ПРОБЛЕМЫ ПЕРЕДАЧИ ИНФОРМАЦИИ

Том 57 **2021** **Вып. 4**

УДК 621.391 : 519.175.4

© 2021 г. **Н.М. Деревянко, М.М. Кошелев**

НОВЫЕ ОЦЕНКИ МОДУЛЯРНОСТИ ГРАФОВ $G(n,r,s)$ И $G_p(n,r,s)$ [1]

Исследуется поведение модулярности графов $G(n,r,s)$ для случая $r = o(\sqrt{n})$ и $n \to \infty$, а также графов $G_p(n,r,s)$ при фиксированных r, s и $n \to \infty$. Для графов $G(n,r,s)$ при $r \geqslant cs^2$ получены существенные улучшения предыдущих верхних оценок. На семейство графов $G_p(n,r,s)$ при $p = p(n) = \omega\big(n^{-\frac{r-s-1}{2}}\big)$ и фиксированных r и s перенесены верхние и нижние оценки, полученные ранее для графов $G(n,r,s)$.

Ключевые слова: модулярность, графы Джонсона, кластеризация, случайные графы.

DOI: 10.31857/S0555292321040082

§ 1. Введение

Модулярность графа – величина, впервые возникшая в работе Ньюмана и Гирвана [1], которая впоследствии оказалась полезной для оценки качества кластеризационных алгоритмов (см. [2–5]). Такие алгоритмы играют важную роль в биологии, физике, социологии и программировании (см. [6]).

Благодаря широкому применению в программировании задача подсчета модулярности получила большое развитие в последние годы. Поскольку вычисление модулярности с помощью компьютера представляется сложной задачей (в [7] было доказано, что задача оценки модулярности является NP-полной), интересным представляется получение оценок для различных классов графов. В последние годы было получено множество оценок для различных классов графов, таких как звезды, гиперкубы, графы, удовлетворяющие степенному закону [8], графы, близкие к полным [9], деревья с небольшой максимальной степенью [10], а также большого числа моделей случайных графов. В частности, были получены результаты для различных вероятностных моделей веб-графов, таких как модели предпочтительного присоединения [11], а также графов $G(n,d)$ (случайные d-регулярные графы) [12] и классической модели Эрдёша – Реньи [13].

Говоря о модели Эрдёша – Реньи, нельзя не упомянуть широкий класс теорем, называемых теоремами о стабильности. В теоремах подобного типа доказывается, что результаты, полученные для определенного графа, можно перенести и на его случайные подграфы. Одним из ярких примеров теорем такого класса является теорема о стабильности числа независимости кнезеровского графа, доказанная в работе [14].

Данная статья посвящена исследованию стабильности модулярности для семейства графов $G(n,r,s)$, которые являются обобщением кнезеровских графов и графов

[1] Работа выполнена за счет гранта Президента Российской Федерации для государственной поддержки ведущих научных школ (номер гранта НШ-2540.2020.1), а также гранта Фонда развития теоретической физики и математики "БАЗИС".

Джонсона. Эти графы нашли свое применение в задачах о кодах с одним запрещенным расстоянием (см. [15]), а также в различных задачах комбинаторной геометрии, таких как гипотезы Борсука и Нелсона–Эрдёша–Хадвигера (см. [16–18]).

Оценки модулярности, полученные ранее для $G(n,r,s)$, удалось перенести на случайные графы $G_p(n,r,s)$, т.е. подграфы $G(n,r,s)$, в которых каждое ребро выбрасывается из графа с вероятностью $1-p$. Помимо этого, удалось получить новые оценки модулярности $G(n,r,s)$, которые также были распространены на случайные подграфы.

Прежде чем перейти к определению основных понятий, введем некоторые обозначения, которыми мы будем пользоваться:

- $e(V)$ – количество ребер, оба конца которых лежат в V;
- $e(U,V)$ – количество ребер, один конец которых лежит в U, а другой – в V;
- V_G – множество всех вершин графа;
- $\deg(v)$ – степень вершины v;
- $e(V,p)$ – случайная величина, равная количеству ребер, оба конца которых лежат в V, для случайного графа $G_p(n,r,s)$, определение которого будет дано ниже;
- $e(U,V,p)$ – случайная величина, равная количеству ребер, один конец которых лежит в U, а другой – в V, для случайного графа $G_p(n,r,s)$;
- $\deg(v,p)$ – случайная величина, равная степени вершины v, для случайного графа $G_p(n,r,s)$.

Настало время дать формальные определения основных объектов исследования.

Определение 1. *Модулярность графа G* – характеристика графа, которая показывает, насколько оптимально можно разбить граф на части, где вершины внутри одной части сильно связаны, а связь между различными частями мала. Модулярность графа G обозначается через $q^*(G)$ и выражается формулой

$$q^*(G) := \max_{\mathcal{A}} \left(\sum_{A\in\mathcal{A}} \frac{e(A)}{e(G)} - \sum_{A\in\mathcal{A}} \frac{\Big(\sum\limits_{v\in A} \deg(v)\Big)^2}{4e^2(G)} \right),$$

где максимум берется по всем разбиениям множества вершин графа

$$\mathcal{A} = \{A_1, A_2, \ldots, A_k\}.$$

Помимо этого, определяется понятие *модулярности разбиения*, а именно

$$q(\mathcal{A}) := \sum_{A\in\mathcal{A}} \frac{e(A)}{e(G)} - \sum_{A\in\mathcal{A}} \frac{\Big(\sum\limits_{v\in A} \deg(v)\Big)^2}{4e^2(G)}.$$

При этом первую сумму $\sum\limits_{A\in\mathcal{A}} \frac{e(A)}{e(G)}$ принято называть *реберным вкладом*, а вторую сумму $\sum\limits_{A\in\mathcal{A}} \frac{\Big(\sum\limits_{v\in A} \deg(v)\Big)^2}{4e^2(G)}$ – *степенным штрафом.*

Замечание 1. Если граф d-регулярный, то сумму степенного штрафа можно записать в более удобном виде:

$$\sum_{A\in\mathcal{A}} \frac{\Big(\sum\limits_{v\in A} \deg(v)\Big)^2}{4e^2(G)} = \sum_{A\in\mathcal{A}} \frac{|A|^2 d^2}{|V|^2 d^2} = \sum_{A\in\mathcal{A}} \frac{|A|^2}{|V|^2}.$$

Определение 2. Пусть $1 \leqslant r \leqslant n$, $0 \leqslant s \leqslant r$, $n, r, s \in \mathbb{N}$, $\Gamma := \{1, 2, \ldots, n\}$, тогда $G(n, r, s) := (V_G(n, r), E(n, r, s))$ является *регулярным графом*, в котором:

$$V_G = V_G(n, r) := \binom{\Gamma}{r} \text{ – все } r\text{-элементные подмножества множества } \Gamma;$$

$$E(n, r, s) := \{(u, v) : \ u, v \in V_G(n, r), \ |u \cap v| = s\}.$$

Множество Γ называется *множеством элементов.*

Замечание 2. Несложно заметить, что для числа вершин и числа ребер в графе $G(n, r, s)$ верно следующее:

$$|V_G| = \binom{n}{r}, \quad |E(n, r, s)| = \frac{1}{2}\binom{n}{r}\binom{r}{s}\binom{n-r}{r-s}.$$

При этом степень каждой вершины равна $\binom{r}{s}\binom{n-r}{r-s}$.

