Московский государственный университет имени М.В. Ломоносова
Механико-математический факультет

На правах рукописи
УДК 519.21

Ярмухаметов Андрей Ринатович

Предельные распределения характеристик случайных дистанционных графов

01.01.05 — теория вероятностей и математическая статистика

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата физико-математических наук

Москва — 2012

Работа выполнена на кафедре математической статистики и случайных процессов механико-математического факультета Московского государственного университета имени М.В. Ломоносова.

Научный руководитель: доктор физико-математических наук
профессор Райгородский Андрей Михайлович,

Официальные оппоненты: доктор физико-математических наук
Кабатянский Григорий Анатольевич,
Институт проблем передачи информации
имени А.А. Харкевича РАН,
главный научный сотрудник

кандидат физико-математических наук
Шабанов Дмитрий Александрович,
Московский государственный университет
имени М.В.Ломоносова,
механико-математический факультет,
ассистент кафедры теории вероятностей

Ведущая организация: Хабаровское отделение
института прикладной математики
Дальневосточного отделения РАН

Защита диссертации состоится 7 декабря 2012 года в 16 часов 45 минут на заседании диссертационного совета Д 501.001.85 при Московском государственном университете имени М.В. Ломоносова по адресу: 119991, ГСП-1, Москва, Ленинские горы, МГУ, механико-математического факультет, аудитория 16-24.

С диссертацией можно ознакомиться в библиотеке механико-математического факультета (Главное здание, 14 этаж).

Автореферат разослан 6 ноября 2012 г.

Ученый секретарь диссертационного
совета Д 501.001.85 при МГУ,
доктор физико-математических наук,
профессор В.Н. Сорокин

Актуальность

Открытие того, что детерминированные утверждения могут быть доказаны с помощью вероятностных соображений, позволило уже в первой половине XX в. установить ряд замечательных фактов из анализа, теории чисел, комбинаторики и теории информации. Вскоре стало ясно, что метод, который сейчас называется *вероятностным*, является весьма мощным инструментом получения результатов в дискретной математике.

Активное изучение случайных графов началось после того, как П. Эрдеш установил, что вероятностный метод помогает решать многие задачи экстремальной теории графов. П. Эрдеш и А. Реньи в 1959 году [1,2] рассмотрели модель случайного графа $G(N,p)$, в которой ребра в полном графе на N вершинах проводятся независимо друг от друга с вероятностью $p = p(N)$. Огромное количество работ посвящено различным задачам, связанным с исследованиями случайного графа $G(N,p)$. Среди них отметим работы Б. Боллобаша, С. Шелла, Е. Семереди, Дж. Спенсера, З. Палка, А.Д. Барбура, Е.Н. Гильберта, И.Н. Коваленко, Г.И. Ивченко, В.Ф. Колчина, В.Е. Степанова и др.

Диссертация посвящена изучению экстремальных характеристик случайных подграфов полного дистанционного графа $\mathcal{G}_N$, у которого вершины — векторы из $\{0,1\}^n$ с условием $||x|| = \sqrt{n/2}$, а ребра — пары векторов, отстоящих друг от друга на расстояние $\sqrt{n/2}$. В дальнейшем вероятностное пространство таких случайных подграфов называется пространством случайных дистанционных графов и обозначается $\mathcal{G}^{dist}(N,p)$, где N — количество вершин в полном графе, p — вероятность появления ребра в случайном подграфе этого графа. Таким образом, модель, рассматриваемая в работе, является обобщением классической модели Эрдеша – Реньи.

Рассмотрение дистанционных графов мотивировано классической задачей комбинаторной геометрии о хроматическом числе пространства[3].

[1]Erdős P., Rényi A., *On random graphs I*, Publ. Math. Debrecen, 6 (1959), 290 – 297.

[2]Erdős P., Rényi A., *On the evolution of random graphs*, Magyar Tud. Akad. Mat. Kutató Int. Közl., 5 (1960), 17 – 61.

