УДК 519.179.4

М. М. Кошелев[1,2], *Д. А. Шабанов*[2]

[1]Московский государственный университет им. М. В. Ломоносова, механико-математический факультет

[2]Московский физико-технический институт (национальный исследовательский университет), лаборатория комбинаторных и геометрических структур

О размере и сложности компонент связности случайного гиперграфа

В работе исследуются предельные распределения размеров и сложностей компонент связности случайного гиперграфа в биномиальной модели $H(n,k,p)$. Рассматривается ситуация «внутри фазового перехода», где $p = p(n)$ представляется в виде $p = \frac{\lambda}{(k-1)\binom{n-1}{k-1}}$ при $\lambda = \lambda(n)$, удовлетворяющем соотношению $(\lambda - 1)n^{1/3} \sim (k-1)^{2/3}\alpha$ при фиксированном $\alpha \in \mathbb{R}$. Основной результат работы состоит в получении обобщения результата Д. Олдоса (1997) о совместных предельных распределениях размеров и сложностей компонент случайного графа на случай $H(n,k,p)$.

Ключевые слова: случайные гиперграфы, компоненты связности, броуновское движение, слабая сходимость

M. M. Koshelev[1,2], *D. A. Shabanov*[2]

[1]Lomonosov Moscow State University

[2]Moscow Institute of Physics and Technology

On the size and complexity of connectivity components of a random hypergraph

The paper deals with finding the limit distributions of sizes and complexities of connectivity components of a random hypergraph in the binomial model $H(n,k,p)$. We consider the situation «inside the phase transition», when $p = p(n)$ is equal to $p = \frac{\lambda}{(k-1)\binom{n-1}{k-1}}$ with $\lambda = \lambda(n)$ satisfying the relation $(\lambda - 1)n^{1/3} \sim (k-1)^{2/3}\alpha$ for fixed $\alpha \in \mathbb{R}$. The main result is the generalization to $H(n,k,p)$ of the result of D. Aldous (1997) concerning the joint distributions of sizes and complexities of a random graph.

Key words: random hypergraphs, connectivity components, Brownian motion, weak convergence

1. Введение и история задачи

В работе исследуются предельные распределения размеров и сложностей компонент связности в случайном гиперграфе в биномиальной модели $H(n,k,p)$. Сначала напомним основные определения.

1.1. Определения и модели

В рамках работы мы следуем базовым понятиям теории графов (см., например [1]). Одной из классических моделей случайных графов является знаменитая биномиальная модель случайного графа $G(n,p)$, также известной как модель Эрдеша – Реньи. В данной модели имеется множество из n вершин, а ребра между ними проводятся по схеме Бернулли: случайно и независимо друг от друга каждая пара вершин соединяется ребром с вероятностью p.

Естественным обобщением модели $G(n,p)$ является биномиальная модель случайного k-однородного гиперграфа $H(n,k,p)$. Напомним, что в k-однородном гиперграфе ребрами являются k-подмножества вершин, а не пары, как в графе. В модели $H(n,k,p)$ снова имеется n вершин, и уже каждое k-подмножество включается в качестве ребра в $H(n,k,p)$ независимо от других с вероятностью p. Модели $G(n,p)$ и $H(n,k,p)$ — это центральные объекты изучения теории случайных графов и гиперграфов.

Нам также понадобятся понятия, связанные с компонентами связности в графах и гиперграфах. В графах они хорошо известны (см. [1]), определим в гиперграфах их естественные обобщения. *Путем* в гиперграфе, соединяющим вершины u и v, называется такая чередующаяся последовательность вершин и ребер $(v_1, A_1, \ldots, v_{k-1}, A_{k-1}, v_k)$, что $v_1 = v$, $v_k = u$ и для любого $i = 1, \ldots, k-1$ обе вершины v_i и v_{i+1} содержатся в ребре A_i. Вершины u и v называются *связанными* в гиперграфе, если в нем есть путь, соединяющий эти вершины. Понятие связанности, очевидно, является отношением эквивалентности (считаем, что вершина связана сама с собой), и, стало быть, вершиным гиперграфа разбиваются на классы эквивалентности — *компоненты связности*. Наконец, напомним понятие *сложности компоненты связности*. Если C — это компонента связности в k-однородном гиперграфе, имеющая t вершин и l ребер, то ее сложностью называется величина

$$r_k(C) = (k-1)l - t + 1.$$

Тем самым, например, в графе ($k = 2$) древесная компонента имеет нулевую сложность, а унициклическая — единичную. В k-однородном гиперграфе нулевую сложность будут иметь компоненты, являющиеся гипердеревьями.

Нам также понадобятся некоторые определения из теории случайных процессов. Пусть $(B_t, t \geq 0)$ — случайный процесс с непрерывными траекториями. Определим для него неотрицательный процесс с непрерывными траекториями: $B'_t = B_t - \min_{s \leq t} B_s$, $t \geq 0$. Заметим, что $B'_t = 0$ тогда и только тогда, когда процесс B_t достиг своего минимума на отрезке $[0,t]$ в точке t. Теперь рассмотрим все такие пары чисел (l_i, r_i), для которых выполняются следующие свойства:

1) $l_i < r_i$;

2) $B'_{l_i} = B'_{r_i} = 0$;

3) для всех $l_i < m < r_i$ выполняется неравенство $B'_m > 0$.

Упорядочим множество пар (l_i, r_i), $i \in \mathbb{N}$, по невозрастанию длин. Полученная последовательность называется последовательностью *экскурсий* процесса B_t.

Далее, для указанного выше процесса B_t определим также точечный процесс $(N_t, t \geq 0)$, как считающий процесс с условием, что процесс $N_t - \int_0^t B'_s ds$ является мартингалом. Наконец, обозначим через δ_i число точек процесса N_t на отрезке (l_i, r_i).

1.2. Известные результаты

Размеры и сложности компонент случайного графа $G(n,p)$ достаточно хорошо изучены. Безусловно, ответ сильно зависит от того, как ведет себя функция $p = p(n)$. Феномен фазового перехода в структуре компонент $G(n,p)$ был открыт еще П. Эрдешем и А. Реньи [2].

