УДК 519.175.4

МАКСИМАЛЬНЫЕ ИНДУЦИРОВАННЫЕ ПОДГРАФЫ БИНОМИАЛЬНОГО СЛУЧАЙНОГО ГРАФА

М.Е. Жуковский

Московский физико-технический институт (государственный университет), лаборатория продвинутой комбинаторики и сетевых приложений

Россия, 141701, Московская область, г. Долгопрудный, Институтский пер., 9

E-mail: zhukmax@gmail.com

Ключевые слова: биномиальный случайный граф, индуцированные деревья, максимальный размер индуцированного подграфа.

Аннотация: Как известно, наибольший размер клики в биномиальном случайном графе принимает одно из двух (неслучайных) значений с вероятностью, стремящейся к 1. Возникает следующий общий вопрос. Пусть дана последовательность $\mathcal{F}_k$ семейств графов на k вершинах. Каково наибольшее k такое, что в случайном графе найдется индуцированный подграф, изоморфный некоторому графу из семейства $\mathcal{F}_k$, с вероятностью, стремящейся к 1? Для некоторых семейств графов (путей, простых циклов и некоторых других) известно, что, как и в случае клики, искомая величина сконцентрирована в двух точках. Во-первых, оставался открытым вопрос, справедлива ли настолько сильная концентрация для деревьев? Во-вторых, до настоящего момента не было известно ни одной “естественной” последовательности семейств, для которой концентрация в конечном множества не верна. Нам удалось как ответить на первый вопрос, так и построить требуемую во втором вопросе последовательность семейств.

1. Постановка задачи

Биномиальный случайный граф (который часто называют графом Эрдеша–Реньи) — это случайный элемент $G(n,p)$, принимающий значения во множестве всех неориентированных графов без петель и кратных ребер на множестве вершин $\{1,\ldots,n\}$, с распределением $\mathsf{P}(G(n,p)=H)=p^{e(H)}(1-p)^{C_n^2-e(H)}$, где $e(H)$ — количество ребер графа H, заданного на том же множестве вершин $\{1,\ldots,n\}$. Иными словами, все ребра случайного графа проводятся независимо, с вероятностью p каждое. Об основных характеристиках биномиального случайного графа можно почитать, например, в [1,2].

Рассмотрим последовательность $\mathcal{F}_k$ семейств графов на k вершинах (т.е. для каждого $k\in\mathbb{N}$ задано некоторое семейство $\mathcal{F}_k=\{G,\ |V(G)|=k\}$). Обозначим X_n случайную величину, равную наибольшему k, для которого существует граф $F\in\mathcal{F}_k$ и изоморфный ему индуцированный подграф H в $G(n,p)$.

Как правило, возникает следующий вопрос об асимптотическом распределении

X_n. Существует ли последовательность множеств Σ_n, $n \in \mathbb{N}$, и положительное число C такие, что, во-первых, для всех $n \in \mathbb{N}$ выполнено $|\Sigma_n| < C$, а, во-вторых, с вероятностью, стремящейся к 1 при $n \to \infty$, $X_n \in \Sigma_n$?

В следующем разделе мы поговорим о том, в каких случаях ответ на этот вопрос известен.

2. История задачи

Первый ответ на этот вопрос получен для семейства пустых (или, наоборот, полных графов).

Иными словами, речь идет о числе независимости (наибольшем размере независимого множества, т.е. множества попарно не смежных вершин) и кликовом числе (наибольшем количестве вершин в клике, т.е. в подграф, все пары вершин которого смежны) графа $G(n,p)$ [3–5]. Утверждается, что для любой константы $p \in (0,1)$ (т.е. p не зависит от n) найдется такая функция $f(n)$, что с вероятностью, стремящейся к 1, кликовое число случайного графа $G(n,p)$ принадлежит множеству $\{f(n), f(n)+1\}$ (ниже в таких ситуациях мы будем говорить о *концентрации в двух точках*). Из соображений симметрии то же самое верно и для числа независимости (но, разумеется, с некоторой другой функцией $f(n)$).

Разумеется, этот результат имеет прямое отношение к поставленному в предыдущем разделе вопросу. Действительно, X_n является числом независимости (кликовым числом), если каждое множество $\mathcal{F}_k$ состоит из единственного графа, являющегося пустым (полным).

