Раздел I
Случайные графы

УДК 519.175.4

А. Р. Ярмухаметов

Кафедра дискретной математики факультета инноваций и высоких технологий МФТИ;
Кафедра математической статистики механико-математического факультета
МГУ им. М. В. Ломоносова

О некоторых свойствах случайных дистанционных графов специального вида

В настоящей работе рассматриваются случайные подграфы полного дистанционного графа, у которого вершины — векторы $\mathbf{x} \in \{0,1\}^n$ с условием $\|\mathbf{x}\| = \sqrt{n/2}$, а ребра — пары векторов, отстоящих друг от друга на расстояние $\sqrt{n/2}$. Ранее была известна пороговая вероятность для свойства связности таких случайных графов, а также пороговая вероятность для возникновения гигантской компоненты в них. Мы доказываем теперь, что, как и в классической модели Эрдеша—Реньи, фазовый переход от связности к ее отсутствию совпадает с переходом от связности к наличию изолированных вершин. Также мы формулируем результат о предельной вероятности связности в предположении, что вероятность ребра находится «внутри» фазового перехода.

Ключевые слова: случайный граф, связность, изолированная вершина, гигантская компонента, дистанционный граф, пороговая вероятность.

1. Введение

1.1. Классические результаты Эрдеша и Реньи

В 1959 году П. Эрдеш и А. Реньи предложили следующую модель случайного графа (см. [1—5]): рассматривается вероятностное пространство

$$G(N,p) = (\Omega_N, \mathscr{F}_N, \mathscr{P}_{N,p}),$$

где Ω_N — множество всех графов $G = (V, E)$ на N вершинах без петель, кратных ребер и ориентации (т. е. $|\Omega_N| = 2^{C_N^2}$), $\mathscr{F}_N = 2^{\Omega_N}$,

$$\mathscr{P}_{N,p}(G) = p^{|E|}(1-p)^{C_N^2 - |E|}, \quad p \in (0,1).$$

Иными словами, мы проводим то или иное ребро между вершинами случайного графа с вероятностью p независимо от остальных ребер.

В дальнейшем мы будем употреблять выражение «модель $G(N,p)$», подразумевая всю последовательность вероятностных пространств $G(N,p)$ при $N \in \mathbb{N}$. Отметим в этой связи, что вероятность ребра p есть, вообще говоря, функция от N.

Будем говорить, что случайный граф в модели $G(N,p)$ обладает некоторым свойством *асимптотически почти наверное* (кратко *а.п.н.*), если вероятностная мера $\mathscr{P}_{N,p}$ множества графов из Ω_N, обладающих этим свойством, стремится к 1 при $N \to \infty$.

В работах [2—3] Эрдеша и Реньи были доказаны следующие теоремы 1—5.

Настоящая работа выполнена при финансовой поддержке гранта РФФИ № 12-01-00683 и гранта Президента РФ № МД-666.2012.1.

Теорема 1. *Величина* $p^* = \dfrac{\ln N}{N}$ *является «пороговой вероятностью» для свойства связности случайного графа, то есть если существует такое* $c > 1$*, что при всех достаточно больших* N *выполнено неравенство* $p \geqslant cp^*$*, то случайный граф в модели* $G(N,p)$ *а.п.н. связен, а если существует такое* $c < 1$*, что при всех достаточно больших* N *выполнено неравенство* $p \leqslant cp^*$*, то случайный граф в модели* $G(N,p)$ *а.п.н. не связен.*

Теорема 2. *Пусть* $k \geqslant 2$*, а* $H_{N,k}$ *— число* k*-вершинных древесных компонент в случайном графе в модели* $G(N,p)$*. Тогда имеют место следующие утверждения:*

1) *если* $p = o(N^{-k/(k-1)})$*, то а.п.н. выполнено равенство* $H_{N,k} = 0$*;*

2) *если существует такое* $c > 0$*, что* $p \sim c \cdot N^{-k/(k-1)}$ *при* $N \to \infty$*, то величина* $H_{N,k}$ *имеет асимптотически пуассоновское распределение с параметром*

$$\lambda_1 = \frac{c^{k-1}k^{k-2}}{k!};$$

3) *если* $pN^{k/(k-1)} \to +\infty$ *и* $pkN - \ln N - (k-1)\ln\ln N \to -\infty$ *при* $N \to \infty$*, то для любого* $l \in \mathbb{R}$ *а.п.н. имеет место неравенство* $H_{N,k} \geqslant l$*;*