Они показали, что здесь пороговым является значение $1/n$. Обозначим через $C_j(n)$ j-й по величине размер компоненты случайного графа $G(n,p)$, а через $\sigma_j(n)$ — сложность j-й по размеру компоненты. В данных обозначениях классические результаты из работ [2], [3], [4] и монографий [5], [6] можно суммировать следующим образом.

1) Если $p = o(1/n)$, то $C_1(n) = o_P(\ln n)$, а все компоненты с вероятностью, стремящейся к 1, являются деревьями.

2) Если $p = c/n$ и c — фиксированное число из $(0,1)$, то
$$\frac{C_1(n)}{\ln n} \xrightarrow{P} \alpha(c) = \frac{1}{c - \ln c - 1}, \quad n \to \infty.$$
Кроме того, с вероятностью, стремящейся к 1, компонент сложности более 1 в $G(n,p)$ нет, а уницикличеcкие компоненты имеют общий размер $O_P(1)$ (как следствие, получаем, что $\sigma_1(n) = 0$).

3) Если $p = c/n$ и c — фиксированное число из $(1,+\infty)$, то
$$\frac{C_1(n)}{n} \xrightarrow{P} \beta(c), \quad \frac{C_2(n)}{\ln n} \xrightarrow{P} \alpha(c) = \frac{1}{c - \ln c - 1}, \quad n \to \infty,$$
где $\beta(c)$ — это единственное решение уравнения $\beta + e^{-\beta c} = 1$ на интервале $(0,1)$. Кроме того, с вероятностью, стремящейся к 1, все компоненты, кроме самой большой, имеют сложность не более 1, а уницикличеcкие компоненты — снова общий размер $O_P(1)$ (как следствие, получаем, что $\sigma_2(n) = 0$).

4) Если $p = 1/n$, то
$$C_1(n) = \Theta_P(n^{2/3}).$$
Кроме того, максимальная сложность компонент ограничена по вероятности, а древесные компоненты в среднем занимают $n - O(n^{2/3})$ вершин.

5) Если $np - \ln n \to +\infty$, то с вероятностью, стремящейся к 1, случайный граф является связным.

Перечисленные результаты многократно усиливались, исследователей интересовали точные предельные распределения перечисленных величин, а не только их асимптотическое поведение. Например, в работе В. Е. Степанова [3] была доказана асимптотическая нормальность $C_1(n)$

Теорема 1. (Д. Олдос, [7]). *Пусть $pn = 1 + \alpha n^{-1/3}$ для фиксированного $\alpha \in \mathbb{R}$. Пусть $(W_t, t \geq 0)$ — стандартное броуновское движение. Положим*
$$B_t = W_t + \alpha t - \frac{t^2}{2}$$
и введем последовательность $(\gamma_j, \delta_j, j \in \mathbb{N})$ — последовательность экскурсий этого процесса вместе с числами точек соответствующего точечного процесса. Тогда для любого $m \in \mathbb{N}$ выполнена следующая сходимость по распределению:
$$\Big((n^{-2/3} \cdot C_j(n), \sigma_j(n)), j \leq m\Big) \xrightarrow{d} ((|\gamma_j|, \delta_j), j \leq m).$$

Размеры и сложности компонент связности случайного гиперграфа $H(n,k,p)$ изучены заметно хуже. Первыми эволюцию случайного гиперграфа подробно исследовали Дж. Шмидт-Прузан и Э. Шамир [8], которые показали, что фазовый переход в модели $H(n,k,p)$ происходит при $p = c/((k-1)\binom{n}{k-1})$, где $c > 0$ не зависит от n. Обозначим через $C_j^{(k)}(n)$, $j \geq 1$ — последовательность размеров компонент $H(n,k,p)$, отсортированных по убыванию. Тогда авторы в [8] показали, что

1) если $c < 1$, то с вероятностью, стремящейся к 1, $C_1^{(k)}(n) = O(\ln n)$, а все компоненты имеют сложность не более 1;

2) если $c = 1$, то $C_1^{(k)}(n) = O_P(n^{2/3})$;

3) если $c > 1$, то существует такая величина $a = a(c) > 0$, что с вероятностью, стремящейся к 1, $C_1^{(k)}(n) > a \cdot n$.

Тем самым снова при переходе некоторой границы максимальный размер компоненты $H(n, k, p)$ меняется с логарифмического порядка на линейный, а в граничном значении имеет порядок $n^{2/3}$. Однако никаких точных распределений авторами [8] не было найдено. В работе [9] была доказана асимптотическая нормальность $C_1^{(k)}(n)$ в третьем случае при $c > 1$. Более простое доказательство этого же результата было предложено в работе Б. Боллобаша и О. Риордана [10]. Авторы [10] получили частичное обобщение теоремы 1 на случай модели $H(n, k, p)$, в которой было найдено предельное совместное распределение самых больших m размеров компонент в ситуации, когда мы находимся внутри фазового перехода.

Теорема 2. (Б. Боллобаш, О. Риордан, [10]). *Пусть* $p = \frac{\lambda}{(k-1)\binom{n}{k-1}}$, *где* $\lambda = \lambda(n)$ *удовлетворяет соотношению* $(\lambda - 1)^3 n \to (k-1)^2 \alpha^3$ *при* $n \to \infty$ *для фиксированного* $\alpha \in \mathbb{R}$. *Пусть* $(W_t, t \geq 0)$ *— стандартное броуновское движение. Положим*

$$B_t = W_t + \alpha t - \frac{t^2}{2},$$

и пусть $(\gamma_j, j \in \mathbb{N})$ *— последовательность экскурсий этого процесса. Тогда для любого* $m \in \mathbb{N}$ *выполнена следующая сходимость по распределению:*

$$\left((k-1)^{1/3} n^{-2/3} \cdot C_j^{(k)}(n), j \leq m\right) \xrightarrow{d} (|\gamma_j|, j \leq m).$$

Однако вопрос сложности компонент в работе [10] не обсуждался. Целью настоящей работы является закрытие данного пробела.

1.3. Новый результат

Основной результат настоящей работы состоит в усилении теоремы 2 и получении полного обобщения теоремы 1. Обозначим через $\nu_j(n)$ сложность j-й по размеру компоненты случайного гиперграфа $H(n, k, p)$.

