превосходит $2s$. Также показано, что любая вершина из U_{j+2} соединена ребром с какой-нибудь вершиной из U_{j+1} , но не имеет прямых соединений с U_j . Таким образом, непересекающиеся множества U_j расположены концентрически вокруг центральной вершины i^* и удаляются от нее с ростом j .

Введем случайные величины $\nu(C)$ и $\nu(U_j)$, равные числу вершин в множествах C и U_j соответственно. Из (1), (22)–(26) следует, что

$$\begin{aligned} p(C) &= \mathbf{P}\{i \in C\} = N^{-Q\tau}, \\ p(U_1) &= \mathbf{P}\{i \in U_1\} = N^{-(b_1 + \varepsilon(N))\tau}, \\ p(U_j) &= \mathbf{P}\{i \in U_j\} = N^{-(b_j + \varepsilon(N))\tau} - N^{-(b_{j-1} + \varepsilon(N))\tau}, \quad j = 2, 3, \dots, s. \end{aligned} \quad (27)$$

Нетрудно видеть, что

$$\mathbf{P}\{\nu(C) = k\} = \binom{N}{k} p^k(C) (1 - p(C))^{N-k}. \quad (28)$$

Используя (1), (23) и нормальное приближение биномиального распределения, из (27), (28) получаем следующее утверждение.

Теорема 2. Пусть $N \rightarrow \infty$. Тогда для целых неотрицательных k равномерно относительно $u = (k - Np(C)) / \sqrt{Np(C)(1 - p(C))}$ в любом фиксированном конечном интервале

$$\mathbf{P}\{\nu(C) = k\} = \frac{1 + o(1)}{\sqrt{2\pi Np(C)(1 - p(C))}} e^{-u^2/2}.$$

Аналогичным образом получаем следующую теорему.

Теорема 3. Пусть $N \rightarrow \infty$. Тогда для целых неотрицательных k равномерно относительно $u_j = (k - Np(U_j)) / \sqrt{Np(U_j)(1 - p(U_j))}$, $j = 1, 2, \dots, s$, в любом фиксированном конечном интервале

$$\mathbf{P}\{\nu(U_j) = k\} = \frac{1 + o(1)}{\sqrt{2\pi Np(U_j)(1 - p(U_j))}} e^{-u_j^2/2}.$$

5. Алгоритм построения графа

Пусть на первом этапе формирования графа определено число полуребер, выходящих из каждой вершины, то есть

$$\xi_1 = k_1, \dots, \xi_N = k_N.$$

Рассмотрим возможную процедуру формирования графа на втором этапе, начинающуюся с некоторой фиксированной вершины, скажем, с вершины 1. Для полуребра начальной вершины, имеющего наименьший номер, из всех других полуребер графа выбираем равновероятно одно и, соединив их, образуем первое ребро графа. В дальнейшем рассмотрении эти полуребра не участвуют. Заметим, что если второе полуребро, как и первое, инцидентно начальной вершине, то образуется петля. Далее, если у начальной вершины

еще остались свободные полуребра, то для такого полуребра с наименьшим номером опять ищем пару и образуем второе ребро. Если второе ребро соединит вершину 1 с той же вершиной, что и первое, то появляются кратные ребра. Эта процедура повторяется до тех пор, пока не будут исчерпаны все полуребра начальной вершины. Все вершины, соединенные таким образом с вершиной 1 и не совпадающие с ней, естественно называть первым поколением вершин. Обозначим $Z_1^{(N)}$ число вершин первого поколения. Понятно, что если $Z_1^{(N)} = 0$, то вершина 1 имеет только петли и работа алгоритма на этом заканчивается. После образования первого поколения вершин по той же схеме проводится формирование второго поколения. Для этого сначала берется вершина первого поколения с наименьшим номером и для первого (по номеру) свободного полуребра этой вершины (если такое есть) среди всех оставшихся свободных полуребер графа равновероятно ищется пара и образуется ребро. Если ребро соединит две вершины первого поколения, то в графе образуется цикл. Если же новое ребро соединит вершину первого поколения с вершиной, не принадлежащей этому поколению, то образуется путь длины 2, соединяющий эту вершину с начальной, поэтому ее естественно считать вершиной второго поколения. Далее процесс образования ребер продолжается до исчерпания свободных полуребер первой вершины первого поколения. Затем процедура последовательно повторяется для всех оставшихся вершин первого поколения. Обозначим $Z_2^{(N)}$ число вершин второго поколения. После построения второго поколения аналогично формируется третье и так далее. Этот процесс заканчивается, когда число вершин очередного поколения оказывается равным нулю. Обозначим $Z_t^{(N)}$ число вершин t -го поколения, $t = 1, 2, \dots$