Семейство случайных графов $G_p(n, r, s)$ состоит из всех графов с множеством вершин $V_G(n, r)$ и ребрами из множества $E(n, r, s)$. Каждое ребро этого множества добавляется в граф независимо с вероятностью p. Таким образом, каждый граф с множеством вершин $V_G(n, r)$ и множеством ребер $E \subseteq E(n, r, s)$ может быть выбран с вероятностью $p^{|E|}(1-p)^{|E(n,r,s)|-|E|}$.

Ранее в работах [19–21] были получены следующие оценки на модулярность графов $G(n, r, s)$.

Теорема 1 (см. [19]). *Пусть $r \geqslant 2$ и $1 \leqslant s \leqslant \left[\frac{r}{2}\right]$. Тогда*

$$\limsup_{n\to\infty} q^*(G(n, r, s)) \leqslant 1 - \frac{\binom{[r/2]}{s}}{2\binom{r}{s}}.$$

Теорема 2 (см. [20]). *Справедливо равенство*

$$q^*(G(n, 2, 1)) = \frac{1}{3} + \frac{2w(w-1)(w-2)}{3n(n-1)(n-2)} - \frac{w^2(w-1)^2}{n^2(n-1)^2} - \frac{4n-2}{3n(n-1)} + \\ + \frac{w(w-1)(4w-2)}{3n^2(n-1)^2}$$

при всех $n \geqslant 5$, где $w = \left\lceil\frac{n}{2}\right\rceil + 1$. Предел выражения в правой части при $n \to \infty$ равен $\frac{17}{48}$.

Теорема 3 (см. [21]). *Пусть $r > s \geqslant 1$. Тогда*

$$\liminf_{n\to\infty} q^*(G(n, r, s)) \geqslant \frac{s}{2r-s}\left(1 + \left(\frac{r-s}{r}\right)^{\frac{2r}{s}}\right).$$

Отметим, что верхние и нижние оценки модулярности графов $G(n, r, s)$ на данный момент весьма далеки друг от друга. Так, при $r = 3$, $s = 1$ нижняя оценка равна $\frac{793}{3645} < 0{,}22$, в то время как верхняя оценка составляет всего лишь $1 - \frac{1}{6} > 0{,}83$. Более того, при больших r и s разница между оценками становится все более существенной. Так, например, уже при $s = 10$, $r = 20$ нижняя оценка равна $\frac{17}{48}$, в то время как верхняя оценка – всего лишь $1 - \frac{1}{2\binom{20}{10}} > 0{,}99999$.

Таким образом, единственный случай, когда верхние и нижние оценки модулярности $G(n,r,s)$ близки между собой, – это случай $r=2$, $s=1$, в котором благодаря теореме 2 известно точное значение модулярности. Отметим также, что для таких параметров оценка из теоремы 3 дает нижнюю оценку предела модулярности $\frac{17}{48}$, совпадающую с настоящим пределом модулярности в этом случае, в то время как теорема 1 дает лишь оценку $\frac{3}{4}$.

В § 2 будут получены новые верхние оценки модулярности графов $G(n,r,s)$. Затем § 3 будет посвящен переносу результатов § 2, а также теорем 1 и 3 на графы $G_p(n,r,s)$.

§ 2. Верхняя оценка модулярности графов $G(n,r,s)$

2.1. Формулировка основных результатов. Первая теорема, которую мы докажем, дает верхнюю оценку величины $e(U)$ для всех $U \subset V_G$, размер которых хотя бы в константу раз меньше $|V_G|$.

Теорема 4. *Пусть* $\alpha,\beta \in (0,1)$, $\alpha<\beta^2$, $s \geqslant 1$,

$$r \geqslant -\frac{1}{\ln\left(\dfrac{1-\beta}{1-\alpha/\beta}\right)}s^2+2s-1.$$

Пусть также $U \subseteq V_G(n,r)$, $|U| < \alpha\binom{n}{r}$. *Тогда*

$$e(U) \leqslant \frac{1+\beta-\beta^2}{2(2-\beta)}\binom{r}{s}\binom{n-s}{r-s}|U|.$$

Также сформулируем следствия теоремы 4.

Следствие 1. *В условиях теоремы* 4 *при* $n\to\infty$, $r=o(\sqrt{n})$ *верно неравенство*

$$e(U,\overline{U}) \geqslant (1+o(1))\frac{(1-\beta)^2}{2-\beta}\binom{r}{s}\binom{n-s}{r-s}|U|.$$

Следствие 2. *В условиях теоремы* 4 *при* $n\to\infty$, $r=o(\sqrt{n})$ *верно неравенство*

$$\frac{e(U)}{e(G)} \leqslant (1+o(1))\frac{1+\beta-\beta^2}{2-\beta}\frac{|U|}{|V_G|}.$$

Доказательство следствия 2 мы опустим в силу его тривиальности.

Сформулируем, наконец, основной результат этого параграфа:

Теорема 5. *Пусть* $\varepsilon \in (0,1)$, $s=s(n)\geqslant 1$,

$$r=r(n) \geqslant -\frac{1}{\ln(1-\varepsilon)}s^2+2s-1,$$

$r=o(\sqrt{n})$. *Тогда*

$$\limsup_{n\to\infty} q^*(G(n,r,s)) \leqslant f(\varepsilon),$$

где $f(\varepsilon)=\max\limits_{x\in[0,1]}\left(\dfrac{1+x-x^2}{2-x}-\max\left(\dfrac{x^2-\varepsilon x}{1-\varepsilon},0\right)\right)$.

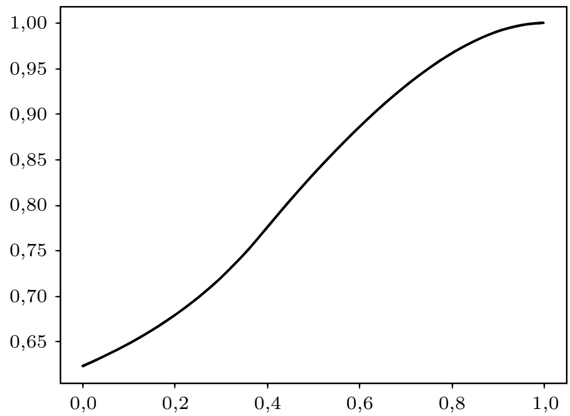

Рис. 1. График функции $f(\varepsilon)$

Таблица 1

ε	$r \geqslant$	$f(\varepsilon)$
0,01	$99{,}4992s^2 + 2s - 1$	0,6246
0,1	$9{,}4912s^2 + 2s - 1$	0,6469
0,2	$4{,}4814s^2 + 2s - 1$	0,6783
0,3	$2{,}8037s^2 + 2s - 1$	0,7195
0,4	$1{,}9576s^2 + 2s - 1$	0,7750
0,5	$1{,}4427s^2 + 2s - 1$	0,8333
0,6	$1{,}0914s^2 + 2s - 1$	0,8857
0,7	$0{,}8306s^2 + 2s - 1$	0,9308
0,8	$0{,}6213s^2 + 2s - 1$	0,9667
0,9	$0{,}4343s^2 + 2s - 1$	0,9909
0,99	$0{,}2171s^2 + 2s - 1$	0,9999

На рис. 1 приведен график, показывающий зависимость функции f от ε.

