УДК 519.175.4

МАКСИМАЛЬНЫЕ ИНДУЦИРОВАННЫЕ ДЕРЕВЬЯ В РАЗРЕЖЕННЫХ СЛУЧАЙНЫХ ГРАФАХ

© 2024 г. Х. К. Буйтраго Оропеса^{1, *}

Представлено академиком РАН В.В. Козловым

Поступило 13.03.2024 г.

После доработки 28.03.2024 г.

Принято к публикации 01.04.2024 г.

Мы доказали, что для любого $\varepsilon > 0$ и $n^{-\frac{e-2}{3e-2}+\varepsilon} \leq p = o(1)$, максимальный размер индуцированного дерева в биномиальном случайном графе $G(n, p)$ сконцентрирован в двух последовательных значениях с вероятностью, стремящейся к 1, при $n \rightarrow \infty$.

Ключевые слова: биномиальный случайный граф, максимальный подграф, концентрация

DOI: 10.31857/S2686954324020133, EDN: XHSTXC

В этой статье мы обсудим некоторые характеристики биномиального случайного графа $G(n, p)$ (см. [1]–[6]), где $p = p(n)$ является функцией от n . Рассмотрим набор помеченных вершин $[n] := \{1, \dots, n\}$. В $G(n, p)$ каждая из $\binom{n}{2}$ пар

вершин соединяется ребром независимо с вероятностью p (более формально, $G(n, p)$ – это случайный элемент, принимающий значения во множестве всех графов на $\{1, \dots, n\}$, с распределением $P(G(n, p) = H) = p^{e(H)}(1-p)^{\binom{n}{2}-e(H)}$, где $e(H)$ – количество ребер в H).

Напомним, что множество $U \subset [n]$ вершин графа, в котором никакие две вершины не смежны, называется *независимым множеством*. Размер наибольшего независимого множества называется *числом независимости графа*. Распределение размера самого большого независимого множества в случайном графе тщательно изучено. В [7]–[8] было доказано, что число независимости $G(n, p = \text{const})$ с вероятностью, стре-

мящейся к 1, принимает одно из двух значений, $f_0(n)$ и $f_0(n) + 1$, где

$$f_0(n) = \left\lfloor 2\log_b n - 2\log_b \log_b n + 2\log_b \frac{e}{2} + 0.9 \right\rfloor,$$
$$b = 1/(1-p).$$

Некоторые уточнения этого результата можно найти в [9]. В таких случаях говорят, что число независимости *сконцентрировано в двух точках*.

В дальнейшем концентрация в двух точках была доказана и для других характеристик случайного графа. Естественным обобщением задачи о числе независимости является следующий вопрос: каков максимальный размер *индуцированного* подграфа (множество $A \subset V(G)$ вершин графа G индуцирует в нем подграф $G|_A$, множество вершин которого равно A , а множество ребер образовано всеми ребрами графа G , оба конца которых принадлежат A) в $G(n, p)$, обладающего заданным свойством? В частности, верна ли концентрация в двух точках при некотором более слабом ограничении на число ребер индуцированного подграфа?

Пусть $t: \mathbb{Z}_{\geq 0} \rightarrow \mathbb{Z}_{\geq 0}$ – некоторая функция. Пусть $X_n[t]$ – максимальный размер k множества вершин в $G(n, p)$, индуцирующего подграф с не более $t(k)$ ребрами, а $Y_n[t]$ – максимальный размер k множества вершин в $G(n, p)$, индуцирующего подграф с ровно $t(k)$ ребрами. В [10]

¹ Московский физико-технический институт (национальный исследовательский университет), кафедра дискретной математики, лаборатория продвинутой комбинаторики и сетевых приложений, Долгопрудный, Московская область, Россия

*E-mail: buitrago.okh@phystech.edu

доказано, что при некоторых ограничениях на t величина $X_n[t]$ сконцентрирована в двух точках.

Теорема 1 (N. Fountoulakis, R.J. Kang, C. McDiarmid, 2014). Пусть $t = t(k) = o\left(\frac{k\sqrt{\ln k}}{\sqrt{\ln \ln k}}\right)$, $p = \text{const} \in (0, 1)$,

$$f_t(n) = \left[\begin{array}{l} 2\log_b n + (t-2)\log_b \log_b(np) - \\ -t\log_b t + t\log_b \frac{2pe}{1-p} + 2\log_b \frac{e}{2} + 0.9 \end{array} \right].$$

Тогда с вероятностью, стремящейся к 1, $X_n[t] \in \{f_t(n), f_t(n) + 1\}$.