4) *если существует такое* $c \in \mathbb{R}$*, что* $pkN - \ln N - (k-1)\ln\ln N \to c$ *при* $N \to \infty$*, то величина* $H_{N,k}$ *имеет асимптотически пуассоновское распределение с параметром*

$$\lambda_2 = \frac{e^{-c}}{k \cdot k!};$$

5) *если* $pkN - \ln N - (k-1)\ln\ln N \to +\infty$ *при* $N \to \infty$*, то а.п.н. имеет место равенство* $H_{N,k} = 0$*.*

Теорема 3. *Величина* $p_1^* = 1/N$ *является «пороговой вероятностью» для возникновения «гигантской компоненты» в случайном графе, а именно: если существует такое* $c < 1$*, что при всех достаточно больших* N *выполнено неравенство* $p \leqslant c/N$*, то найдется такая константа* $\beta > 0$*, что случайный граф в модели* $G(N,p)$ *а.п.н. будет состоять из компонент, размер (количество вершин) каждой из которых не превосходит* $\beta \ln N$*; если же существует такое* $c > 1$*, что при всех достаточно больших* N *выполнено неравенство* $p \geqslant c/N$*, то найдутся такие константы* $\gamma \in (0,1)$ *и* $\beta > 0$*, что случайный граф в модели* $G(N,p)$ *а.п.н. будет содержать ровно одну компоненту размера не меньше* γN *(ее, по понятным причинам, и называют гигантской) и еще несколько компонент, размер каждой из которых не превосходит* $\beta \ln N$*.*

В дальнейшем термин «гигантская компонента» мы будем употреблять в аналогичных ситуациях: есть последовательность графов G_N на N вершинах, $N \to \infty$, и есть такая константа $\gamma \in (0,1)$, что в графе G_N (возможно, при $N \geqslant N_0$) присутствует компонента размера не меньше γN.

Теорема 4. *Пусть* p^* *— пороговая вероятность из теоремы 1. Если существует такое* $c > 1/2$*, что при всех достаточно больших* N *выполнено неравенство* $p \geqslant cp^*$*, то случайный граф в модели* $G(N,p)$ *а.п.н. состоит из гигантской компоненты и, возможно, из изолированных вершин.*

Другими словами, условие теоремы 4 (совместно с теоремой 1) можно переформулировать как совпадение моментов возникновения связности и исчезновения изолированных вершин.

Теорема 5. *Предположим, существует константа* $c \in \mathbb{R}$*, с которой* $p = (\ln N + c + o(1))/N$*. Тогда*

$$\mathscr{P}_{N,p}(G \text{ связен}) \to \exp(-\exp(-c)) \quad \text{при} \quad N \to \infty.$$

Комментарий. Если пункт 4 теоремы 2 доказать для $k \geqslant 1$, то этот пункт при $k = 1$ совместно с теоремой 4 даст нам теорему 5.

1.2. Постановка основной задачи

Введем новое вероятностное пространство, которое будем называть *пространством случайных дистанционных графов*. Для этого положим $n = 4k$, $k \in \mathbb{N}$, $N = C_n^{n/2}$ и рассмотрим *полный дистанционный граф* $\mathscr{G}_N = (\mathscr{V}_N, \mathscr{E}_N)$, у которого

$$\mathscr{V}_N = \{\mathbf{x} = (x_1, \ldots, x_n) \colon x_i \in \{0,1\}, x_1 + \ldots + x_n = 2k = n/2\},$$
$$\mathscr{E}_N = \{\{\mathbf{x}, \mathbf{y}\} \in \mathscr{V}_N \times \mathscr{V}_N \colon (\mathbf{x}, \mathbf{y}) = k = n/4\},$$

где $(\mathbf{x}, \mathbf{y})$ — евклидово скалярное произведение векторов $\mathbf{x}, \mathbf{y}$.

Таким образом, вершины полного дистанционного графа являются точками из $\{0,1\}^n$ и этих вершин ровно N. При этом ребра графа $\mathscr{G}_N$ суть пары его вершин, удаленных друг от друга на расстояние $\sqrt{n/2}$. Именно этим и обусловлено название графа. Рассмотрение подобных графов глубоко мотивировано задачами комбинаторной геометрии (см. [6, 7]).