Теорема 3. *Пусть* $p = \frac{\lambda}{(k-1)\binom{n}{k-1}}$, *где* $\lambda = \lambda(n)$ *удовлетворяет соотношению* $(\lambda - 1)^3 n \to (k-1)^2 \alpha^3$ *при* $n \to \infty$ *для фиксированного* $\alpha \in \mathbb{R}$. *Пусть* $(W_t, t \geq 0)$ *— стандартное броуновское движение. Положим*

$$B_t = W_t + \alpha t - \frac{t^2}{2},$$

и введем последовательность $(\gamma_j, \delta_j, j \in \mathbb{N})$ *— последовательность экскурсий этого процесса вместе с числами точек соответствующего точечного процесса. Тогда для любого* $m \in \mathbb{N}$ *выполнена следующая сходимость по распределению:*

$$\left(((k-1)^{1/3} n^{-2/3} \cdot C_j^{(k)}(n), \nu_j(n)), j \leq m\right) \xrightarrow{d} ((|\gamma_j|, \delta_j), j \leq m).$$

В теореме 3 найдено, тем самым, не только предельное распределение сложности компонент, но и совместное распределение первых m по размеру компонент вместе с их сложностями. Перейдем к доказательству теоремы 3, которое будет разбито на несколько частей.

2. Алгоритм BFS

Доказательство теоремы 3 следует идеям из работы [7]. В силу того, что Боллобаш и Риордан дали доказательство теоремы 2 весьма кратко, иногда просто ссылаясь на доказательство теоремы 1 из [7] без деталей, мы приведем доказательство сходимости по распределению и для размеров компонент.

В основе доказательства [7] лежит рассмотрение алгоритма BFS набора компонент связности случайного графа, который затем можно хорошо аппроксимировать случайным процессом с непрерывным временем. Для каждого момента времени $t \in \mathbb{Z}_+$ рассмотрим последовательность троек (U_t, Q_t, O_t) множеств вершин гиперграфа. Они преобразуются по следующему правилу.

- В начальный момент времени множество U_0 пусто, Q_0 состоит из одной вершины, а O_0 — из всех остальных вершин гиперграфа.
- Если Q_t непусто, то берем из него вершину v и рассматриваем все ребра, которые полностью лежат в $Q_t \cup O_t$ и содержат v. Пусть их объединение дает множество E_{t+1}.
- Тогда $U_{t+1} = U_t \cup \{v\}$, $Q_{t+1} = Q_t \setminus \{v\} \cup (E_{t+1} \cap O_t)$, $O_{t+1} = O_t \setminus E_{t+1}$.
- Если же Q_t пусто, то мы берем вершину v из O_t, E_{t+1} определяем аналогично, после чего полагаем $U_{t+1} = U_t \cup \{v\}$, $Q_{t+1} = E_{t+1} \cap O_t$, $O_{t+1} = O_t \setminus E_{t+1}$.

Определим также C_t как количество компонент связности, полностью лежащих внутри U_t. Суть множеств, тем самым, проста. Множество Q_t — это очередь вершин, из которых мы еще не осуществляли поиск, U_t — уже рассмотренные вершины, а O_t — еще неактивные вершины. Отметим простые свойства алгоритма BFS.

1) $|U_t| = t$.

2) $|Q_t| = 0$ тогда и только тогда, когда мы обошли очередную компоненту связности.

3) Пусть $T = T(n) = O(n^{3/4})$. Тогда с вероятностью $q = 1 - o(1)$ любая пара ребер, рассмотренная для перехода от t к $t+1$ при $t < T$ пересекается лишь по вершине v. Действительно, посчитаем математическое ожидание пар ребер, которые пересекаются по еще какой-нибудь вершине. На t-м шаге оно не превосходит

$$n\binom{n}{k-2}^2 p^2(n) \le n^{2k-3}p^2(n) = O\left(\frac{1}{n}\right).$$

Тогда суммарное среднее число таких пар по $t < T$ не превосходит $O(T/n) = O(n^{-1/4})$. По неравенству Маркова $1 - q = O(n^{-1/4})$.

4) Пусть $X_t = |Q_t| - C_t$. Тогда $X_{t+1} - X_t = \eta_{t+1} - 1$, где η_{t+1} — количество вершин, добавленных в очередь на шаге t.

3. Свойства процесса X_t

Обсудим свойства процесса X_t. Обозначим через $(\mathcal{F}_t, t \ge 0)$ фильтрацию, порожденную $((U_t, Q_t, O_t, C_t), t \ge 0)$. Вычислим $E(\eta_{t+1}|\mathcal{F}_t)$. Обозначим $o'_t = |O_t|$, если Q_t непусто и $|O_t| - 1$ в противном случае. Тогда $E(\eta_{t+1}|\mathcal{F}_t) = o'_t(1 - (1-p)^{\binom{n-t-2}{k-2}})$, откуда получаем

$$E(\eta_{t+1}|\mathcal{F}_t) = o'_t\left(p\binom{n-t-2}{k-2} + O\left(p^2\binom{n-t-2}{k-2}^2\right)\right) = \frac{o'_t\lambda(1-\frac{t+2}{n})^{k-2}}{n}(1 + O(n^{-1})).$$

Положим $\alpha_{t+1} = \frac{\lambda(1-\frac{t+2}{n})^{k-2}}{n}$. Тогда для $D_{t+1} = E(\eta_{t+1} - 1|\mathcal{F}_t)$ имеем представление

$$D_{t+1} = \alpha_{t+1}(n - t - X_t - C_{t+1} + O(1)) - 1.$$

Также определим $\Delta_{t+1} = X_{t+1} - X_t - D_{t+1}$. Нетрудно видеть, что Δ_{t+1} измерима относительно $\mathcal{F}_{t+1}$, а $E(\Delta_{t+1}|\mathcal{F}_t) = 0$.

После подстановки получаем, что

$$X_{t+1} = (1 - \alpha_{t+1})X_t + \alpha_{t+1}(n - t) - 1 + \Delta_{t+1} - \alpha C_{t+1} + R_{t+1},$$

где $R_{t+1} = O(n^{-1})$. Теперь определим последовательность x_t рекурсивным образом:

$$x_{t+1} = (1 - \alpha_{t+1})x_t + \alpha_{t+1}(n - t) - 1, x_0 = 0.$$

Рассмотрим теперь $X_{t+1} - x_{t+1} = (1 - \alpha_{t+1})(X_t - x_t) + \Delta_{t+1} - \alpha_{t+1}C_{t+1} + R_{t+1}$. Применяя эту формулу рекурсивно, получим явное представление для $X_t - x_t$:

$$X_t - x_t = \sum_{i=1}^{t} \frac{\beta_t}{\beta_i}(\Delta_i - \alpha_i C_i + R_i), \tag{1}$$

где $\beta_t = \prod_{i=1}^{t}(1 - \alpha_i)$. Нам пригодится следующее простое утверждение.