[3]Райгородский А.М., *Проблема Борсука и хроматические числа метрических пространств*,

Впервые полный дистанционный граф $\mathcal{G}_N$, свойства которого изучаются в диссертационной работе, в геометрическом контексте рассмотрели в 1981 году П. Франкл и Р.М. Уилсон[4]. С помощью этого графа они показали, что хроматическое число пространства $\mathbb{R}^n$ растет экспоненциально с ростом n. В 1991 году Дж. Кан и Г. Калаи[5] применили результаты Франкла и Уилсона для опровержения классической гипотезы Борсука о том, что всякое ограниченное неодноточечное множество в $\mathbb{R}^n$ может быть разбито на $n+1$ часть меньшего диаметра. Таким образом, изучение внутренней структуры дистанционного графа и его подграфов играет исключительно важную роль.

Также этот граф имеет непосредственное отношение к теории кодирования, в частности максимальные клики в нем это матрицы Адамара[6].

В диссертации получена пороговая вероятность для свойства связности случайных графов в вероятностном пространстве $\mathcal{G}^{dist}(N,p)$, а также пороговая вероятность для возникновения гигантской компоненты в них. В работе также изучено предельное распределение числа древесных компонент в случайном дистанционном графе в зависимости от вероятности ребра p. Кроме того, доказано, 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 получен результат о предельной вероятности связности в предположении, что вероятность ребра находится "внутри" этого фазового перехода.

Цель работы

Цель данной диссертации состоит в нахождении пороговых вероятностей связности и наличия "гигантской компоненты" случайного графа в модели $\mathcal{G}^{dist}(N,p)$, исследовании предельного распределения числа древесных компонент в случайном графе в этой же модели в зависимости от вероятности ребра p.

Успехи Матем. Наук, 56(1) (2001), 107 – 146.

[4]Frankl P., Wilson R., *Intersection theorems with geometric consequences*, Combinatorica, 1 (1981), 357 – 368.

[5]Kahn J., Kalai G., *A counterexample to Borsuk's conjecture*, Bulletin (new series) of the AMS, 29(1) (1993), 60 – 62.

[6]Холл М., *Комбинаторика*, М.: Мир, 1970.

Научная новизна

Все результаты диссертации являются новыми. Перечислим основные из них:

1. Доказано, что пороговая вероятность связности случайного графа в модели $\mathcal{G}^{dist}(N,p)$ равна $\ln N/N_1$, где N_1 — степень вершины полного дистанционного графа (который является регулярным).

2. Найдено предельное распределение количества древесных компонент в случайном графе в модели $\mathcal{G}^{dist}(N,p)$ при различных вероятностях ребра p.

3. Доказано, что пороговая вероятность наличия гигантской компоненты в случайном графе в модели $\mathcal{G}^{dist}(N,p)$ равна $1/N_1$.

Результаты диссертации обоснованы в виде строгих математических доказательств и получены автором самостоятельно. Точные формулировки установленных автором утверждений приведены ниже.

Методы исследования

При доказательстве результатов автором использовалась разнообразная техника. Так при решении задачи о гигантской компоненте и распределении числа деревьев потребовалась разработка новой методики подсчета деревьев в полном дистанционном графе. Используются методы теории вероятностей и случайных процессов: метод моментов, теория ветвящихся процессов и др. Все теоремы в той или иной степени потребовали применения сложной комбинаторной техники в сочетании с теорией вероятностей.

Теоретическая и практическая ценность

Диссертация носит теоретический характер. Результаты и методы работы могут быть полезными при изучении вероятностного метода в комбинаторике, в исследованиях свойств случайных дистанционных графов, а также полного дистанционного графа.

Апробация работы

По теме диссертации были сделаны доклады на следующих семинарах механико-математического факультета МГУ имени М.В. Ломоносова:

- Кафедральном семинаре кафедры математической статистики и случайных процессов под рук. профессора А.М. Зубкова (2009 и 2011 гг.),
- Семинаре под руководством профессора А.М. Райгородского (2008 – 2011 гг., неоднократно)

Результаты диссертации докладывались на десятом международном семинаре "Дискретная математика и ее приложения" (г. Москва, 2010 г.), на международной конференции "8-я французская конференция по комбинаторике" (г. Париж, Франция, 2010 г.), на международной конференции "Конечные и бесконечные множества", посвященной 80-летию Хайнала (г. Будапешт, Венгрия, 2011).