Определим новое вероятностное пространство

$$\mathscr{G}^{dist}(N, p) = (\Omega_N^{dist}, \mathscr{F}_N^{dist}, \mathscr{P}_{N,p}^{dist}),$$

где Ω_N^{dist} — множество всех остовных подграфов $G = (\mathscr{V}_N, E)$ полного дистанционного графа $\mathscr{G}_N$, $\mathscr{F}_N^{dist} = 2^{\Omega_N^{dist}}$,

$$\mathscr{P}_{N,p}^{dist}(G) = p^{|E|}(1-p)^{|\mathscr{E}_N| - |E|}, \quad \text{где} \quad p \in (0,1).$$

Несмотря на близость модели, рассматриваемой в данной работе, и классической модели Эрдеша—Реньи случайных графов между ними имеются существенные различия.

2. Формулировки результатов

Сделаем несколько предварительных замечаний.

В соответствии с формулой Стирлинга общее число вершин в полном дистанционном графе $\mathscr{G}_N$ есть

$$N = C_n^{n/2} = \sqrt{\frac{2}{\pi}} \cdot \frac{2^n}{\sqrt{n}} \cdot (1 + o(1)), \quad n = \frac{\ln N}{\ln 2} \cdot (1 + o(1)).$$

Граф $\mathscr{G}_N$ — регулярный. Обозначим через N_1 степень каждой из его вершин. В соответствии с формулой Стирлинга

$$N_1 = (C_{n/2}^{n/4})^2 = \frac{2\sqrt{2\ln 2}}{\sqrt{\pi}} \cdot \frac{N}{\sqrt{\ln N}} \cdot (1 + o(1)).$$

Сформулируем результаты для модели $\mathscr{G}^{dist}(N, p)$, аналогичные результатам в классической модели.

Теорема 6. *Пусть* $p_* = \dfrac{\ln N}{N_1}$. *Тогда*

а) *если существует такое* $c > 1$, *что при всех достаточно больших* N *выполнено неравенство* $p \geqslant cp_*$, *то случайный граф в модели* $\mathscr{G}^{dist}(N, p)$ *а.п.н. связен;*

б) *если существует такое* $c < 1$, *что при всех достаточно больших* N *выполнено неравенство* $p \leqslant cp_*$, *то случайный граф в модели* $\mathscr{G}^{dist}(N, p)$ *а.п.н. не связен.*

Теорема 6 доказана в [8].

Теорема 7. *Пусть $k \geqslant 2$, а $X_{N,k}$ — число k-вершинных древесных компонент в случайном графе в модели $\mathscr{G}^{dist}(N,p)$. Тогда имеют место следующие утверждения:*

1) если $p = o(N^{-1/(k-1)} \cdot N_1^{-1})$, то а.п.н. выполнено равенство $X_{N,k} = 0$;

2) если существует такое $c > 0$, что $p \sim c \cdot N^{-\frac{1}{k-1}} \cdot N_1^{-1}$ при $N \to \infty$, то величина $X_{N,k}$ имеет асимптотически пуассоновское распределение с параметром

$$\lambda_1 = \frac{c^{k-1}k^{k-2}}{k!};$$

3) если $p \cdot N^{\frac{1}{k-1}} \cdot N_1 \to +\infty$ и $pkN_1 - \ln N - (k-1)\ln\ln N \to -\infty$ при $N \to \infty$, то для любого $l \in \mathbb{R}$ а.п.н. имеет место неравенство $X_{N,k} \geqslant l$;

4) если существует такое $c \in \mathbb{R}$, что $pkN_1 - \ln N - (k-1)\ln\ln N \to c$ при $N \to \infty$, то величина $X_{N,k}$ имеет асимптотически пуассоновское распределение с параметром

$$\lambda_2 = \frac{e^{-c}}{k \cdot k!};$$

5) если $pkN_1 - \ln N - (k-1)\ln\ln N \to +\infty$ при $N \to \infty$, то а.п.н. имеет место равенство $X_{N,k} = 0$.

Теорема 7 доказана в [9].

Теорема 8. *Пусть $p_* = \dfrac{1}{N_1}$. Тогда*

а) если существует такое $c > 1$, что при всех достаточно больших N выполнено равенство $p = cp_*$, то найдутся такие функции θ_1 и θ_2, равные $o(1)$ при $N \to \infty$, что случайный граф в модели $\mathscr{G}^{dist}(N,p)$ а.п.н. будет содержать ровно одну гигантскую компоненту, имеющую размер

$$N(1-t_c)(1+\theta_1(N)), \quad t_c = \frac{1}{c}\sum_{k=1}^{\infty}\frac{k^{k-1}}{k!}(ce^{-c})^k,$$

и еще несколько компонент, размер каждой из которых не превосходит $N \cdot \theta_2(N)$;

б) если существует такое $c < 1$, что при всех достаточно больших N выполнено неравенство $p \leqslant cp_*$, то найдется такая константа $\beta > 0$, что случайный граф в модели $\mathscr{G}^{dist}(N,p)$ а.п.н. будет состоять из компонент, размер каждой из которых не превышает $\beta \ln N$.