Утверждение 1. *Пусть* $S_t = \sum_{i=1}^{t} \frac{1}{\beta_i}\Delta_i$, $\hat{X}_t = x_t + \beta_t S_t$. *Тогда* $|X_t - \hat{X}_t| = O\left(\frac{tC_t}{n}\right)$.

Доказательство. Очевидно из (1), ведь $\alpha_i = O(1/n)$, а величины C_i возрастают:

$$|X_t - \hat{X}_t| = \left|\sum_{i=1}^{t} \frac{\beta_t}{\beta_i}(-\alpha_i C_i + R_i)\right| = O\left(\frac{tC_t}{n}\right).$$

□

Положим $y_t = x_t - n + t$, тогда будет верно соотношение $y_{t+1} = (1 - \alpha_{t+1})y_t$, откуда $y_t = -n\beta_t$, а $x_t = n - t - \beta_t n$. Вычислим теперь асимптотику величины β_t:

$$\begin{aligned}
\ln \beta_t &= \sum_{i=1}^{t} \ln(1 - \alpha_i) = \sum_{i=1}^{t} -\alpha_i - O\left(\sum_{i=1}^{t} \alpha_i^2\right) = \\
&= -\frac{\lambda}{n}\sum_{i=1}^{t}\left(1 - \frac{i+2}{n}\right)^{k-2} - O(tn^{-2}) = \\
&= -\frac{\lambda}{n}\sum_{i=1}^{t} e^{(k-2)\left(-\frac{i+2}{n} + O\left(\left(\frac{i+2}{n}\right)^2\right)\right)} - O(tn^{-2}) = \\
&= -\frac{\lambda}{n}\sum_{i=1}^{t} e^{-\frac{(k-2)(i+2)}{n} + O\left(\left(\frac{t+2}{n}\right)^2\right)} - O(tn^{-2}) =
\end{aligned}$$

$$\begin{aligned}
&= -\frac{\lambda e^{-\frac{2k-4}{n}}}{n}\sum_{i=1}^{t} e^{-\frac{(k-2)i}{n}} + O(t^3n^{-3} + tn^{-2}) = \\
&= -\frac{\lambda e^{-\frac{3k-6}{n}}}{n}\frac{1 - e^{-\frac{(k-2)t}{n}}}{1 - e^{-\frac{k-2}{n}}} + O(t^3n^{-3} + tn^{-2}) = \\
&= -\frac{\lambda(1 + O(\frac{1}{n}))}{n}\frac{t - \frac{(k-2)t^2}{2n} + O\left(\frac{t^3}{n^2}\right)}{1 + O\left(\frac{1}{n}\right)} + O(t^3n^{-3} + tn^{-2}) = \\
&= -\lambda\left(\frac{t}{n} - \frac{(k-2)t^2}{2n^2}\right) + O(t^3n^{-3} + tn^{-2}).
\end{aligned}$$

Следовательно,

$$x_t = n - t - ne^{-\lambda\left(\frac{t}{n} - \frac{(k-2)t^2}{2n^2}\right) + O(t^3 n^{-3} + tn^{-2})}.$$

Введем функцию $g(x) = 1 - x - e^{-\lambda\left(x - \frac{k-2}{2}x^2\right)}$. Тогда при $t \le Cn^{2/3}$ для некоторого фиксированного $C > 0$ имеем $x_t = ng\left(\frac{t}{n}\right) + O(1)$. Исследуем свойства функции $g(x)$ через ее производные:

$$g'(x) = \lambda(1 - (k-2)x)e^{-\lambda\left(x - \frac{k-2}{2}x^2\right)} - 1,$$

откуда $g'(0) = \lambda - 1$. Аналогично,

$$g''(x) = -\lambda^2(1 - (k-2)x)^2 e^{-\lambda\left(x - \frac{k-2}{2}x^2\right)} - \lambda(k-2)e^{-\lambda\left(x - \frac{k-2}{2}x^2\right)},$$

то есть $g''(0) = -\lambda(k-2) - \lambda^2$. Отсюда получаем представление вида

$$g(x) = (\lambda - 1)x - (\lambda(k-2) + \lambda^2)x^2/2 + O(x^3).$$

Для перехода в непрерывное время нам остается получить оценки хвостов распределений Δ_i и S_i.

Лемма 1. *При всех достаточно больших n для любого i выполняется неравенство $P(\Delta_i \ge 2n^{1/5}) \le e^{-n^{1/5}}$.*

Доказательство. Заметим, что $D_i + 1 = E(\eta_i|\mathcal{F}_{i-1}) = o'_{i-1}(1 - p^{\binom{n-i-1}{k-2}}) \le 2$ при всех достаточно больших n. Остается оценить $X_i - X_{i-1} = \eta_i - 1$:

$$P(\eta_i \ge kn^{1/5}) \le \binom{\binom{n-1}{k-1}}{n^{1/5}} p^{n^{1/5}} \le \frac{\lambda^{n^{1/5}}}{(k-1)^{n^{1/5}} n^{1/5}!} \le e^{-n^{1/5}}$$

при всех достаточно больших n. □

Лемма 2. *Для фиксированного $T > 0$ при всех достаточно больших n выполняется*

$$\max_{i \le [(k-1)^{-1/3} n^{2/3} T]} |S_i| \le Cn.$$

Доказательство. Очевидно в силу равномерной ограниченности (и равномерной же отделенности от 0) величин β_i, а также того, что

$$\sum_{1 \le i \le n} X_i - X_{i-1} \le n, D_i \le 2.$$

□

Лемма 3. *Для некоторого $C > 0$ при всех t и n выполняется соотношение*

$$E(\max_{i \le t} S_i^2) \le Ct.$$

Доказательство. Из неравенств Дуба следует, что

$$E(\max_{i \le t} S_i^2) \le 4E|S_t|^2 = 4\sum_{i=1}^{t} D\frac{\Delta_i}{\beta_i} = O\left(\sum_{i=1}^{t} D\Delta_i\right) = O(t).$$

□

4. Переход в непрерывное время

В данном разделе осуществляется предельный переход от процесса X_t к броуновскому движению. Наиболее удобно это сделать с помощью перехода в непрерывное время. А именно, определим для $s \geq 0$ процессы

$$X^*(s) = \frac{X_{[(k-1)^{-1/3}n^{2/3}s]}}{(k-1)^{1/3}n^{1/3}}, \quad \hat{X}^*(s) = \frac{\hat{X}_{[(k-1)^{-1/3}n^{2/3}s]}}{(k-1)^{1/3}n^{1/3}}. \tag{2}$$

Из утверждения 1 следует, что $|X^*(s) - \hat{X}^*(s)| = O(s/n^{1/3})$.