Описанную процедуру построения графа, начинающуюся с конкретной вершины, далее будем называть просто алгоритмом. Понятно, что в построенном с помощью этого алгоритма графе могут быть петли, кратные ребра, циклы и он может не содержать все N вершин. В последнем случае построенный граф является компонентой связности некоторой реализации рассматриваемого случайного графа, содержащей начальную вершину. Для того, чтобы построить всю реализацию, можно выбрать какую-то из вершин, не вошедших в первую компоненту связности и, считая ее начальной, снова использовать алгоритм. Повторив эту процедуру до тех пор, пока не будут исчерпаны все N вершин, мы достигнем цели.

Нетрудно видеть, что любой элемент множества реализаций случайного графа можно построить по предложенной схеме, при этом вероятность каждой реализации определяется равенством (3).

6. Связь с ветвящимися процессами

Рассмотрим два начинающихся с одной частицы ветвящихся процесса B и B^* . Будем считать, что число прямых потомков начальной частицы этих процессов имеет распределение (2). Процесс B поставим в соответствие рассматриваемому случайному графу, полагая, что число прямых потомков каждой частицы процесса (за исключением начальной) имеет случайное распределение

$$p_k(B) = \frac{1}{\eta_N} \sum_{i=1}^N \xi_i \mathbf{I}(\xi_i = k + 1) = \frac{k + 1}{\eta_N} \sum_{i=1}^N \mathbf{I}(\xi_i = k + 1), \quad (29)$$

где $k = 0, 1, 2, \dots$, а $\mathbf{I}(A)$ означает индикатор события A .

Распределение числа прямых потомков частиц процесса B^* (кроме начальной) задается равенствами (7). Обозначим Z_t и Z_t^* число частиц t -го поколения процессов B и B^* соответственно, $t = 1, 2, \dots$

Обозначим $A_t^{(N)}$ событие, состоящее в том, что построенный с помощью алгоритма подграф до поколения t включительно является деревом и при этом степень любой вершины меньше, чем N^Q . Связь между случайным графом и ветвящимися процессами можно описать с помощью следующей леммы.

Лемма 5. Пусть $N \rightarrow \infty$ и $t \leq \ln \ln \ln \ln N$. Тогда справедливы следующие утверждения.

1. $\mathbf{P}\{A_t^{(N)}\} = 1 + O(t / \ln \ln \ln N)$.
2. $\mathbf{P}\{Z_t^{(N)} = m\} = \mathbf{P}\{Z_t = m\} + O(t / \ln \ln \ln N)$, $m = 0, 1, 2, \dots$

Доказательство. Очевидно, что

$$\begin{aligned} \mathbf{P}\{A_t^{(N)}\} &= \mathbf{P}\{A_t^{(N)}, A_{t-1}^{(N)}, \dots, A_1^{(N)}\} \\ &= \mathbf{P}\{A_1^{(N)}\} \mathbf{P}\{A_2^{(N)} \mid A_1^{(N)}\} \dots \mathbf{P}\{A_t^{(N)} \mid A_{t-1}^{(N)}, \dots, A_1^{(N)}\}. \end{aligned}$$

Последовательную оценку вероятностей, являющихся сомножителями правой части этого равенства, можно провести, используя непосредственные вычисления. Они элементарны, но весьма громоздки, поэтому здесь не приводятся. В ходе этих вычислений необходимо проводить и оценку вероятностей вида $\mathbf{P}\{Z_t^{(N)} = m\}$, что, с учетом (29) и лемм 3, 4, приводит к соотношениям

$$\mathbf{P}\{A_t^{(N)}\} = (1 - x)^t, \quad \mathbf{P}\{Z_t^{(N)} = m\} = \mathbf{P}\{Z_t = m\}(1 - x)^t,$$

где $m = 0, 1, 2, \dots$, $x > 0$ и $x = O(N^{Q(1-\tau)})$. Отсюда и из (19)–(22), (24) очевидным образом следуют оба утверждения леммы.