Теорема 8 доказана в [10].

Теорема 9. *Пусть вероятность p_* такая же, как в теореме 6. Если существует такое $c > 2/3$, что при всех достаточно больших N выполнено неравенство $p \geqslant cp_*$, то случайный граф в модели $\mathscr{G}^{dist}(N,p)$ а.п.н. состоит из гигантской компоненты и, возможно, из изолированных вершин.*

Теорема 9 доказана в разделе 3.

Теорема 10. *Предположим, существует такая константа $c \in \mathbb{R}$, что $p = (\ln N + c + o(1))/N_1$. Тогда*

$$\mathscr{P}_{N,p}^{dist}(G \text{ связен}) \to \exp(-\exp(-c)) \quad \text{при} \quad N \to \infty.$$

Теорема 10 следует из пункта 4 теоремы 7 для случая $k = 1$ и теоремы 9.

Замечание. Несмотря на то, что пункт 4 теоремы 7 сформулирован для $k \geqslant 2$, нетрудно видеть, что ее доказательство (см. [9]) остается верным и для случая $k = 1$.

3. Доказательство теоремы 9

В соответствии с теоремой 6 можно считать, что $p \leqslant \frac{2\ln N}{N_1}$. При этом, конечно, N достаточно велико и $p \geqslant cp_*$, где $c > 2/3$.

Приведем некоторые вспомогательные обозначения из [8]: $X_{N,i}$ — число i-элементных компонент связности случайного графа в пространстве $\mathcal{G}^{dist}(N,p)$, $f(i, N-i)$ — минимальное число ребер между подмножеством множества вершин графа $\mathcal{G}_N$ из i элементов и его дополнением из $N-i$ элементов, где минимум берется по всем i-элементным подмножествам множества вершин графа $\mathcal{G}_N$.

Положим также

$$Y_N = \sum_{i=2}^{[N/2]+1} X_{N,i}.$$

Поскольку из теоремы 8 следует существование гигантской компоненты, то для завершения доказательства теоремы достаточно показать, что $EY_N \to 0$ при $N \to \infty$.

Зафиксируем $\alpha \in (0, 1/10)$ и $\beta \geqslant 10^6$. Тогда найдется такое N_0, что при всех $N \geqslant N_0$ верна следующая последовательность неравенств:

$$EY_N = \sum_{i=2}^{[N/2]+1} EX_{N,i} \leqslant \sum_{i<\alpha N_1} EX_{N,i} + \sum_{\alpha N_1 \leqslant i < \beta N_1} C_N^i (1-p)^{f(i,N-i)} + \\ + \sum_{\beta N_1 \leqslant i \leqslant [N/2]+1} C_N^i (1-p)^{f(i,N-i)} = \sum\nolimits_1^* + \sum\nolimits_2^* + \sum\nolimits_3^*.$$

Утверждение 1. *В условиях теоремы 9 и при фиксированном выше α имеем $\sum_1^* \to 0$, коль скоро $N \to \infty$.*

Доказательство. Воспользуемся частью 2 утверждения 1 из [10]. Как нетрудно видеть, доказательство этой части утверждения остается прежним, если заменить условие $2 \leqslant k \leqslant N_1^{3/4}$ на условие $2 \leqslant k \leqslant N$.

Таким образом, при $2 \leqslant k \leqslant N$ имеет место неравенство

$$T_{k,N} \leqslant N \cdot N_1^{k-1} \cdot \frac{k^{k-2}}{k!},$$

где $T_{k,N}$ — число различных k-вершинных деревьев (не обязательно индуцированных) в полном дистанционном графе $\mathcal{G}_N$.