Теперь заметим, что $\hat{X}^*(s) = \beta^*(s)S^*(s) + x^*(s)$, где $\beta^*(s) = \frac{\beta_{[(k-1)^{-1/3}n^{2/3}s]}}{(k-1)^{1/3}n^{1/3}}$,

$$S^*(s) = \sum_{i=1}^{[(k-1)^{-1/3}n^{2/3}s]} \frac{1}{\beta_i}\Delta_i,$$

$$\begin{aligned} x^*(s) = \frac{x_{[(k-1)^{-1/3}n^{2/3}s]}}{(k-1)^{1/3}n^{1/3}} &= \frac{n^{2/3}}{(k-1)^{1/3}} g\left(\frac{[(k-1)^{-1/3}n^{2/3}s]}{n}\right) + o(1) = \\ &= (\lambda-1)n^{1/3}\frac{(k-1)^{-1/3}}{(k-1)^{1/3}}s - (\lambda(k-2)+\lambda^2)\frac{(k-1)^{-2/3}}{(k-1)^{1/3}}\frac{s^2}{2} + o(1) = \\ &= \alpha s - \frac{s^2}{2} + o(1), \end{aligned}$$

где $o(.)$ равномерно мало по всем $s < s_0$.

Осталось показать, что процесс $\beta^*(s)S^*(s)$ сходится по распределению к броуновскому движению W_s. Это следует из следующей фундаментальной теоремы из теории мартингалов.

Теорема 4. *Пусть $M_n = (M_n(t), t \geq 0)$, $n \in \mathbb{N}$ — последовательность локальных мартингалов с непрерывными справа траекториями и $M_n(0) = 0$. Пусть также $(A_n(t), t \geq 0)$ — последовательность процессов, обладающая следующими свойствами:*

1) $A_n(t)$ п.н. возрастает при каждом n;

2) $\lim_{n\to\infty} E\left(\sup_{t\leq T} |A_n(t) - A_n(t-0)|\right) = 0$;

3) $\lim_{n\to\infty} E\left(\sup_{t\leq T}(M_n(t) - M_n(t-0))^2\right) = 0$;

4) $M_n^2(t) - A_n(t)$ — локальный мартингал;

5) для любого $t \geq 0$ выполняется $A_n(t) \xrightarrow{P} c(t)$, где $c(t)$ — неубывающая функция.

Тогда $M_n \xrightarrow{d} X$, где X — это центрированный гауссовский процесс с дисперсией $c(t)$.

Начнем с построения $A_n(t)$. Зафиксируем n и рассмотрим последовательность $\tilde{D}_i = \beta_i^2 S_i^2 - \beta_{i-1}^2 S_{i-1}^2$. Преобразуем это выражение:

$$\begin{aligned} \tilde{D}_i &= \beta_i^2 S_i^2 - \beta_{i-1}^2 S_{i-1}^2 = (\beta_i S_i - \beta_{i-1}S_{i-1})(\beta_i S_i + \beta_{i-1}S_{i-1}) = \\ &= \left(\Delta_i + (\beta_i - \beta_{i-1})\sum_{j=1}^{i-1}\frac{\Delta_j}{\beta_j}\right)\left(\Delta_i + (\beta_i + \beta_{i-1})\sum_{j=1}^{i-1}\frac{\Delta_j}{\beta_j}\right) = \\ &= \Delta_i^2 + 2\beta_i S_{i-1}\Delta_i + (\beta_i^2 - \beta_{i-1}^2)S_{i-1}^2. \end{aligned}$$

Теперь вычислим $E(\tilde{D}_i|\mathcal{F}_{i-1})$:

$$E(\tilde{D}_i|\mathcal{F}_{i-1}) = E(\Delta_i^2|\mathcal{F}_{i-1}) + (\beta_i^2 - \beta_{i-1}^2)S_{i-1}^2 = D(\eta_i|\mathcal{F}_{i-1}) + (\beta_i^2 - \beta_{i-1}^2)S_{i-1}^2.$$

Положим теперь $\tilde{\Delta}_i = \tilde{D}_i - E(\tilde{D}_i|\mathcal{F}_{i-1}) = \tilde{D}_i - D(\eta_i|\mathcal{F}_{i-1}) - (\beta_i^2 - \beta_{i-1}^2)S_{i-1}^2$. Нетрудно видеть, что $\tilde{\Delta}_i$ образуют мартингальные разности. Стало быть, последовательность

$$\tilde{Z}_i = \sum_{j=1}^{i} \tilde{\Delta}_j = \sum_{j=1}^{i} \tilde{D}_j - \sum_{j=1}^{i} E(\tilde{D}_j|\mathcal{F}_{j-1}) = \beta_i^2 S_i^2 - \sum_{j=1}^{i} \left(D(\eta_j|\mathcal{F}_{j-1}) + (\beta_j^2 - \beta_{j-1}^2)S_{j-1}^2\right)$$

является мартингалом. Наконец, для данного зафиксированного n мы готовы определить $A_n(t)$ следующим образом:

$$A_n(t) = \frac{1}{(k-1)^{2/3}n^{2/3}} \sum_{i=1}^{[(k-1)^{-1/3}n^{2/3}t]} \left(D(\eta_i|\mathcal{F}_{i-1}) + (\beta_i^2 - \beta_{i-1}^2)S_{i-1}^2\right).$$