Работа автора поддержана грантами РФФИ 09-01-00294 и 12-01-00683.

Публикации

Результаты диссертации опубликованы в 6 работах автора, из них 4 — из списка ВАК. Работ в соавторстве нет. Список работ приведен в конце автореферата.

Структура диссертации

Диссертация состоит из введения, четырех глав и списка литературы, насчитывающего 73 наименования. Общий объем диссертации составляет 63 страницы.

КРАТКОЕ СОДЕРЖАНИЕ ДИССЕРТАЦИИ

Во введении излагается история исследований, относящихся к случайным графам Эрдеша и Реньи, случайным дистанционным графам, полному дистанционному графу, а также описывается структура диссертации.

В главе 1 сформулированы основные определения и результаты диссертации. Для точных формулировок этих результатов введем ряд обозначений. Пусть $N \in \mathbb{N}$. Рассмотрим множество Ω_N всех графов $G = (V_N, E)$ без петель, кратных ребер и ориентации с множеством вершин $V_N = \{1, ..., N\}$. Назовем *случайным графом в модели Эрдеша–Реньи* вероятностное пространство

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

где $\mathcal{F}_N = 2^{\Omega_N}$, $\mathcal{P}_{N,p}(G) = p^{|E|}(1-p)^{C_N^2 - |E|}$, $p \in (0,1)$.

Положим $n = 4k$, $k \in \mathbb{N}$, $N = C_n^{n/2}$ и рассмотрим *полный дистанционный граф* $\mathcal{G}_N = (\mathcal{V}_N, \mathcal{E}_N)$, у которого

$$\mathcal{V}_N = \left\{ \mathbf{x} = (x_1, \ldots, x_n) : x_i \in \{0,1\}, x_1 + \ldots + x_n = 2k = \frac{n}{2} \right\},$$

$$\mathcal{E}_N = \left\{ \{\mathbf{x}, \mathbf{y}\} \in \mathcal{V}_N \times \mathcal{V}_N : \langle \mathbf{x}, \mathbf{y} \rangle = k = \frac{n}{4} \right\},$$

где $\langle \mathbf{x}, \mathbf{y} \rangle$ — скалярное произведение векторов $\mathbf{x}$ и $\mathbf{y}$.

Таким образом, вершины полного дистанционного графа являются точками из $\{0,1\}^n$ и этих вершин ровно N. При этом ребра графа $\mathcal{G}_N$ суть пары его вершин, удаленных друг от друга на расстояние $\sqrt{\frac{n}{2}}$. Именно этим и обусловлено название графа. Изучение такого рода графов мотивировано не только задачей о хроматическом числе пространства, но и исследованиями равновесных кодов с запрещенным расстоянием[7].

Определим новое вероятностное пространство

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

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

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

[7] Мак-Вильямс Ф.Дж., Слоэн Н.Дж.А., Теория кодов, исправляющих ошибки, Москва, «Связь» 1979.

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

$$N_1 = \left(C_{\frac{n}{2}}^{\frac{n}{4}}\right)^2 = \frac{2\sqrt{2\ln 2}}{\sqrt{\pi}} \cdot \frac{N}{\sqrt{\ln N}} \cdot (1 + \delta_1(N)),$$

где $\delta_1(N) = o(1)$.

Имеется некоторая литература о моделях случайных подграфов регулярных графов (у нас как раз частный случай такой модели)[8]. Однако наш граф достаточно сложный для анализа его свойств с помощью этих результатов напрямую.

Далее, как в диссертации, так и в данном автореферате, там, где это не создает путаницы, под понятием *случайный граф* подразумевается как само вероятностное пространство, так и элемент этого пространства — граф G.