Приведем также таблицу приближенных значений $f(\varepsilon)$ и ограничений на r, порождаемых соответствующим ε (табл. 1).

Сравним новые результаты с имеющимися. Заметим, что верхняя оценка из теоремы 1 имеет вид

$$1 - \frac{\binom{[r/2]}{s}}{2\binom{r}{s}} \geqslant 1 - 2^{-s-1}.$$

В частности, из этого следует, что при любом $s \geqslant 1$ такая оценка не лучше, чем $\frac{3}{4}$. Тогда, взяв, например, $r \geqslant -\frac{1}{\ln 0,7}s^2 + 2s - 1$, можно увидеть, что при таких параметрах r и s новая оценка, примерно равная 0,7195, существенно лучше предыдущей. При возрастании s разрыв между оценками становится еще более драматическим. Например, при $s = 10$ и $r \geqslant 129$ (для таких r выполняется условие теоремы 5 с $\varepsilon = 0{,}6$) новая теорема дает оценку $f(0{,}6) \approx 0{,}8857$, в то время как теорема 2 не позволяет оценить модулярность даже числом 0,999.

К сожалению, зазор между верхними и нижними оценками модулярности все еще достаточно велик. Нетрудно видеть, что нижняя оценка модулярности зависит лишь от отношения s и r, в то время как в теореме 5 требуется условие

$$r \geqslant Cs^2 + 2s - 1.$$

Таким образом, если мы зафиксируем C, то с ростом r и s нижняя оценка будет стремиться к 0, в то время как верхняя будет оставаться на одном и том же уровне.

2.2. Вспомогательные леммы и определения.

Лемма 1. *Пусть* $\alpha \in (0,1)$, $r \geqslant -\dfrac{1}{\ln \alpha} s^2 + 2s - 1$, $s \geqslant 1$. *Тогда*

$$\binom{r-s}{s} \geqslant \alpha \binom{r}{s}.$$

Доказательство. Перепишем требуемое неравенство в виде

$$\alpha r!\,(r-2s)! \leqslant ((r-s)!)^2 \iff \prod_{k=0}^{s-1} \frac{r-k}{r-s-k} \leqslant \frac{1}{\alpha}.$$

Оценим сверху левую часть требуемого неравенства:

$$\prod_{k=0}^{s-1} \frac{r-k}{r-s-k} = \prod_{k=0}^{s-1} \left(1 + \frac{s}{r-s-k}\right) \leqslant \left(1 + \frac{s}{r-2s+1}\right)^s =$$

$$= \left(1 + \frac{s}{r-2s+1}\right)^{\frac{r-2s+1}{s} \frac{s^2}{r-2s+1}} \leqslant e^{\frac{s^2}{r-2s+1}}.$$

Осталось понять, при каких условиях $e^{\frac{s^2}{r-2s+1}} \leqslant \dfrac{1}{\alpha}$. Это эквивалентно тому, что

$$\frac{s^2}{r-2s+1} \leqslant \ln \frac{1}{\alpha},$$

или

$$r \geqslant -\frac{1}{\ln \alpha} s^2 + 2s - 1. \quad \blacktriangle$$

Определение 3. Пусть V – некоторое множество вершин графа $G(n,r,s)$, а S – некоторое множество s-элементных подмножеств множества $\{1, 2, \ldots, n\}$.

Вкладом вершин V в множество S назовем следующую величину:

$$W(V,S) := \sum_{v \in V} |\{S_j \in S :\ S_j \subset v\}| = \sum_{S_j \in S} |\{v \in V :\ S_j \subset v\}|.$$

Далее для удобства будем использовать следующие обозначения:

- $S_{\text{all}} := \{S_1, S_2, \ldots, S_{\binom{n}{s}}\}$, $|S_{\text{all}}| = \dbinom{n}{s}$, – всевозможные s-элементные множества;
- $\overline{S} := S_{\text{all}} \setminus S$ для $S \subseteq S_{\text{all}}$;
- $V_{S_i} := \{v \in V_G :\ S_i \subset v\}$, $S_i \in S_{\text{all}}$;
- $W_{S_i}(V) := W(V, \{S_i\}) = |\{v \in V :\ S_i \subset v\}|$.

Лемма 2. *Пусть* $\alpha, \beta \in (0,1)$, $\alpha < \beta$, $s \geqslant 1$,

$$r \geqslant -\frac{1}{\ln\left(\dfrac{1-\beta}{1-\alpha}\right)} s^2 + 2s - 1.$$

Пусть также $U \subseteq V_G$, $S \subseteq S_{\mathrm{all}}$, $|S| < \alpha\binom{n}{s}$. *Тогда для любого* $S_i \in S$, *такого что*

$$W_{S_i}(U) = \beta\binom{n-s}{r-s} + d, \quad d \geqslant 0,$$

верно неравенство

$$W(V_{S_i} \cap U, \overline{S}) \geqslant d\binom{r}{s}.$$

Доказательство. Определим $V_{\mathrm{left}} := V_{S_i} \setminus (V_{S_i} \cap U)$. Имеем

$$|V_{\mathrm{left}}| = (1-\beta)\binom{n-s}{r-s} - d.$$

Оценим $W(V_{S_i}, \overline{S})$:

$$W(V_{S_i}, \overline{S}) = \sum_{S_j \in \overline{S}} W_{S_j}(V_{S_i}) \geqslant \sum_{S_j \in \overline{S}} \binom{n-2s}{r-2s} = |\overline{S}|\binom{n-2s}{r-2s} >$$
$$> (1-\alpha)\binom{n}{s}\binom{n-2s}{r-2s} > (1-\alpha)\binom{n-s}{s}\binom{n-2s}{r-2s} = (1-\alpha)\binom{r-s}{s}\binom{n-s}{r-s}.$$

Применяя лемму 1, получаем, что

$$W(V_{S_i}, \overline{S}) > (1-\alpha)\binom{r-s}{s}\binom{n-s}{r-s} \geqslant (1-\beta)\binom{r}{s}\binom{n-s}{r-s}.$$

С другой стороны, имеем

$$W(V_{S_i}, \overline{S}) = W(V_{S_i} \cap U, \overline{S}) + W(V_{\mathrm{left}}, \overline{S}) \leqslant W(V_{S_i} \cap U, \overline{S}) +$$
$$+ (1-\beta)\binom{n-s}{r-s}\binom{r}{s} - d\binom{r}{s} = \left(W(V_{S_i} \cap U, \overline{S}) - d\binom{r}{s}\right) + (1-\beta)\binom{r}{s}\binom{n-s}{r-s},$$

откуда тривиально следует требуемое неравенство. ▲

2.3. Доказательство основных результатов.