Также для последующих выкладок нам понадобится определить

$$Z_n(t) = \frac{1}{(k-1)^{2/3}n^{2/3}} \tilde{Z}_{[(k-1)^{-1/3}n^{2/3}t]}.$$

Заметим, что для каждого n выполняется тождество

$$(\beta^*(t)S^*(t))^2 = Z_n(t) + A_n(t).$$

Теперь начнем проверять условия теоремы 4. Мартингальность $\beta^*(t)S^*(t)$ относительно фильтрации $\mathcal{F}^*(s) = \mathcal{F}_{[(k-1)^{-1/3}n^{2/3}s]}$ очевидна из определения величин Δ_i. Также по определению $\beta^*(0)S^*(0) = 0$. Как известно, квадрат мартингала является субмартингалом, поэтому $E(\tilde{D}_i|\mathcal{F}_{i-1}) \geq 0$, и, следовательно, $A_n(t)$ является монотонно возрастающей. Таким образом, свойство 1 доказано. Проверим свойство 2:

$$A_n(t) - A_n(t-0) = \frac{1}{(k-1)^{2/3}n^{2/3}} \sum_{i=1}^{[(k-1)^{-1/3}n^{2/3}t]} \left(D(\eta_i|\mathcal{F}_{i-1}) + (\beta_i^2 - \beta_{i-1}^2)S_{i-1}^2\right) -$$
$$- \lim_{\varepsilon\to 0+} \frac{1}{(k-1)^{2/3}n^{2/3}} \sum_{i=1}^{[(k-1)^{-1/3}n^{2/3}(t-\varepsilon)]} \left(D(\eta_i|\mathcal{F}_{i-1}) + (\beta_i^2 - \beta_{i-1}^2)S_{i-1}^2\right).$$

Заметим, что $D(\eta_i|\mathcal{F}_{i-1}) = O(1)$, поэтому

$$\sup_{t\leq T}|A_n(t) - A_n(t-)| \leq \frac{1}{(k-1)^{2/3}n^{2/3}} \sup_{i\leq (k-1)^{-1/3}Tn^{2/3}} |\beta_i^2 - \beta_{i-1}^2|S_{i-1}^2 + O(n^{-2/3}).$$

В силу $\beta_i^2 - \beta_{i-1}^2 \leq \frac{4}{n}$ имеем

$$\sup_{t\leq T}|A_n(t) - A_n(t-)| \leq \frac{4}{(k-1)^{2/3}n^{5/3}} \sup_{i\leq (k-1)^{-1/3}tn^{2/3}} S_{i-1}^2.$$

Остается заметить, что в силу леммы 2 математическое ожидание данного выражения стремится к 0.

Проверим теперь свойство 3 теоремы 4. Распишем разность $\beta^*(s)S^*(s) - \beta^*(s-)S^*(s-)$ чуть подробнее:

$$\beta^*(s)S^*(s) - \beta^*(s-0)S^*(s-0) =$$
$$= \lim_{\varepsilon\to 0} \frac{1}{(k-1)^{1/3}n^{1/3}} \Big((\beta_{[(k-1)^{-1/3}n^{2/3}s]} - \beta_{[(k-1)^{-1/3}n^{2/3}(s-\varepsilon)]})S_{[(k-1)^{-1/3}n^{2/3}(s-\varepsilon)]} +$$
$$+ \Delta_{[(k-1)^{-1/3}n^{2/3}s]}\Big).$$

Переходя к супремуму по t, получаем, что

$$\sup_{t\le T}|\beta^*(s)S^*(s)-\beta^*(s-)S^*(s-)|\le\left|\frac{\sup\limits_{i\le[(k-1)^{-1/3}n^{2/3}T]}(\beta_i-\beta_{i-1})S_{i-1}}{(k-1)^{1/3}n^{1/3}}\right|+\left|\frac{\sup\limits_{i\le[(k-1)^{-1/3}n^{2/3}T]}\Delta_i}{(k-1)^{1/3}n^{1/3}}\right|,$$

откуда тривиально следует, что

$$\sup_{t\le T}|\beta^*(s)S^*(s)-\beta^*(s-)S^*(s-)|\le\left|\frac{4\sup\limits_{i\le[(k-1)^{-1/3}n^{2/3}T]}S_{i-1}}{(k-1)^{1/3}n^{4/3}}\right|+\left|\frac{\sup\limits_{i\le[(k-1)^{-1/3}n^{2/3}T]}\Delta_i}{(k-1)^{1/3}n^{1/3}}\right|.$$

Применяя лемму 2, получаем

$$\sup_{t\le T}|\beta^*(s)S^*(s)-\beta^*(s-)S^*(s-)|\le\frac{C}{n^{1/3}}+\left|\frac{\sup\limits_{i\le[(k-1)^{-1/3}n^{2/3}T]}\Delta_i}{(k-1)^{1/3}n^{1/3}}\right|.$$

Тогда

$$\begin{aligned}E\sup_{t\le T}|\beta^*(s)S^*(s)-\beta^*(s-)S^*(s-)|^2&=E(\sup_{t\le T}|\beta^*(s)S^*(s)-\beta^*(s-)S^*(s-)|)^2\le\\&\le E\left(\frac{C}{n^{1/3}}+\left|\frac{\sup\limits_{i\le[(k-1)^{-1/3}n^{2/3}T]}\Delta_i}{(k-1)^{1/3}n^{1/3}}\right|\right)^2=\\&=\frac{C^2}{n^{2/3}}+2CE\left|\frac{\sup\limits_{i\le[(k-1)^{-1/3}n^{2/3}T]}\Delta_i}{(k-1)^{1/3}n^{2/3}}\right|+\\&+E\left|\frac{\sup\limits_{i\le[(k-1)^{-1/3}n^{2/3}T]}\Delta_i}{(k-1)^{1/3}n^{1/3}}\right|^2.\end{aligned}$$

Первое слагаемое, очевидно, стремится к нулю, поэтому оценим второе и третье, опираясь на лемму 1. Имеем:

$$E\left|\frac{\sup\limits_{i\le[(k-1)^{-1/3}n^{2/3}T]}\Delta_i}{(k-1)^{1/3}n^{2/3}}\right|=O\left(n^{-2/3}\cdot\left[n^{1/5}+[(k-1)^{-1/3}n^{2/3}T]e^{-n^{1/5}}n\right]\right)=o(1),$$

$$E\left|\frac{\sup\limits_{i\le[(k-1)^{-1/3}n^{2/3}T]}\Delta_i}{(k-1)^{1/3}n^{2/3}}\right|^2=O\left(n^{-4/3}\cdot\left[n^{2/5}+[(k-1)^{-1/3}n^{2/3}T]e^{-n^{1/5}}n^2\right]\right)=o(1),$$

что завершает обоснование третьего условия теоремы 4.