Поскольку в любом связном графе можно так убрать несколько ребер, чтобы вершины вместе с оставшимися ребрами образовывали дерево, то, пользуясь вышесказанным, имеем

$$\sum\nolimits_1^* \leqslant \sum_{k=2}^{[\alpha N_1]} T_{k,N} p^{k-1} (1-p)^{kN_1 - 2C_k^2} \leqslant \sum_{k=2}^{[\alpha N_1]} N N_1^{k-1} \frac{k^{k-2}}{k!} p^{k-1} (1-p)^{kN_1 - 2C_k^2} \leqslant$$

$$(\text{т.к. } k \leqslant [\alpha N_1], \text{ то } 2C_k^2 \leqslant k^2 \leqslant k\alpha N_1)$$

$$\leqslant \sum_{k=2}^{[\alpha N_1]} N \frac{k^{k-2}}{k!} (2\ln N)^{k-1} \left(1 - \frac{c\ln N}{N_1}\right)^{(1-\alpha)kN_1} \leqslant$$

$$\leqslant \sum_{k=2}^{[\alpha N_1]} N \frac{k^{k-2}}{k!}(2\ln N)^{k-1}\exp(-c(1-\alpha)k\ln N) = \sum_{k=2}^{[\alpha N_1]} \frac{k^{k-2}}{k!}\frac{(2\ln N)^{k-1}}{N^{c(1-\alpha)k-1}} =$$

(при $k \geqslant 2$ имеем $c(1-\alpha)k - 1 > 2/3 \cdot 9/10 \cdot 2 - 1 = 1/5 > 0$)

$$= \delta(N) + \sum_{k=10}^{[\alpha N_1]} \frac{k^{k-2}}{k!}\frac{(2\ln N)^{k-1}}{N^{c(1-\alpha)k-1}},$$

где $\delta(N) = o(1)$.

Воспользовавшись неравенством $k! \geqslant (k/e)^k$ и неравенством

$$c(1-\alpha)k - 1 \geqslant k/2 \text{ при } k \geqslant 10,$$

имеем

$$\sum\nolimits_1^* \leqslant \delta(N) + \sum_{k=10}^{[\alpha N_1]} \frac{k^{k-2}}{k^k \cdot e^{-k}}\frac{(2\ln N)^{k-1}}{\sqrt{N}^k} \leqslant \delta(N) + \sum_{k=10}^{[\alpha N_1]} \left(\frac{2e\ln N}{\sqrt{N}}\right)^k \to 0 \quad \text{при} \quad N \to \infty.$$

Таким образом, утверждение доказано. □

Утверждение 2. *В условиях теоремы 9 и при фиксированных выше α, β имеем $\sum_2^* \to 0$, коль скоро $N \to \infty$.*

Утверждение 2 непосредственно следует из доказательства утверждения 5 из [8]. Как нетрудно видеть, утверждение 5 из [8] остается верным при $c > 2/3$ (и даже при $c > 1/2$), а также $\sum_2^* < \sum_2$, где $\sum_2$ из [8].

Утверждение 3. *В условиях теоремы 9 и при фиксированном выше β имеем $\sum_3^* \to 0$, коль скоро $N \to \infty$.*

Утверждение 3 непосредственно следует из доказательства утверждения 6 из [8] небольшим его изменением. Как нетрудно видеть, $\sum_3^* < \sum_3$, где $\sum_3$ из [8].

Таким образом, из утверждений 1, 2 и 3 следует, что $EY_N \to 0$ при $N \to \infty$. Теорема 9 доказана.

Литература

1. *Алон Н.*, *Спенсер Дж.* Вероятностный метод.— М.: Бином, 2007.
2. *Erdős P.*, *Rényi A.* On random graphs 1 // Publ. Math.— 1959.— Debrecen 6.— P. 290—297.
3. *Erdős P.*, *Rényi A.* On the evolution of random graphs // Publ. Math. Inst. Hungar. Acad. Sci.— 1960.— V 5.— P. 17—61.
4. *Erdős P.*, *Rényi A.*. On the strength of connectedness of a random graph // Acta Math. Acad. Sci. Hungar.— 1961.— V. 12.— P. 261—267.
5. *Bollobás B.*. Random Graphs.— New York: Academic Press., 2001.
6. *Райгородский А. М.* Линейно-алгебраический метод в комбинаторике.— М.: МЦНМО, 2007.
7. *Райгородский А. М.* Проблема Борсука и хроматические числа некоторых метрических пространств // УМН. — 2001.— Т. 56, вып. 1 (337).— С. 107—146.

8. *Ярмухаметов А. Р.* О связности случайных дистанционных графов специального вида // Чебышевский сборник.— 2009.— Т. X, вып. 1 (29).— С. 95—108.

9. *Ярмухаметов А. Р.* Древесные компоненты в случайных дистанционных графах спецаильного вида // Современная математика и ее приложения. — 2011. — Т. XX. — С. 98—110.

10. *Ярмухаметов А. Р.* Гигантская компонента в случайных дистанционных графах специального вида // Математические заметки, в печати.

Поступила в редакцию 15.06.2011