Доказательство теоремы 4. Определим $e_{S_i}(U)$ – множество ребер с концами в множестве U, оба конца которых содержат s-элементное множество S_i.

Мы знаем, что $e(U) = \sum_{i=1}^{\binom{n}{s}} e_{S_i}(U)$. Очевидно, что $e_{S_i}(U) \leqslant \dfrac{W_{S_i}^2(U)}{2}$. Отсюда имеем оценку $e(U) \leqslant \sum_{i=1}^{\binom{n}{s}} \dfrac{W_{S_i}^2(U)}{2}$.

Будем считать, что $W_{S_1}(U) \geqslant W_{S_2}(U) \geqslant \ldots$ Пусть m таково, что

$$W_{S_m}(U) \geqslant \beta\binom{n-s}{r-s}, \quad W_{S_{m+1}}(U) < \beta\binom{n-s}{r-s},$$

или $m = \binom{n}{s}$, если первое неравенство верно для всех S_i. Множество $\{S_1, \ldots, S_m\}$ будем обозначать через $S_{\max}$. Оценим m. Очевидно, что $\sum_{i=1}^{\binom{n}{s}} W_{S_i}(U) = \binom{r}{s}|U|$, откуда имеем оценку

$$m\beta\binom{n-s}{r-s} \leqslant \binom{r}{s}|U| \iff m \leqslant \frac{\binom{r}{s}|U|}{\beta\binom{n-s}{r-s}} < \frac{\binom{r}{s}\alpha\binom{n}{r}}{\beta\binom{n-s}{r-s}} = \frac{\alpha}{\beta}\binom{n}{s}.$$

Представим $W_{S_i}(U)$, $i \leqslant m$, в виде

$$W_{S_i}(U) = \beta\binom{n-s}{r-s} + d_i, \quad d_i \geqslant 0.$$

Оценим $W(U, \overline{S_{\max}})$. Применяя лемму 2 с $S = S_{\max}$, $\alpha = \alpha/\beta$, $\beta = \beta$ (и суммируя результат по всем $S_i \in S_{\max}$), получаем неравенство

$$\sum_{S_i \in S_{\max}} W(V_{S_i} \cap U, \overline{S_{\max}}) \geqslant \sum_{S_i \in S_{\max}} d_i\binom{r}{s}.$$

Заметим, что в левой части неравенства каждое вхождение множества из $\overline{S_{\max}}$ в вершину из U посчитано не более $\binom{r}{s}$ раз. Отсюда получаем

$$W(U, \overline{S_{\max}}) \geqslant \frac{1}{\binom{r}{s}} \sum_{S_i \in S_{\max}} d_i\binom{r}{s} = \sum_{S_i \in S_{\max}} d_i.$$

Теперь можно получить оценку на $\sum\limits_{S_i \in S_{\max}} d_i$. Действительно,

$$\binom{r}{s}|U| = W(U, S_{\text{all}}) = \sum_{S_i \in S_{\max}} W_{S_i}(U) + W(U, \overline{S_{\max}}) \geqslant$$

$$\geqslant \sum_{S_i \in S_{\max}} \left(\beta\binom{n-s}{r-s} + d_i\right) + \sum_{S_i \in S_{\max}} d_i,$$

откуда имеем

$$\sum_{S_i \in S_{\max}} d_i \leqslant \frac{1}{2}\left(\binom{r}{s}|U| - \beta\binom{n-s}{r-s}m\right).$$

Получим еще одну оценку на $\sum\limits_{S_i \in S_{\max}} d_i$. Она тривиальна, ибо $d_i \leqslant (1-\beta)\binom{n-s}{r-s}$. Отсюда имеем оценку

$$\sum_{S_i \in S_{\max}} d_i \leqslant (1-\beta)\binom{n-s}{r-s}m.$$

Нетрудно видеть, что при таких ограничениях на $\sum\limits_{S_i \in S_{\max}} d_i$ верна следующая оценка на $\sum\limits_{i=1}^{\binom{n}{s}} \frac{W_{S_i}^2(U)}{2}$ (она вытекает из того, что максимум такой суммы, очевидно, достигается на последовательности W_{S_i}, равной $\binom{n-s}{r-s}, \ldots, \binom{n-s}{r-s}, \beta\binom{n-s}{r-s}$, $\ldots, \beta\binom{n-s}{r-s}, 0, 0, \ldots, 0$):

$$\sum_{i=1}^{\binom{n}{s}} \frac{W_{S_i}^2(U)}{2} \leqslant \frac{\min\left((1-\beta)\binom{n-s}{r-s}m, \frac{1}{2}\left(\binom{r}{s}|U| - \beta\binom{n-s}{r-s}m\right)\right)}{(1-\beta)\binom{n-s}{r-s}} \frac{\binom{n-s}{r-s}^2}{2} +$$

$$+\frac{\binom{r}{s}|U|-\dfrac{\min\left((1-\beta)\binom{n-s}{r-s}m,\frac{1}{2}\left(\binom{r}{s}|U|-\beta\binom{n-s}{r-s}m\right)\right)}{(1-\beta)\binom{n-s}{r-s}}\binom{n-s}{r-s}}{\beta\binom{n-s}{r-s}}\times$$

$$\times\frac{\beta^2\binom{n-s}{r-s}^2}{2}.$$

Прокомментируем эту оценку. Первое слагаемое в ней соответствует членам вида $\binom{n-s}{r-s}$ в оптимальной последовательности, в то время как второе слагаемое соответствует членам вида $\beta\binom{n-s}{r-s}$.

Упростив выражение в правой части, получаем

$$\sum_{i=1}^{\binom{n}{s}}\frac{W^2_{S_i}(U)}{2}\leqslant\min\left(\binom{n-s}{r-s}m,\frac{1}{2(1-\beta)}\left(\binom{r}{s}|U|-\beta\binom{n-s}{r-s}m\right)\right)\frac{\binom{n-s}{r-s}}{2}+$$

$$+\left(\binom{r}{s}|U|-\min\left(\binom{n-s}{r-s}m,\frac{1}{2(1-\beta)}\left(\binom{r}{s}|U|-\beta\binom{n-s}{r-s}m\right)\right)\right)\frac{\beta\binom{n-s}{r-s}}{2}=$$

$$=\frac{\beta}{2}\binom{r}{s}\binom{n-s}{r-s}|U|+\min\left(\binom{n-s}{r-s}m,\frac{1}{2(1-\beta)}\left(\binom{r}{s}|U|-\beta\binom{n-s}{r-s}m\right)\right)\times$$

$$\times\frac{\binom{n-s}{r-s}(1-\beta)}{2}=\frac{\beta}{2}\binom{r}{s}\binom{n-s}{r-s}|U|+$$

$$+\min\left((1-\beta)\binom{n-s}{r-s}m,\frac{1}{2}\left(\binom{r}{s}|U|-\beta\binom{n-s}{r-s}m\right)\right)\frac{\binom{n-s}{r-s}}{2}.$$