Проверим свойство 4. Оно очевидно из соотношения

$$(\beta^*(t)S^*(t))^2=Z_n(t)+A_n(t),$$

а также мартингальности $Z_n(t)$.

Проверим, наконец, свойство 5 для $c(t) = t$. Сперва заметим, что

$$\left| A_n(t) - \frac{1}{(k-1)^{2/3}n^{2/3}} \sum_{i=1}^{[(k-1)^{-1/3}n^{2/3}t]} D(\eta_i|\mathcal{F}_{i-1}) \right| \xrightarrow{P} 0$$

при $n \to \infty$. Действительно, из леммы 3 следует, что $P(\max_{i \le d} S_i^2 \geq d \ln n) \to 0$. Подставляя $d = [(k-1)^{-1/3}n^{2/3}t]$ и вспоминая оценку $|\beta_i^2 - \beta_{i-1}^2| \leq 4/n$, получаем

$$P\left(\frac{1}{(k-1)^{2/3}n^{2/3}} \sum_{i=1}^{[(k-1)^{-1/3}n^{2/3}t]} |\beta_i^2 - \beta_{i-1}^2| S_{i-1}^2 \geq \frac{d^2 \ln n}{(k-1)^{2/3}n^{2/3}} \cdot \frac{4}{n} \right) \to 0,$$

откуда и вытекает утверждение выше. Осталось доказать, что

$$\frac{1}{(k-1)^{2/3}n^{2/3}} \sum_{i=1}^{[(k-1)^{-1/3}n^{2/3}t]} D(\eta_i|\mathcal{F}_{i-1}) \xrightarrow{P} t.$$

Имеем

$$\begin{aligned}
D(\eta_i|\mathcal{F}_{i-1}) &= E(\eta_i^2|\mathcal{F}_{i-1}) - E^2(\eta_i|\mathcal{F}_{i-1}) = \\
&= o_i' \left(1 - (1-p)^{\binom{n-i-1}{k-2}}\right) + o_i'(o_i' - 1)\left(1 - 2(1-p)^{\binom{n-i-1}{k-2}} + \right. \\
&\left. + (1-p)^{2\binom{n-i-1}{k-2} - \binom{n-i-2}{k-3}}\right) - (o_i')^2 \left(1 - (1-p)^{\binom{n-i-1}{k-2}}\right)^2 = \\
&= (o_i')^2 \binom{n-i-2}{k-3} p + o_i' \binom{n-i-1}{k-2} p + O(n^{-1}) = \\
&= n^2 \binom{n-i-2}{k-3} p + n \binom{n-i-1}{k-2} p + O(n^{-1/4}) = \lambda(k-1) + O(n^{-1/4}),
\end{aligned}$$

причем все O равномерны по $t = O(n^{3/4})$. Суммируя полученный результат по i, получаем необходимую сходимость.

5. Сходимость размеров компонент связности

Подведем итоги рассуждений предыдущих параграфов. Из теоремы 4 следует, что процесс $\beta^*(s)S^*(s)$ сходится к броуновскому движению. Тогда в силу (2) мы получаем, что процесс $X^*(s)$ сходится к процессу B_s из условия нашей теоремы. Далее, заметим, что набор компоненты заканчивается в тот момент, когда X_s и, стало быть, $X^*(s)$, достигает своего очередного минимального значения. Значит, размеры компонент будут соответствовать экскурсиям процесса $X^*(s)$. Напрямую следуя общим рассуждениям Олдоса о слабой сходимости таких функционалов из работы [7] (см. параграф 2.3), мы получаем, что длины экскурсий $X^*(s)$ будут сходиться по распределению к длинам экскурсий B_s. Остается заметить, что коэффициент сжатия времени составляет $(k-1)^{-1/3}n^{-2/3}$, поэтому этот же коэффициент будет соответствовать правильной нормировке компонент:

$$\left((k-1)^{1/3}n^{-2/3} \cdot C_j^{(k)}(n), j \leq m\right) \xrightarrow{d} (|\gamma_j|, j \leq m).$$

6. Сходимость сложностей компонент связности

Перейдем к обсуждению сложностей компонент. Сначала заметим, что по свойству 3 обхода в ширину в ходе первых $n^{3/4}$ итераций обхода в ширину сложность может увеличиваться лишь за счет ребер, которые имеют нетривиальное пересечение с очередью Q_i. Но в отличие от случая графов новое ребро может увеличить сложность компоненты более чем на единицу. Поймем, какие у нас есть варианты.

Оценим вероятность того, что новое ребро, проведенное на шаге i, пересекает множество Q_i хотя бы по трем вершинам. Легко видеть, что такая условная вероятность этого события (относительно $\mathcal{F}_{i-1}$) не превосходит

$$|Q_i|^3\binom{n-i-3}{k-4}p = O(|Q_i|^3 n^{-3}).$$

Мы знаем, что с вероятностью, стремящейся к 1, $|Q_i| = O(n^{2/3}\ln n)$, поэтому в типичной ситуации эта условная вероятность равна $O(\ln^3 n/n)$. Суммируя по $i \le n^{3/4}$, получаем, что с вероятностью, стремящейся к 1, подобные ребра не дают вклад в сложность компонент.