С другой стороны, в [11] было доказано, что для случайной величины $Y_n[t]$ при t , близких к $p\binom{k}{2}$, концентрации в конечном множестве точек нет.

Теорема 2 (J. Balogh, M. Zhukovskii, 2019).

Пусть $t(k) = \binom{k}{2}p + O(k)$.

1. Найдется такое число $\mu > 0$, что для любых $c > \mu$ и $C > 2c + \mu$ выполнено

$$\begin{aligned} & 0 < \liminf_{n \rightarrow \infty} P \times \\ & \times \left(n - C\sqrt{\frac{n}{\ln n}} < Y_n[t] < n - c\sqrt{\frac{n}{\ln n}} \right) \leq \\ & \leq \limsup_{n \rightarrow \infty} P \times \\ & \times \left(n - C\sqrt{\frac{n}{\ln n}} < Y_n[t] < n - c\sqrt{\frac{n}{\ln n}} \right) < 1. \end{aligned}$$

2. Пусть для любой последовательности $m_k = O(\sqrt{k / \ln k})$ целых неотрицательных чисел выполнено

$$\left| \left(t(k) - \binom{k}{2}p \right) - \left(t(k - m_k) - \binom{k - m_k}{2}p \right) \right| = o(k).$$

Тогда для любого $\varepsilon > 0$ найдутся такие c, C , что

$$\begin{aligned} & \liminf_{n \rightarrow \infty} P \times \\ & \times \left(n - C\sqrt{\frac{n}{\ln n}} < Y_n[t] < n - c\sqrt{\frac{n}{\ln n}} \right) > 1 - \varepsilon. \end{aligned}$$

Наконец, в [12] была доказана концентрация случайной величины $Y_n[t]$ при малых t .

Теорема 3 (D. Kamalidinov, A. Skorkin, M. Zhukovskii, 2021). Пусть $R > 0$. Найдется $\varepsilon > 0$ такое, что для любой последовательности неотрицательных целых чисел $t(k)$, удовлетворяющей следующим ограничениям:

- $|t(k+1)/t(k) - 1| \leq \frac{R}{k}$ при всех достаточно

больших k ,

- $t(k) < \varepsilon k^2$ при всех достаточно больших k , найдется $f(n)$, для которого справедливо $|f(n) - 2\log_{1/(1-p)} n| \leq (3\varepsilon \ln(1/\varepsilon)) \ln n$, и с вероятностью, стремящейся к 1, $Y_n[t] \in \{f(n), f(n) + 1\}$.

Кроме того, вопрос ставился и при заданных структурных свойствах индуцированного подграфа. Так, в [12] изучен максимальный размер индуцированного дерева в $G(n, p = \text{const})$ и доказана концентрация в двух точках.

Теорема 4 (D. Kamalidinov, A. Skorkin, M. Zhukovskii, 2021). Пусть $f_\varepsilon(n) = \lfloor 2\log_{1/(1-p)}(enp) + 2 + \varepsilon \rfloor$. Тогда найдется $\varepsilon > 0$, при котором с вероятностью, стремящейся к 1, максимальный размер индуцированного дерева в $G(n, p = \text{const})$ принимает одно из двух значений, $f_\varepsilon(n)$ и $f_\varepsilon(n) + 1$.

В [13] аналогичный размер получен для максимального размера индуцированного леса.

Теорема 5 (M. Krivoshepa, M. Zhukovskii, 2021). Найдется $\varepsilon > 0$, при котором с вероятностью, стремящейся к 1, максимальный размер индуцированного леса в $G(n, p = \text{const})$ принимает одно из двух значений, $\lfloor 2\log_{1/(1-p)}(enp) + 2 + \varepsilon \rfloor$ и $\lfloor 2\log_{1/(1-p)}(enp) + 3 + \varepsilon \rfloor$.

Все предыдущие результаты применимы к модели случайного графа, где вероятность появления ребра является постоянной величиной. Некоторые из этих результатов были расширены для случая $p(n) = o(1)$. Так, в [14] было доказано, что число независимости $G(n, p)$, где $n^{-2/3+\varepsilon} < p < 1 / [\log(n)]^2$, сконцентрировано в двух точках.

Теорема 6 (T. Bohman, J. Hofstad, 2023). Пусть $\varepsilon > 0$ и $n^{-2/3+\varepsilon} < p < 1 / [\log(n)]^2$. Тогда найдется $k_z = k_z(n) \in \mathbb{Z}$, при котором с вероятностью, стремящейся к 1, число независимости случайного графа $G(n, p)$ принимает одно из двух значений, k_z и $k_z + 1$.