В дальнейшем будем употреблять выражение "модель $\mathcal{G}^{dist}(N, p)$", подразумевая всю последовательность вероятностных пространств $\mathcal{G}^{dist}(N, p)$ при $N \in \mathbb{N}$.

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

Глава 2 посвящена теореме о связности случайного дистанционного графа. Приведем ее формулировку.

Теорема 1. *Пусть* $p_* = \frac{\ln N}{N_1}$. *Тогда*

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

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

[8] Chung F., Horn P., Lu L., *The giant component in a random subgraph of a given graph*, Proceedings of WAW2009, Lecture Notes in Computer Science 5427, 38 – 49.

Теорема 1 по сути аналогична теореме о связности случайного графа в классической модели Эрдеша – Реньи. Однако, ее доказательство требует существенно более продвинутой техники. В полном графе количество ребер между любым k-элементным подмножеством вершин и его дополнением равняется $k(N-k)$. Поэтому для случайного графа в классической модели можно явно оценить сверху математическое ожидание количества компонент связности следующим образом:

$$\sum_{k=1}^{N} C_N^k (1-p)^{k(N-k)}.$$

Для случайного дистанционного графа такого явного равенства нет, так как оценка количества ребер между подмножеством вершин и его дополнением в случае полного дистанционного графа является существенно более нетривиальной задачей.

При доказательстве пункта б) установлено, что при данных условиях количество изолированных вершин в случайном графе в модели $\mathcal{G}^{dist}(N,p)$ стремится к бесконечности.

При доказательстве пункта а), так же как и в классическом случае устанавливается, что математическое ожидание количества компонент связности размера, меньшего чем N, стремится к нулю. Среди всех промежуточных результатов выделим следующее важное свойство полного дистанционного графа.

Утверждение 2.2. *Число общих соседних вершин для двух любых различных вершин графа $\mathcal{G}_N$ при достаточно больших N не меньше, чем*

$$N_2 = 10^{-4}\frac{N}{\ln N}.$$

С помощью этого свойства и регулярности доказывается, что количество ребер между любым i-элементным подмножеством вершин графа $\mathcal{G}_N$ и его дополнением не меньше, чем $\min(i, N-i)\frac{N_2}{3}$.

Глава 3 посвящена теореме о предельном распределении количества древесных компонент в случайном графе в модели $\mathcal{G}^{dist}(N,p)$. Приведем формулировку теоремы.

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

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

2. *если* $p \sim c \cdot N^{-\frac{1}{k-1}} \cdot N_1^{-1}$ *для некоторого* $c > 0$, *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 фиксированного* $l \in \mathbb{R}$ *а.п.н. выполняется неравенство* $X_{N,k} \geq l$;

4. *если* $pkN_1 - \ln N - (k-1)\ln\ln N \to c$ *при* $N \to \infty$ *для некоторого* $c \in \mathbb{R}$, *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* $X_{N,k} = 0$.

Результат теоремы очень похож на результат для классической модели Эрдеша – Реньи, хотя и с некоторыми отличиями. Распределение количества древесных компонент в классическом случае можно подсчитать, используя формулу Кэли, утверждающую, что число различных деревьев на k нумерованных вершинах равняется k^{k-2}. Откуда следует, что количество различных k-вершинных деревьев в полном графе размера N равняется $C_N^k k^{k-2}$. В диссертации доказано утверждение о количестве деревьев в графе $\mathcal{G}_N$.

Утверждение 3.2. *Обозначим через $T_{k,N}$ число различных k-вершинных деревьев в полном дистанционном графе $\mathcal{G}_N$ (т.е. деревьев, которые можно получить из $\mathcal{G}_N$ отбрасыванием некоторых его вершин и ребер).*

1. *Пусть $2 \leq k \leq N_1^{1/4}$. Тогда при достаточно больших N имеет место равенство*

$$T_{k,N} = N \cdot N_1^{k-1} \cdot \frac{k^{k-2}}{k!}(1 + \delta^{'}(N)),$$

где $|\delta^{'}(N)| \leq \frac{2}{\sqrt{N_1}}$.