Тем самым у нас остаются только ребра, которые пересекают множество Q_i по одной или двум вершинам. Далее, мы следуем рассуждениям из [7]. На каждом шаге i работы BFS мы погрузим процесс набора новых ребер в компоненту в непрерывное время. А именно, для каждого ребра, выходящего из текущей вершины v_i и содержащегося в $Q_i \cup O_i$, рассмотрим независимую равномерную случайную величину на отрезке $[i, i+1]$, которую мы будем считать временем включения этого ребра в компоненту. Рассмотрим считающий процесс $N_n(s)$, который ставит «метку» в тот момент времени, когда появляется ребро, повышающее сложность компоненты. Тогда при $|Q_{\lfloor s\rfloor}| \ge 1$

$$\begin{aligned} P\left(N_n(s) \text{ имеет точку на } [s, s+\Delta s]|\mathcal{F}_{\lfloor s\rfloor}\right) &= (1-p\Delta s)^{\binom{n-\lfloor s\rfloor-1}{k-1}-\binom{n-|Q_{\lfloor s\rfloor}|+1-\lfloor s\rfloor-1}{k-1}} = \\ &= (|Q_{\lfloor s\rfloor}|-1)\frac{pn^{k-2}}{(k-2)!}(1+O(1/n))\Delta s = \\ &= \frac{|Q_{\lfloor s\rfloor}|-1}{n}(1+o(1/n))\Delta s. \end{aligned}$$

Рассмотрим и второй считающий процесс $N'_n(s)$, который ставит «метку» в тот момент времени, когда появляется ребро, повышающее сложность компоненты ровно на 1. Несложно понять, что верно такое утверждение:

$$P\left(N'_n(s) \text{ имеет точку на } [s, s+\Delta s]|\mathcal{F}_{\lfloor s\rfloor}\right) = (|Q_{\lfloor s\rfloor}|-1)\frac{pn^{k-2}}{(k-2)!}(1+O(1/n))\Delta s.$$

Далее, заметим, что

$$\max(|Q_s|-1, 0) = X_s - \min_{t\le s} X_s.$$

Тогда

$$X^*(u) - \min_{t\le u} X^*(u) = (k-1)^{-1/3} n^{-1/3}\left(X_s - \min_{j\le s} X_j\right) = (k-1)^{-1/3}n^{-1/3}\max(|Q_s|-1,0),$$

где $s = [(k-1)^{-1/3}n^{2/3}u]$. Тогда

$$\begin{aligned} &P\left(N_n(s) \text{ имеет точку на } [(k-1)^{-1/3}n^{2/3}u, (k-1)^{-1/3}n^{2/3}(u+\Delta u)]|\mathcal{F}_{(k-1)^{-1/3}n^{2/3}u}\right) = \\ &= \frac{(k-1)^{1/3}n^{1/3}}{n}\left(X^*(u) - \min_{t\le u}X^*(u)\right)(1+o(1))(k-1)^{-1/3}n^{2/3}\Delta u = \\ &= \left(X^*(u) - \min_{t\le u}X^*(u)\right)(1+o(1))\Delta u. \end{aligned}$$

Согласно уже доказанному, процесс $X^*(u) - \min\limits_{t\le u} X^*(u)$ сходится по распределению к $B'_u = B_u - \min_{t\le u} B_t$. Из работы [7] известно, что подобной сходимости уже достаточно, чтобы доказать сходимость процессов $N_n(t)$ и $N'_n(t)$ к точечному процессу N_t с интенсивностью B'_t. Учитывая, что предел один и тот же, мы получаем, что каждый из них корректно считает сложность компонент. Следовательно, доказана искомая совместная сходимость по распределению размеров компонент случайного гиперграфа и их сложностей:

$$\left(((k-1)^{1/3}n^{-2/3} \cdot C_j^{(k)}(n), \nu_j(n)), j \leq m\right) \xrightarrow{d} ((|\gamma_j|, \delta_j), j \leq m).$$

Теорема 3 доказана.

Исследование выполнено за счет гранта Российского научного фонда № 22-21-00411.

Список литературы

1. *Харари Ф.* Теория графов / пер. с англ. Москва : УРСС, 2018. 304 с.
2. *Erdős P., Rényi A.* On the evolution of random graphs // Publication of the Mathematical Institute of the Hungarian Academy of Sciences. 1960. V. 5. P. 17–61.
3. *Степанов В.Е.* О вероятности связности случайного графа $\mathcal{G}_m(t)$ // Теория вероятн. и ее примен. 1970. Т. 15, № 1. С. 55–67.
4. *Łuczak T., Pittel B., Wierman J.* The structure of a random graph at the point of phase transition // Transactions of the American Mathematical Society. 1994. V. 341. P. 721–748.
5. *Bollobás B.* Random graphs. Cambridge University Press, 2001.
6. *Jansen S., Łuczak T., Rucinski A.* Random graphs. New York : Wiley-Interscience, 2000.
7. *Aldous D.* Brownian excursions, critical random graphs and the multiplicative coalescent // The Annals of Probability. 1997. V. 25. P. 812–854.
8. *Schmidt-Pruzan J., Shamir E.* Component structures in the evolution of random hypergraphs // Combinatorica. 1985. V. 5. P. 81–94.
9. *Behrisch M., Coja-Oghlan A., Kang M.* The order of the giant component of random hypergraphs // Random Structures and Algorithms. 2010. V. 36. P. 149–184.
10. *Bollobás B., Riordan O.* Asymptotic normality of the size of the giant component in a random hypergraph // Random Structures and Algorithms. 2013. V. 41. P. 441–450.

References

1. *Harary F.* Graph theory. transl. from eng. Moscow : URSS, 2018. 304 p.
2. *Erdős P., Rényi A.* On the evolution of random graphs. Publication of the Mathematical Institute of the Hungarian Academy of Sciences. 1960. V. 5. P. 17–61.
3. *Stepanov V.E.* On the Probability of Connectedness of a Random Graph $\mathcal{G}_m(t)$. Theory Probab. Appl. 1970. V. 15. P. 55–67.
4. *Łuczak T., Pittel B., Wierman J.* The structure of a random graph at the point of phase transition. Transactions of the American Mathematical Society. 1994. V. 341. P. 721–748.
5. *Bollobás B.* Random graphs. Cambridge University Press, 2001.
6. *Jansen S., Łuczak T., Rucinski A.* Random graphs. New York : Wiley-Interscience, 2000.
7. *Aldous D.* Brownian excursions, critical random graphs and the multiplicative coalescent. The Annals of Probability. 1997. V. 25. P. 812–854.
8. *Schmidt-Pruzan J., Shamir E.* Component structures in the evolution of random hypergraphs. Combinatorica. 1985. V. 5. P. 81–94.

9. *Behrisch M., Coja-Oghlan A., Kang M.* The order of the giant component of random hypergraphs. Random Structures and Algorithms. 2010. V. 36. P. 149–184.

10. *Bollobás B., Riordan O.* Asymptotic normality of the size of the giant component in a random hypergraph. Random Structures and Algorithms. 2013. V. 41. P. 441–450.

Поступила в редакцию 05.12.2023