Осталось оценить этот минимум. Так как одна из величин монотонно возрастает по m, а другая монотонно убывает по m, то максимум минимума достигается в точке равенства этих величин. Получаем уравнение на m_{opt}:

$$(1-\beta)\binom{n-s}{r-s}m_{\text{opt}}=\frac{1}{2}\left(\binom{r}{s}|U|-\beta\binom{n-s}{r-s}m_{\text{opt}}\right)\iff$$

$$\iff(1-\frac{\beta}{2})\binom{n-s}{r-s}m_{\text{opt}}=\frac{1}{2}\binom{r}{s}|U|\iff m_{\text{opt}}=\frac{1}{2\left(1-\frac{\beta}{2}\right)}\frac{\binom{r}{s}|U|}{\binom{n-s}{r-s}}.$$

Подставляя m_{opt} в оценку для $\sum_{i=1}^{\binom{n}{s}}\frac{W^2_{S_i}(U)}{2}$, получаем

$$\sum_{i=1}^{\binom{n}{s}}\frac{W^2_{S_i}(U)}{2}\leqslant\frac{\beta}{2}\binom{r}{s}\binom{n-s}{r-s}|U|+\frac{1-\beta}{2(2-\beta)}\binom{r}{s}\binom{n-s}{r-s}|U|=$$

$$=\frac{1+\beta-\beta^2}{2(2-\beta)}\binom{r}{s}\binom{n-s}{r-s}|U|.\quad\blacktriangle$$

Д о к а з а т е л ь с т в о с л е д с т в и я 1. Заметим, что при $r = o(\sqrt{n})$ выполняется соотношение $\binom{n-r}{r-s} = (1+o(1))\binom{n-s}{r-s}$. Действительно,

$$1 > \frac{\binom{n-r}{r-s}}{\binom{n-s}{r-s}} = \frac{(n-r)!^2}{(n-s)!(n-2r+s)!} = \frac{(n-r)\dots(n-2r+s+1)}{(n-s)\dots(n-r+1)} \geqslant$$
$$\geqslant \left(\frac{n-2r+s+1}{n-r+1}\right)^{r-s} = \left(1 - \frac{r-s}{n-r+1}\right)^{r-s}.$$

Обозначая $\dfrac{r-s}{n-r+1} = t$, получаем

$$1 > \frac{\binom{n-r}{r-s}}{\binom{n-s}{r-s}} \geqslant (1-t)^{\frac{(r-s)^2}{t(n-r+1)}}.$$

При достаточно больших n имеем $(1-t)^{1/t} > 1/3$, т.е.

$$1 > \frac{\binom{n-r}{r-s}}{\binom{n-s}{r-s}} \geqslant \left(\frac{1}{3}\right)^{\frac{(r-s)^2}{n-r+1}}.$$

Осталось заметить, что при $n \gg r^2$ выполняется соотношение $(1/3)^{\frac{(r-s)^2}{n-r+1}} = 1 + o(1)$.

Следствие теперь тривиально вытекает из равенства

$$2e(U) + e(U,\overline{U}) = \sum_{v\in U} \deg(v) = \binom{r}{s}\binom{n-r}{r-s}|U| = (1+o(1))\binom{r}{s}\binom{n-s}{r-s}|U|.$$

и теоремы 4. ▲

Д о к а з а т е л ь с т в о т е о р е м ы 5. Введем функцию $\alpha(x) = x\dfrac{x-\varepsilon}{1-\varepsilon}$. Проверим, что в условиях теоремы $\alpha(x) < x^2$ всюду на $(0,1)$. Действительно,

$$x\frac{x-\varepsilon}{1-\varepsilon} = \frac{x^2}{1-\varepsilon} - \frac{\varepsilon x}{1-\varepsilon} = x^2 - \frac{\varepsilon(x-x^2)}{1-\varepsilon}.$$

Так как второе слагаемое положительно, то оценка очевидна. Заметим также, что

$$\frac{1-x}{1-\frac{\alpha(x)}{x}} = \frac{1-x}{1-\frac{x-\varepsilon}{1-\varepsilon}} = 1-\varepsilon.$$

Зафиксируем $k \in \mathbb{N}$ и введем обозначения: $\beta_i = \dfrac{i}{k}$, $i \in \{0,\dots,k\}$, $\alpha_i = \alpha(\beta_i)$. Рассмотрим оптимальное разбиение $\mathcal{A} = \{A_1,\dots,A_\ell\}$, $\ell > 1$, множества вершин графа $G(n,r,s)$ на части. Представим наше разбиение в виде объединения множеств C_i, $i \in \{1,\dots,k\}$, таких что $A_j \in C_i \Leftrightarrow \dfrac{|A_j|}{|V_G|} \in [\alpha_{i-1},\alpha_i)$. Стоит отметить, что некоторые интервалы могут быть некорректными, т.е. такими, что $\alpha_{i-1} > \alpha_i$. В таком случае мы полагаем этот интервал равным пустому множеству. Нетрудно видеть, что любое число из полуинтервала $[0,1)$ попадет ровно в один отрезок. Тогда модулярность можно переписать в следующем виде:

$$q(\mathcal{A}) = \sum_{A\in\mathcal{A}} \left(\frac{e(A)}{e(G)} - \frac{|A|^2}{|V_G|^2}\right) = \sum_{i=1}^{k} \sum_{A\in C_i} \left(\frac{e(A)}{e(G)} - \frac{|A|^2}{|V_G|^2}\right).$$

Пусть m – минимальное число, для которого $\alpha_m > 0$. Тогда $C_1, \ldots, C_{m-1}$ будут пустыми, поэтому нижний индекс суммирования положим равным m. Предположим теперь, что C_k не пусто. В этом случае имеем

$$q(\mathcal{A}) \leqslant 1 - \sum_{A\in\mathcal{A}} \frac{|A|^2}{|V_G|^2} \leqslant 1 - \alpha^2 \left(1 - \frac{1}{k}\right).$$

Поскольку $\alpha(1 - 1/k) \to 1$ при $k \to \infty$, при достаточно больших k необходимая оценка тривиальна. Таким образом, можно считать, что C_k пусто, а суммирование ведется до $k - 1$. Получим

$$q(\mathcal{A}) = \sum_{i=m}^{k-1} \sum_{A\in C_i} \left(\frac{e(A)}{e(G)} - \frac{|A|^2}{|V_G|^2}\right).$$