Зададимся вопросом, а что происходит с другими одноэлементными семействами? Иными словами, путь для каждого $k \in \mathbb{N}$ $\mathcal{F}_k = \{F_k\}$. В [6] концентрация в двух точках доказана для $F_k = P_k$ (простого пути на k вершинах) и $F_k = C_k$ (простого цикла на k вершинах). В той же работе был сформулирован вопрос (и оставлен без ответа) о концентрации величины X_n в случае семейств деревьев (разумеется, в этом случае семейство $\mathcal{F}_k$ состоит из нескольких графов). В следующем разделе мы даем ответ на этот вопрос.

Обратимся теперь к большим семействам $\mathcal{F}_k$. В различных работах были рассмотрены семейства деревьев, регулярных графов, полных двудольных графов и полных многодольных графов. К сожалению, насколько нам известно, ни для какого из этих семейств не была доказана ни концентрация в двух точках, ни концентрация в m точках ни для какого натурального m. Приведем основные известные результаты.

В 1983 году в работе [7] были рассмотрены семейства деревьев (то есть $\mathcal{F}_k$ состоит из всех деревьев на k вершинах) и доказано, что $\frac{X_n}{\ln n} \xrightarrow{\mathsf{P}} \frac{2}{\ln[1/(1-p)]}$ при $n \to \infty$ (здесь и далее мы обозначаем $\xrightarrow{\mathsf{P}}$ сходимость по вероятности). В 1987 году в работе [8] были получены похожие результаты для относительно широкого класса семейств $\mathcal{F}_k$. Частными случаями, например, являются семейства $ck(1+o(1))$-регулярных графов (граф называется *d-регулярным*, если все его вершины имеют одинаковую степень, равную d) — для них $\frac{X_n}{\ln n} \xrightarrow{\mathsf{P}} \frac{2}{c\ln[1/p]+(1-c)\ln[1/(1-p)]}$ при $n \to \infty$. Для некоторых семейств полных двудольных графов и полных многодольных графов подобные результаты были получены в [8, 9].

Кроме того, в [10] были рассмотрены семейства графов, имеющих ограниченное сверху количество ребер. Более формально, для заданной последовательности $e = e(k)$ пусть $\mathcal{F}_k = \mathcal{F}_k(e)$ — это множество всех графов на k вершинах с не более $e(k)$ ребрами. Основной результат работы [10] утверждает, в частности, следующее. Пусть $n^{-1/3+\varepsilon} < p < 1-\varepsilon$ для некоторого $\varepsilon \in (0, 1/3)$. Пусть, кроме того, $e = e(k) \geqslant 0$ — такая последовательность, что $e = o\left(\frac{pk\ln k}{\ln\ln k}\right)$. Тогда для X_n имеет место концентрация в двух точках.

Несложно показать, используя, так называемый, метод второго момента, что аналогичный результат верен и для семейств графов с *в точности* e ребрами: если $0 \leqslant e(k) = O(k)$ (эта оценка легко может быть улучшена, но в настоящей работе мы не ставим перед собой цели оптимизации этого результата), а $\mathcal{F}_k = \mathcal{F}_k(e)$ — семейство всех графов на k вершинах с ровно $e(k)$ ребрами, то, опять же, справедлив тот же концентрационный результат. Для удобства мы будем обозначать случайную величину X_n, определяемую последним рассмотренным семейством, $\mathcal{X}_n[e]$.

Одна из основных целей нашего исследования состоит в нахождении такой естественной последовательности графов, что, для соответствующей величины X_n нет концентрации ни в каком конечном множестве точек. Иными словами, не найдется последовательности множеств Σ_n, $n \in \mathbb{N}$, и положительного числа C таких, что, во-первых, для всех $n \in \mathbb{N}$ выполнено $|\Sigma_n| < C$, а, во-вторых, с вероятностью, стремящейся к 1 при $n \to \infty$, $X_n \in \Sigma_n$.

Разумеется, мы начали с поиска такой последовательности $e = e(k)$, что $\mathcal{X}_n[e]$ обладает требуемым свойством. Вполне естественно проверить, подойдет ли нам «среднее» количество ребер $e(k) = p\binom{k}{2} + O(1)$. Нам удалось доказать, что эта последовательности действительно является требуемой и что класс таких последовательностей можно довольно сильно расширить. Строгие формулировки полученных результатов приведены в следующем разделе.