Заметим, что связь между случайным графом и ветвящимся процессом B является более тонкой, чем это утверждается леммой 5. При тщательном рассмотрении доказательства этой леммы становится ясно, что каждой реализации ветвящегося процесса соответствует целый класс построенных с помощью алгоритма деревьев, отличающихся друг от друга только нумерацией вершин. Эта ситуация подобна установленному В. А. Ватутиным [13] потраекторному соответствию между случайными деревьями с занумерованными вершинами и ветвящимися процессами Гальтона–Ватсона с пуассоновским распределением числа потомков частиц.

Рассмотрим теперь связь между процессами B и B^* . Используя (2), (6), (7), (29), лемму 3, неравенство Чебышева и независимость входящих в правую часть (29) индикаторов, легко получить следующее утверждение.

Лемма 6. При достаточно больших N и $k = 0, 1, 2, \dots$

$$\mathbf{P}\{|p_k(B) - p_k^*| > N^{(1-\tau)/2\tau}\} < c N^{-1/\tau},$$

где c — некоторая положительная постоянная.

Применяя (2), (7), (29) и лемму 6, для ветвящихся процессов B и B^* можно получить утверждение, подобное лемме 5, при этом доказательство проводится аналогично. Это значит, что при $N \rightarrow \infty$ и $t \leq \ln \ln \ln \ln N$ в первых t поколениях этих процессов с

вероятностью, стремящейся к единице, нет частиц, имеющих более чем N^Q прямых потомков. Кроме того, при этих же условиях для любых $m = 0, 1, 2, \dots$

$$\mathbf{P}\{Z_t = m\} = \mathbf{P}\{Z_t^* = m\} + O(t / \ln \ln \ln N). \quad (30)$$

Сопоставляя лемму 5 и (30), убеждаемся в справедливости следующего результата.

Лемма 7. Пусть $N \rightarrow \infty$ и $t \leq \ln \ln \ln \ln N$. Тогда для $m = 0, 1, 2, \dots$

$$\mathbf{P}\{Z_t^{(N)} = m\} = \mathbf{P}\{Z_t^* = m\} + O(t / \ln \ln \ln N).$$

7. Доказательство теоремы 1

Найдем сначала вероятность вырождения процесса B^* . Следуя [8], введем обозначение

$$F_t^*(z) = \sum_{k=0}^{\infty} \mathbf{P}\{Z_t^* = k\} z^k, \quad t = 0, 1, 2, \dots$$

Пусть

$$F_t(z) = \sum_{k=0}^{\infty} \mathbf{P}\{\mu(t) = k\} z^k,$$

где $\mu(t)$, $t = 0, 1, 2, \dots$, — ветвящийся процесс Гальтона–Ватсона, введенный в разделе 2. Нетрудно видеть, что

$$F_{t+1}^*(z) = \varphi(F_t(z)).$$

Полагая $z = 0$, отсюда получаем, что

$$\mathbf{P}\{Z_{t+1}^* = 0\} = \varphi(\mathbf{P}\{\mu(t) = 0\}),$$

где производящая функция $\varphi(z)$ определена в (4). Переходя в этом равенстве к пределу при $t \rightarrow \infty$ и учитывая (8), (9), находим вероятность вырождения B^* :

$$\lim_{t \rightarrow \infty} \mathbf{P}\{Z_t^* = 0\} = \varphi(q). \quad (31)$$

Обозначим $D_{t,i}^{(N)}$ событие, состоящее в том, что к моменту t закончилось построение связной компоненты графа с помощью алгоритма, то есть $Z_t^{(N)} = 0$, при этом вершина i является начальной и в этом подграфе нет вершин из ядра. Ясно, что

$$\mathbf{P}\{D_{t,i}^{(N)}\} \leq \mathbf{P}\{Z_t^{(N)} = 0\}. \quad (32)$$

Положим $t = [\ln \ln \ln \ln N]$, где $[x]$ означает целую часть числа x . Из леммы 5 следует, что

$$\mathbf{P}\{D_{t,i}^{(N)}\} \geq \mathbf{P}\{D_{t,i}^{(N)}, A_t^{(N)}\} = \mathbf{P}\{Z_t^{(N)} = 0\} + O(t / \ln \ln \ln N). \quad (33)$$

Используя лемму 7 и (30)–(33), нетрудно проверить соотношение (10).