Теперь для каждого элемента A из $C_i, i < k$, воспользуемся следствием 2 с $\alpha = \alpha_i$, $\beta = \beta_i$. Также воспользуемся тем, что для любого A из C_i верно неравенство $\dfrac{|A|}{|G|} \geqslant \max(\alpha_{i-1}, 0)$. Оценка получится такой:

$$q(\mathcal{A}) \leqslant \sum_{i=m}^{k-1} \sum_{A\in C_i} \left((1 + o(1))\frac{1 + \beta_i - \beta_i^2}{2 - \beta_i} \frac{|A|}{|V_G|} - \max(\alpha_{i-1}, 0)\frac{|A|}{|V_G|}\right),$$

или, что тоже самое,

$$q(\mathcal{A}) \leqslant \sum_{i=m}^{k-1} \left((1 + o(1))\frac{1 + \beta_i - \beta_i^2}{2 - \beta_i} - \max(\alpha_{i-1}, 0)\right) \frac{\sum\limits_{A\in C_i} |A|}{|V_G|}.$$

Отсюда

$$q(\mathcal{A}) \leqslant \max_{i\in m,\ldots,k-1} \left((1 + o(1))\frac{1 + \beta_i - \beta_i^2}{2 - \beta_i} - \max(\alpha_{i-1}, 0)\right) \sum_{i=m}^{k-1} \frac{\sum\limits_{A\in C_i} |A|}{|V_G|} =$$
$$= \max_{i\in\{m,\ldots,k-1\}} \left(\frac{1 + \beta_i - \beta_i^2}{2 - \beta_i} - \max(\alpha_{i-1}, 0)\right) + o(1) \leqslant$$
$$\leqslant \max_{i\in\{1,\ldots,k\}} \left(\frac{1 + \beta_i - \beta_i^2}{2 - \beta_i} - \max(\alpha_{i-1}, 0)\right) + o(1).$$

Осталось оценить этот максимум. Действительно,

$$\max_{i\in\{1,\ldots,k\}} \left(\frac{1 + \beta_i - \beta_i^2}{2 - \beta_i} - \max(\alpha_{i-1}, 0)\right) =$$
$$= \max_{i\in\{1,\ldots,k\}} \left(\frac{1 + \frac{i}{k} - \left(\frac{i}{k}\right)^2}{2 - \frac{i}{k}} - \max\left(\left(\frac{i}{k} - \frac{1}{k}\right) \frac{\frac{i}{k} - \frac{1}{k} - \varepsilon}{1 - \varepsilon}, 0\right)\right) =$$
$$= \max_{i\in\{1,\ldots,k\}} \left(\frac{1 + \frac{i}{k} - \left(\frac{i}{k}\right)^2}{2 - \frac{i}{k}} - \max\left(\frac{i}{k}\frac{\frac{i}{k} - \varepsilon}{1 - \varepsilon} - \frac{i}{k}\frac{\frac{1}{k}}{1 - \varepsilon} - \frac{1}{k}\frac{\frac{i}{k} - \frac{1}{k} - \varepsilon}{1 - \varepsilon}, 0\right)\right) \leqslant$$

$$\leqslant \max_{i\in\{1,\dots,k\}} \left(\frac{1+\frac{i}{k}-\left(\frac{i}{k}\right)^2}{2-\frac{i}{k}} - \max\left(\frac{i}{k}\frac{\frac{i}{k}-\varepsilon}{1-\varepsilon}, 0\right) + \frac{1}{k}\frac{1}{1-\varepsilon} + \frac{1}{k} \right) =$$

$$= \max_{i\in\{1,\dots,k\}} \left(\frac{1+\frac{i}{k}-\left(\frac{i}{k}\right)^2}{2-\frac{i}{k}} - \max\left(\frac{i}{k}\frac{\frac{i}{k}-\varepsilon}{1-\varepsilon}, 0\right)\right) + \frac{1}{k}\left(\frac{1}{1-\varepsilon} + 1\right) \leqslant$$

$$\leqslant \max_{x\in[0,1]} \left(\frac{1+x-x^2}{2-x} - \max\left(\frac{x^2-\varepsilon x}{1-\varepsilon}, 0\right)\right) + \frac{1}{k}\left(\frac{1}{1-\varepsilon} + 1\right).$$

Устремляя k к бесконечности, получаем утверждение теоремы. ▲

§ 3. Оценки модулярности графов $G_p(n, r, s)$

3.1. Формулировка основных результатов. Введем формальное определение случайных графов $G_p(n,r,s)$. Множество вершин такого графа совпадает с множеством вершин графа $G(n,r,s)$, а множество ребер есть случайное подмножество $E_p(n,r,s)$ множества ребер графа $G(n,r,s)$, удовлетворяющее следующим свойствам:

1. Для любого e из $E(n,r,s)$ верно $\mathbf{P}(e \in E_p(n,r,s)) = p$;
2. Все индикаторы $\mathbb{1}\{e \in E_p(n,r,s)\}$ независимы в совокупности.

В данном параграфе будем считать, что графы $G_p(n,r,s)$ при всех n определены на одном вероятностном пространстве. Таким образом, элементарным исходом в таком вероятностном пространстве будет последовательность $\{H_i\}_{i=r}^{\infty}$, $H_i \sim G_p(i,r,s)$, причем все H_i независимы в совокупности.

На случай графов $G_p(n,r,s)$ удалось перенести теоремы 5 и 1.

Теорема 6. *Пусть r, s – фиксированные целые числа, для которых выполняется равенство*

$$r \geqslant -\frac{1}{\ln(1-\varepsilon)}s^2 + 2s - 1, \quad \textit{где} \quad \varepsilon \in (0,1).$$

Пусть также $p = p(n) = \omega\bigl(n^{-\frac{r-s-1}{2}}\bigr)$. Тогда почти наверное

$$\limsup_{n\to\infty} q^*(G_p(n,r,s)) \leqslant f(\varepsilon),$$

где $f(\varepsilon) = \max\limits_{x\in[0,1]} \left(\frac{1+x-x^2}{2-x} - \max\left(\frac{x^2-\varepsilon x}{1-\varepsilon}, 0\right)\right)$.

Теорема 7. *Пусть s и $r \geqslant 2s$ – фиксированные целые числа, $p = p(n) = \omega\bigl(n^{-\frac{r-s-1}{2}}\bigr)$. Тогда почти наверное*

$$\limsup_{n\to\infty} q^*(G_p(n,r,s)) \leqslant 1 - \frac{\binom{[r/2]}{s}}{2\binom{r}{s}}.$$

Также удалось перенести теорему 3, в которой, однако, пришлось ввести дополнительное ограничение на r.

Теорема 8. *Пусть $s \geqslant 1$ и $r \geqslant 2s$ – фиксированные целые числа, $p = p(n) = \omega\bigl(n^{-\frac{r-s-1}{2}}\bigr)$. Тогда почти наверное*

$$\liminf_{n\to\infty} q^*(G_p(n,r,s)) \geqslant \frac{s}{2r-s}\left(1 + \left(\frac{r-s}{r}\right)^{\frac{2r}{s}}\right).$$

3.2. Вспомогательные леммы. Для доказательства теорем 6–8 будем использовать следующее классическое неравенство, доказанное в [22].