3. Новые результаты

Во-первых, мы доказали концентрацию размера наибольшего индуцированного дерева $t(G(n,p))$ в случайном графе $G(n,p)$ в четырех точках.[1]

Теорема 1. *Пусть $\hat{k} = \hat{k}(n)$ таково, что*

$$e^{\hat{k}\ln n - \frac{5}{2}\ln\hat{k} + \hat{k} - \binom{\hat{k}}{2}\ln[1/(1-p)] + (\hat{k}-1)\ln[p/(1-p)] - \frac{1}{2}\ln(2\pi)} = 1.$$

Тогда с вероятностью, стремящейся к 1, выполнены неравенства

$$\lfloor\hat{k}\rfloor - 2 \leqslant t(G(n,p)) < \lceil\hat{k}\rceil + 1.$$

Во-вторых, мы доказали, что в случае $e(k) = \binom{k}{2}p + O(k)$ величина $\mathcal{X}_n[e]$ не сконцентрирована ни в каком конечном множестве.[2]

[1] Этот результат получен в совместной работе с Д. Камальдиновым.

[2] Этот результат получен в совместной работе с Дж. Балогом.

Теорема 2. *Пусть* $e(k) = \binom{k}{2}p + O(k)$ — *последовательность неотрицательных целых чисел.*

1. *Существует такое* $t > 0$, *что, если* $c > t$ *и* $C > 2c + t$, *то*

$$0 < \liminf_{n\to\infty} \mathsf{P}\left(n - C\sqrt{\frac{n}{\ln n}} < \mathcal{X}_n(e) < n - c\sqrt{\frac{n}{\ln n}}\right) <$$

$$\limsup_{n\to\infty} \mathsf{P}\left(n - C\sqrt{\frac{n}{\ln n}} < \mathcal{X}_n(e) < n - c\sqrt{\frac{n}{\ln n}}\right) < 1.$$

2. *Пусть для любой последовательности* $m_k = O(\sqrt{k/\ln k})$ *целых неотрицательных чисел выполнено следующее условие:* $\left|\left(e(k) - \binom{k}{2}p\right) - \left(e(k - m_k) - \binom{k-m_k}{2}p\right)\right| = o(k)$. *Тогда для любого* $\varepsilon > 0$ *найдутся такие* c, C, *что*

$$\liminf_{n\to\infty} \mathsf{P}\left(n - C\sqrt{\frac{n}{\ln n}} < \mathcal{X}_n(e) < n - c\sqrt{\frac{n}{\ln n}}\right) > 1 - \varepsilon.$$

Список литературы

1. Bollobás B. Random Graphs / 2nd Edition. Cambridge University Press, 2001.
2. Janson S., Łuczak T., Rucinski A. Random Graphs. New York: Wiley, 2000.
3. Bollobás B., Erdős P. Cliques in random graphs // Math. Proc. Camb. Phil. Soc. 1976. Vol. 80. P. 419-427.
4. Matula D. The employee party problem // Not. Amer. Math. Soc. 1972. Vol. 19, No. 2. P. A-382.
5. Matula D. The largest clique size in a random graph // Tech. Rep. Dept. Comp. Sci. Southern Methodist University, Dallas, Texas, 1976.
6. Dutta K., Subramanian C.R. On Induced Paths, Holes and Trees in Random Graphs // Proc. ANALCO 2018. New Orleans, Louisiana USA, 2018. P. 168-177.
7. Erdős P., Palka Z. Trees in random graphs // Discrete Mathematics. 1983. Vol. 46. P. 145-150.
8. Ruciński A. Induced subgraphs in a random graph // Annals of Discrete Mathematics. 1987. Vol. 33. P. 275-296.
9. Palka Z. Bipartite complete induced subgraphs of a random graph // Annals of Discrete Mathematics. 1985. Vol. 28. P. 209-219.
10. Fountoulakis N., Kang R.J., McDiarmid C. Largest sparse subgraphs of random graphs // European Journal of Combinatorics. 2014. Vol. 35. P. 232-244.