2. *Пусть $2 \leq k \leq N$. Тогда имеет место неравенство*

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

Это утверждение применяется также в теореме о наличии гигантской компоненты в случайном дистанционном графе, содержащейся в главе 4.

Теорема 3. *Пусть $p_* = \frac{1}{N_1}$. Тогда:*

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

б) если существует такое $c > 1$, что при всех достаточно больших N выполнено равенство $p = cp_*$, то найдется такая константа $\beta > 0$, что случайный граф в модели $\mathcal{G}^{dist}(N,p)$ а.п.н. будет содержать ровно одну компоненту размера

$$N(1 - t_c) + O(N^{0,8}),$$

где $t_c = \frac{1}{c}\sum_{k=1}^{\infty} \frac{k^{k-1}}{k!}(ce^{-c})^k$, и при этом все остальные компоненты будут иметь размер, не превосходящий $\beta \ln N$.

Пункт а) теоремы доказывается с помощью теории ветвящихся процессов. Для каждой вершины A графа строится ветвящийся процесс, с помощью которого доказывается, что количество вершин в компоненте связности, содержащей A, превосходит $\beta \ln N$ с вероятностью, меньшей чем $o(N^{-1})$, при достаточно больших N и некотором фиксированном $\beta = \beta(c)$. Так как вероятность объединения меньше, чем сумма вероятностей, то отсюда непосредственно вытекает утверждение пункта а).

Пункт б) является существенно более сложным. Сначала доказывается, что существует $\beta = \beta(c)$, такое, что количество вершин, содержащихся в компонентах связности, размер которых не превосходит $\beta \ln N$, равняется $Nt_c + O(N^{0,8})$. Стоит отметить, что большинство из этих компонент являются деревьями, количество вершин, содержащихся в недревесных компонентах связности, равняется $o(\ln N)$. Также доказывается, что с вероятностью 1 нет компонент связности в интервале $\left(\beta \ln N, \left[N_1^{0,9}\right]\right)$. В завершении мы доказываем, что все оставшиеся вершины а.п.н. образуют связный подграф в случайном графе.

После доказательства теоремы о гигантской компоненте в главе 4 устанавливаются еще 2 теоремы. Приведем их формулировки.

Теорема 4. *При $p \geq cp_*$ (где $c > 2/3$, p_* из теоремы 1) случайный граф в пространстве $\mathcal{G}^{dist}(N,p)$ а.п.н. состоит из гигантской компоненты и, возможно, из изолированных вершин.*

Иными словами, теорема 4 говорит о том, что переход от связности графа к ее отсутствию совпадает с переходом от связности к наличию изолированных вершин.

Теорема 5. *Пусть $p = (\ln N + c + \delta_3(N))/N_1$ (где $c \in \mathbb{R}$, $\delta_3(N) = o(1)$). Тогда*

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

Благодарности. Автор выражает глубокую признательность профессору Андрею Михайловичу Райгородскому за постановку задач, постоянную внимательность к работе, многолетнюю поддержку.

Работы автора по теме диссертации

[1] Yarmukhametov A.R., *Tree components in random distance graphs of special form*, Journal of Mathematical Sciences, 3 (187) (2012), 360 – 373.

[2] Ярмухаметов А.Р., *Гигантская компонента в случайных дистанционных графах специального вида*, Математические заметки, 3 (92) (2012), 463 – 479.

[3] Ярмухаметов А.Р., *О некоторых свойствах случайных дистанционных графов специального вида*, Труды МФТИ, 4 №1 (2012), 4 – 10.

[4] Ярмухаметов А.Р., *Гигантская и мелкие компоненты в случайных дистанционных графах специального вида*, Математические заметки, 6 (92) (2012), 949 – 953.

[5] Ярмухаметов А.Р., *О связности случайных дистанционных графов специального вида*, Чебышевский сборник, X, выпуск 1(29) (2009), 95 – 108.

[6] Ярмухаметов А.Р., *Древесные компоненты в случайных дистанционных графах специального вида*, Современная математика и ее приложения, 20 (2011), 98 – 110.