Лемма 3 (неравенство Хёфдинга). *Пусть* $X_1, \ldots, X_n$ – *независимые случайные величины, причем при всех* i *существует пара чисел* a_i, b_i, *для которой* $\mathbf{P}(X_i \in [a_i, b_i]) = 1$. *Тогда для случайной величины* $S_n = X_1 + \ldots + X_n$ *выполняется следующее неравенство*:

$$\mathbf{P}(|S_n - \mathbf{E}S_n| \geqslant \varepsilon n) < 2\exp\left(-\frac{2\varepsilon^2 n^2}{\sum\limits_{i=1}^{n}(b_i - a_i)^2}\right).$$

Помимо этого, нам понадобится следующее тривиальное утверждение.

Предложение 1. *Пусть* Q_n – *последовательность событий, такая что* $\mathbf{P}(\overline{Q_n}) \leqslant e^{-Cn}$ *для некоторого* $C > 0$. *Тогда почти наверное найдется такое* N, *что все события* Q_n, $n > N$, *реализовались.*

Доказательство. Из условия теоремы следует, что

$$\sum_{i=1}^{\infty} \mathbf{P}\left(\overline{Q_n}\right) < +\infty.$$

Тогда по лемме Бореля – Кантелли почти наверное реализуется лишь конечное число событий $\overline{Q_n}$, откуда утверждение очевидно. ▲

Докажем с помощью этого оценку на число ребер в графах $G_p(n, r, s)$. Далее для краткости будем обозначать $G_n = G(n, r, s)$.

Лемма 4. *Пусть* $p = p(n) = f(n)n^{-\frac{r-s-1}{2}}$, *где* $f(n) \to \infty$. *Тогда почти наверное выполняется свойство*

$$\exists N : \forall n > N : |e(G_n, p) - pe(G_n)| < \frac{p \log f(n)}{f(n)} e(G_n).$$

Доказательство. Зафиксируем сначала достаточно большое $n \in \mathbb{N}$. По лемме 3 для $\varepsilon = \dfrac{p \log f(n)}{f(n)}$, $a_i = 0$, $b_i = 1$ имеем

$$\mathbf{P}(|e(G_n, p) - pe(G_n)| \geqslant \varepsilon e(G_n)) \leqslant 2e^{-2\varepsilon^2 e(G_n)} \leqslant e^{-Cn}.$$

Отметим, что имеет место гораздо более сильная оценка на вероятность

$$\mathbf{P}(|e(G_n, p) - pe(G_n)| \geqslant \varepsilon e(G_n)) \leqslant 2e^{-2\varepsilon^2 e(G_n)} \leqslant e^{-Cn^{r+1}\log^2 f(n)},$$

однако для доказательства требуемого утверждения достаточно оценки e^{-Cn}. Таким образом, при достаточно большом n имеем оценку

$$\mathbf{P}\left(|e(G_n, p) - pe(G_n)| \geqslant \frac{p \log f(n)}{f(n)} e(G_n)\right) \leqslant e^{-Cn}.$$

Для завершения доказательства осталось применить предложение 1. ▲

Отметим, что в дальнейшем мы будем использовать обозначения C_i (или C без индекса) для положительных констант.

Для доказательства теоремы 6 нам понадобятся следующие леммы.

Л е м м а 5. *Пусть r, s удовлетворяют неравенству*

$$r \geqslant -\frac{1}{\ln\left(\frac{1-\beta}{1-\frac{\alpha}{\beta}}\right)} s^2 + 2s - 1$$

при некоторых $\alpha, \beta \in (0,1)$, $\alpha < \beta^2$. Пусть также $p = p(n) = f(n)n^{-\frac{r-s-1}{2}}$, где $f(n) \to \infty$, $c > 0$. Тогда почти наверное выполняется свойство

$$\exists N : \forall n > N, \forall V \subset V_G, \binom{[cn]}{r-1} \leqslant |V| \leqslant \alpha \binom{n}{r} :$$

1. $\left|e(V,\overline{V},p) - pe(V,\overline{V})\right| < \dfrac{p \log f(n)}{f(n)} e(V,\overline{V});$

2. $\left|\sum\limits_{v \in V} \deg(v,p) - p \sum\limits_{v \in V} \deg(v)\right| < \dfrac{p \log f(n)}{f(n)} d_n |V|,$

где $d_n = \binom{r}{s}\binom{n-r}{r-s}$ – степень вершин $G(n,r,s)$.

Д о к а з а т е л ь с т в о. Зафиксируем сначала достаточно большое $n \in \mathbb{N}$. Воспользуемся леммой 3 для $\varepsilon = \dfrac{p \log f(n)}{f(n)}$, $a_i = 0$, $b_i = 1$, а также следствием 1:

$$\mathbf{P}\left(|e(V,\overline{V},p) - pe(V,\overline{V})| \geqslant \frac{p \log f(n)}{f(n)} e(V,\overline{V})\right) \leqslant 2e^{-2\varepsilon^2 e(V,\overline{V})} \leqslant$$
$$\leqslant 2\exp\left(-C_0 n^{-r+s+1} \log^2 f(n) n^{r-s} n^{r-1}\right) \leqslant 2\exp\left(-C_0 n^r \log^2 f(n)\right).$$

Теперь получим аналогичную оценку для утверждения 2. Заметим, что

$$\sum_{v \in V} \deg(v,p) \sim \mathrm{Bin}(e(V,\overline{V}),p) + 2\,\mathrm{Bin}(e(V),p),$$

где биномиальные слагаемые в правой части независимы. Тогда можно оценить эту сумму с помощью неравенства Хёфдинга для суммы $e(V,\overline{V}) + e(V)$ независимых слагаемых, где первые $e(V,\overline{V})$ слагаемых распределены как $\mathrm{Bern}(p)$, а остальные – как $2\,\mathrm{Bern}(p)$, $a_i = 0$, $b_i = 2$. Тогда получаем, что

$$\mathbf{P}\left(\left|\sum_{v \in V} \deg(v,p) - p\sum_{v \in V} \deg(v)\right| \geqslant \varepsilon d_n |V|\right) =$$
$$= \mathbf{P}\left(\left|\sum_{v \in V} \deg(v,p) - p\sum_{v \in V} \deg(v)\right| \geqslant \varepsilon(e(V,\overline{V}) + e(V))\right) \leqslant$$
$$\leqslant 2e^{-\frac{1}{2}\varepsilon^2(e(V,\overline{V})+e(V))} \leqslant 2e^{-\frac{1}{4}\varepsilon^2 d_n |V|}.$$

Подставляя $\varepsilon = \dfrac{p \log f(n)}{f(n)}$, получаем оценку

$$\mathbf{P}\left(\left|\sum_{v \in V} \deg(v,p) - p\sum_{v \in V} \deg(v)\right| \geqslant \varepsilon d_n |V|\right) \leqslant 2\exp\left(-C_1 n^r \log^2 f(n)\right).$$