Следующим естественным шагом служит получение результата о концентрации для других изученных ранее величин при $p = o(1)$. Отметим, что доказательства всех предыдущих результатов легко обобщаются на медленно убывающую p , а именно $p = n^{-o(1)}$.

В настоящей работе нам удалось доказать, что результат теоремы 4 о максимальном раз-

мере индуцированного дерева в $G(n, p)$ можно расширить для достаточно широкого диапазона $p = o(1)$.

Теорема 7. Пусть $\varepsilon > 0$, $n^{-\frac{e-2}{3e-2}+\varepsilon} \leq p = o(1)$.

Тогда найдется такое $\delta > 0$, что с вероятностью, стремящейся к 1, максимальный размер индуцированного дерева в $G(n, p)$ принимает одно из двух значений, $g(n)$ и $g(n) + 1$, где

$$g(n) = \lfloor 2 \log_{1/(1-p)}(enp) \rfloor + \delta \lfloor n \rfloor.$$

Несмотря на то, что в основе доказательства теоремы 7 лежит стандартная техника – применение неравенства Маркова для получения верхней оценки на случайную величину и неравенства Чебышева для получения нижней – доказательство из статьи [12] для случая $p = \text{const}$ на рассмотренный нами случай непосредственно не переносится. Проблема, как обычно, заключается в необходимости достаточно точного оценивания второго момента случайной величины X_k , равной количеству индуцированных деревьев размера k . Эту величину можно представить как сумму индикаторов $I_{B(U)}$ по всем k -элементным множествам $U \subset [n]$, где $B(U)$ – событие заключающееся в том, что U индуцирует дерево в $G(n, p)$. Тогда $\mathbb{E}X_k^2 = \sum_{U, U'} \mathbb{P}(B(U) \cap B(U'))$. Заметим, что $\mathbb{E}X_k^2$ можно представить как

$$\sum_{\ell=0}^{k-1} \sum_{r=0}^{\ell-1} \binom{n}{k} \binom{k}{\ell} \binom{n-k}{k-\ell} N(k, \ell, r) p^{2(k-1)-r} \times \\ \times (1-p)^{2 \binom{k}{2} - \binom{\ell}{2} - 2(k-1)+r}$$

где $N(k, \ell, r)$ – количество способов построить два помеченных дерева на k вершинах на заданном множестве $2k - \ell$ вершин, пересекающихся по заданному множеству ℓ вершин так, что в их пересечении находится ровно r ребер. В [12] был предложен метод оценивания $N(k, \ell, r)$, который работает достаточно успешно при $p = \text{const}$. Как только вероятность $p = n^{-\alpha}$ становится достаточно маленькой, этот метод перестает давать достаточную оценку. Тем не менее, нам удалось доказать, что концентрация в двух точках остается верна в достаточно широком диапазоне вероятностей p , вплоть до $n^{-\frac{e-2}{3e-2}}$, посредством уточнения оценки величины $N(k, \ell, r)$. А именно, пусть $\hat{k} = \lfloor 2 \log_{1/(1-p)}(enp) \rfloor + O(1)$ – наибольшее k , при котором $\mathbb{E}X_k \geq 1$. Мы доказали, что $g(n)$ из формулировки теоремы совпадает с $\hat{k} - 1$. Для этого достаточно доказать, что 1) $\mathbb{E}X_{\hat{k}+1} \rightarrow 0$,

2) $\text{Var}X_{\hat{k}-1} = o((\mathbb{E}X_{\hat{k}-1})^2)$. Первое утверждение не составляет труда, так как математическое ожидание вычисляется явно:

$$\mathbb{E}X_k = \binom{n}{k} k^{k-2} p^{k-1} (1-p)^{\binom{k}{2}-k+1}$$

Для доказательства второго утверждения достаточно показать, что

$$\sum_{\ell=0}^{k-1} \sum_{r=0}^{\ell-1} \binom{n}{k} \binom{k}{\ell} \binom{n-k}{k-\ell} N(k, \ell, r) p^{(k-1)-r} \times \\ \times (1-p)^{2 \binom{k}{2} - \binom{\ell}{2} - 2(k-1)+r} = (1 + o(1)) \mathbb{E}X_k^2,$$

где $k = \hat{k} - 1$. Нетрудно заметить, что слагаемое при $\ell = 0$ уже имеет необходимую асимптотку. Таким образом, остается доказать, что оставшаяся часть суммы асимптотически меньше, чем $\mathbb{E}X_k^2$. Мы разбили эту сумму на две части: $\ell < \ell_0$ и $\ell \geq \ell_0$, где $\ell_0 = k - O(1/p)$, и, воспользовавшись методом из работы [12], мы доказали, что первая часть суммы достаточно мала. Для исследования второй части суммы, мы уточнили оценку величины $N(k, \ell, r)$ с помощью формулы для числа укорененных лесов на заданном множестве размеченных вершин (см. [15]).