Рассмотрим сумму

$$S_N = \sum_{i=1}^N \sigma_{N,i}, \quad (34)$$

где

$$\sigma_{N,i} = \frac{\overline{\mathbf{I}(D_i^{(N)})} - \mathbf{P}\{\overline{D_i^{(N)}}\}}{\sqrt{N\varphi(q)(1-\varphi(q))}}, \quad (35)$$

а $D_i^{(N)}$ означает событие, состоящее в том, что вершина i принадлежит гигантской компоненте, $i = 1, 2, \dots, N$. Используя неравенство Чебышева в виде

$$\mathbf{P}\{|\xi - \mathbf{M}\xi| > \varepsilon\} < \frac{\mathbf{M}|\xi - \mathbf{M}\xi|^3}{\varepsilon^3},$$

получаем, что для любого $\varepsilon > 0$ при $N \rightarrow \infty$

$$\sum_{i=1}^N \mathbf{P}\{|\sigma_{N,i}| > \varepsilon\} \rightarrow 0. \quad (36)$$

Очевидно также, что в силу (35)

$$\sum_{i=1}^N \mathbf{M}(\sigma_{N,i} \mathbf{I}(\sigma_{N,i} \leq \varepsilon)) \rightarrow 0, \quad (37)$$

и, кроме того,

$$\sum_{i=1}^N \mathbf{D}(\sigma_{N,i} \mathbf{I}(\sigma_{N,i} \leq \varepsilon)) \rightarrow 1. \quad (38)$$

Из (36)–(38) следует, как нетрудно проверить, учитывая (3), что для суммы (34) выполнены условия центральной предельной теоремы для серий сумм зависимых случайных величин (теорема 1, §8, гл. VII книги [14]). В силу (10), (34) и (35) это означает, что число вершин графа, не принадлежащих гигантской компоненте, в пределе имеет нормальное распределение с параметрами $(N\mathbf{P}\{\overline{D^{(N)}}\}, \sqrt{N\varphi(q)(1-\varphi(q))})$. Отсюда непосредственно следует утверждение теоремы 1. Заметим, что естественная замена $\mathbf{P}\{D^{(N)}\}$ в формулировке теоремы 1 на $1 - \varphi(q)$ (см. (10)) не удастся, поскольку в силу (33) нарушается условие (37).

Список литературы

1. Faloutsos C., Faloutsos P., Faloutsos M., On power-law relationships of the internet topology. *Computer Communications Rev.* (1999) **29**, 251–262.
 2. Reittu H., Norros I., On the power-law random graph model of massive data networks. *Performance Evaluation* (2004) **55**, 3–23.
 3. Newman M. E. Y., Strogatz S. H., Watts D. J., Random graphs with arbitrary degree distribution and their applications. *Phys. Rev. E* (2001) **64**, 026118.
 4. Janson S., Luczak T., Rucinski A., *Random Graphs*. Wiley, New York, 2000.
 5. Павлов Ю. Л., Степанов М. М., Об асимптотических свойствах случайных графов “Интернет-типа”. *Обзорение прикладной и промышленной математики* (2005) **12**, №3, 677.
- 2 Дискретная математика, т.19 №3

6. Степанов М. М., О предельных распределениях степеней узлов в случайных графах “Интернет-типа”. *Методы математического моделирования и информационные технологии. Труды Института прикладных математических исследований Карельского научного центра РАН* (2005) **6**, 235–242.
7. Stepanov M. M., On the loop number of a node of a random graph of “Internet-type”. In: *Intern. workshop “Optimal stopping and stochastic control”. Extended Abst.*, Petrozavodsk, 2005, pp. 65.
8. Севастьянов Б. А., *Ветвящиеся процессы*. Наука, Москва, 1971.
9. Gnedenko B., Sur la distribution limite du terme maximum d’une série aléatoire. *Ann. Math.* (1943) **44**, №3, 423–453.
10. Дейвид Г., *Порядковые статистики*. Мир, Москва, 1979.
11. Прудников А. П., Брычков Ю. А., Маричев О. И., *Интегралы и ряды*. Наука, Москва, 1981.
12. Ибрагимов И. А., Линник Ю. В., *Независимые и стационарно связанные величины*. Наука, Москва, 1965.
13. Вагутин В. А., Распределение расстояния до корня минимального поддеревя, содержащего все вершины данной высоты. *Теория вероятностей и ее применения* (1993) **38**, №2, 273–287.
14. Ширяев А. Н., *Вероятность-2*. МЦНМО, Москва, 2004.

Статья поступила 6.04.2006.