Sah и Sawney заметили, что после того, как p переступает пороговое значение $n^{-2/3+o(1)}$, число независимости перестает быть настолько хорошо сконцентрированным (см. [14, 16]). В дальнейшем было бы интересно изучить оптимальность нашего результата и установить, в какой момент максимальный размер индуцированного дерева также перестает быть сконцентрированным.

СПИСОК ЛИТЕРАТУРЫ

1. Bollobás B. Random Graphs. 2nd ed. Cambridge University Press, 2001.
2. Janson S., Luczak T., Ruciński A. Random Graphs. New York: Wiley, 2000.
3. Жуковский М.Е., Райгородский А.М. Случайные графы: модели и предельные характеристики // Успехи математических наук. 2015. Т. 70. № 1. С. 35–88.
4. Деревянко Н.М., Киселев С.Г. Числа независимости случайных подграфов некоторого дистанционного графа // Проблемы передачи информации. 2017. Т. 53. № 4. С. 307–318.
5. Егорова А.Н., Жуковский М.Е. Опровержение закона нуля или единицы для экзистенциальных

- монадических свойств разреженного биномиального случайного графа // Доклады Академии наук. Т. 99. № 1. С. 68–70.
6. *Ostrovsky L.B., Zhukovskii M.E.* Monadic second-order properties of very sparse random graphs // *Annals of Pure and Applied Logic*. 2017. V. 168. N 11. P. 2087–2101.
 7. *Bollobás B., Erdős P.* Cliques in random graphs // *Math. Proc. Camb. Phil. Soc.* 1976. V. 80. P. 419–427.
 8. *Matula D.* The largest clique size in a random graph // *Tech. Rep. Dept. Comp. Sci. Dallas, Texas: Southern Methodist University*, 1976.
 9. *Krivelevich M., Sudakov B., Vu V.H., Wormald N.C.* On the probability of independent sets in random graphs // *Random Structures & Algorithms*. 2003. V. 22. N 1. P. 1–14.
 10. *Fountoulakis N., Kang R.J., McDiarmid C.* Largest sparse subgraphs of random graphs // *European Journal of Combinatorics*. 2014. V. 35. P. 232–244.
 11. *Balogh J., Zhukovskii M.* On the sizes of large subgraphs of the binomial random graph // *Discrete Mathematics*. 2022. V. 345. N 2. 112675. ISSN 0012-365X.
 12. *Kamaldinov D., Skorkin A., Zhukovskii M.* Maximum sparse induced subgraphs of the binomial random graph with given number of edges // *Discrete Mathematics*. 2021. V. 344. N 2. 112205. ISSN 0012-365X.
 13. *Krivoshapko M., Zhukovskii M.* Maximum induced forests in random graphs // *Discrete Applied Mathematics*. 2021. V. 305. P. 211–213.
 14. *Bohman T., Hofstad J.* Two-Point Concentration of the Independence Number of the Random Graph // 2022. arXiv:2208.00117.
 15. *Moon J.W.* Counting Labelled Trees. Canadian Mathematical Monograph. 1970.
 16. *Bohman T., Warnke L., Zhu E.* Two-point concentration of the domination number of random graphs // 2024. arXiv:2401.10486.

MAXIMUM INDUCED TREES IN SPARSE RANDOM GRAPHS

J. C. Buitrago Oropeza^{a, *}

^a*Moscow Institute of Physics and Technology (National Research University), Department of Discrete Mathematics, Advanced Combinatorics and Network Applications Laboratory, Dolgoprudny, Moscow region, Russia*

Presented by Academician of the RAS V.V. Kozlov

We prove that a.a.s. for any $\varepsilon > 0$ and $n^{-\frac{e-2}{3e-2}+\varepsilon} \leq p = o(1)$ the maximum size of an induced subtree of the binomial random graph $G(n, p)$ is concentrated in 2 consecutive points.

Keywords: binomial random graph, maximum subgraph, concentration