Оценим вероятность того, что хотя бы для одного из V, удовлетворяющему условию леммы, нарушилось какое-либо из неравенств. Для этого обозначим индикатор события, соответствующего нарушению хотя бы одного из неравенств для некоторого множества V, через $\mathbb{1}_V$. Из предыдущих оценок очевидно, что

$$\mathbf{P}(\mathbb{1}_V = 1) \leqslant 4e^{-C_2 n^r \log^2 f(n)}.$$

Имеем

$$\mathbf{P}\left(\exists V : \binom{[cn]}{r-1} \leqslant |V| \leqslant \alpha\binom{n}{r},\ \mathbb{1}_V = 1\right)$$
$$\leqslant \sum_{\binom{cn}{r-1} \leqslant |V| \leqslant \alpha\binom{n}{r}} \binom{\binom{n}{r}}{|V|} 4e^{-C_2 n^r \log^2 f(n)} \leqslant \binom{n}{r}\binom{\binom{n}{r}}{\frac{1}{2}\binom{n}{r}} 4e^{-C_2 n^r \log^2 f(n)}.$$

Воспользуемся неравенством $\binom{2a}{a} < 2^{2a}$ для дальнейшей оценки вероятности:

$$\binom{n}{r}\binom{\binom{n}{r}}{\frac{1}{2}\binom{n}{r}} 4e^{-C_2 n^r \log^2 f(n)} < 2^n 2^{\binom{n}{r}} 4e^{-C_2 n^r \log^2 f(n)} \leqslant$$
$$\leqslant \exp\left(n \ln 2 + \binom{n}{r} \ln 2 + 2\ln 2 - C_2 n^r \log^2 f(n)\right) \leqslant e^{-C_3 n^r \log^2 f(n)} \leqslant e^{-Cn}.$$

Для завершения доказательства осталось применить предложение 1. ▲

Л е м м а 6. *Пусть r, s удовлетворяют неравенству*

$$r \geqslant -\frac{1}{\ln\left(\dfrac{1-\beta}{1-\alpha/\beta}\right)} s^2 + 2s - 1$$

при некоторых $\alpha, \beta \in (0,1)$, $\alpha < \beta^2$. Пусть также $p = p(n) = f(n) n^{-\frac{r-s-1}{2}}$, где $f(n) \to \infty$, $c > 0$. Тогда почти наверное выполняется свойство

$$\exists N : \forall n > N,\ \forall V \subset V_G,\ \binom{[cn]}{r-1} \leqslant |V| \leqslant \alpha\binom{n}{r} :$$

1. $\dfrac{e(V, \overline{V}, p)}{e(G_n, p)} = (1 + o(1)) \dfrac{e(V, \overline{V})}{e(G_n)}$;

2. $\dfrac{\left(\sum\limits_{v \in V} \deg(v, p)\right)^2}{4e^2(G_n, p)} = (1 + o(1)) \dfrac{|V|^2}{|V_G|^2}$.

Д о к а з а т е л ь с т в о. Докажем сначала утверждение 1. Воспользовавшись леммами 5 и 4, оценим $\dfrac{e(V, \overline{V}, p)}{e(G_n, p)}$ сверху и снизу:

$$\frac{e(V, \overline{V})\left(p - \dfrac{p \log f(n)}{f(n)}\right)}{e(G_n)\left(p + \dfrac{p \log f(n)}{f(n)}\right)} \leqslant \frac{e(V, \overline{V}, p)}{e(G_n, p)} \leqslant \frac{e(V, \overline{V})\left(p + \dfrac{p \log f(n)}{f(n)}\right)}{e(G_n)\left(p - \dfrac{p \log f(n)}{f(n)}\right)},$$

откуда

$$\frac{e(V, \overline{V}, p)}{e(G_n, p)} = (1 + o(1)) \frac{e(V, \overline{V})}{e(G_n)}.$$

Аналогично для утверждения 2 выполняется цепочка неравенств

$$\frac{\Big(\sum\limits_{v\in V}\deg(v)\Big)^2\Big(p-\dfrac{p\log f(n)}{f(n)}\Big)^2}{4e^2(G_n)\Big(p+\dfrac{p\log f(n)}{f(n)}\Big)^2}\leqslant\frac{\Big(\sum\limits_{v\in V}\deg(v,p)\Big)^2}{4e^2(G_n,p)}\leqslant$$

$$\leqslant\frac{\Big(\sum\limits_{v\in V}\deg(v)\Big)^2\Big(p+\dfrac{p\log f(n)}{f(n)}\Big)^2}{4e^2(G_n)\Big(p-\dfrac{p\log f(n)}{f(n)}\Big)^2},$$

откуда

$$\frac{\Big(\sum\limits_{v\in V}\deg(v,p)\Big)^2}{4e^2(G_n,p)}=(1+o(1))\frac{\Big(\sum\limits_{v\in V}\deg(v)\Big)^2}{4e^2(G_n)}. \quad \blacktriangle$$

Для доказательства теоремы 7 нам понадобятся аналогичные леммы. Однако прежде чем формулировать и доказывать эти леммы, сформулируем аналог следствия 1, доказанный в работе [19].

Лемма 7. *Пусть* $k\geqslant 2$, $k\in\mathbb{N}$, $s\leqslant\left\lceil\frac{r}{2}\right\rceil$, $r,s\in\mathbb{N}$, $\varepsilon\in(0,1)$. *Тогда существует* $N:=N(\varepsilon,k)$, *такое что для любого* $n>N$ *и любого множества вершин* $V\in\in V_G(n,r)$, *такого что*

$$|V|\in\left[\frac{1}{k}\binom{n}{r},\frac{k-1}{k}\binom{n}{r}\right],$$

верно неравенство

$$e(V,\overline{V})\geqslant\frac{\binom{[r/2]}{s}\binom{n-2r}{r-s}|V||\overline{V}|(1-\varepsilon)}{2\binom{n}{r}}.$$

Здесь $\overline{V}$ – *дополнение* V *до вершин графа* $G(n,r,s)$.

Теперь мы готовы сформулировать необходимые леммы.

Лемма 8. *Пусть* $r\geqslant 2s$, $p=p(n)=f(n)n^{-\frac{r-s-1}{2}}$, *где* $f(n)\to\infty$, $k\in\mathbb{N}$. *Тогда почти наверное выполняется свойство*

$$\exists N:\ \forall n>N,\ \forall V\subset V_G,\ \frac{1}{k}\binom{n}{r}\leqslant|V|\leqslant\frac{k-1}{k}\binom{n}{r}:$$

1. $\big|e(V,\overline{V},p)-pe(V,\overline{V})\big|<\dfrac{p\log f(n)}{f(n)}e(V,\overline{V});$

2. $\left|\sum\limits_{v\in V}\deg(v,p)-pd_n|V|\right|<\dfrac{p\log f(n)}{f(n)}d_n|V|.$

Доказательство. Зафиксируем сначала достаточно большое $n\in\mathbb{N}$. Воспользуемся леммой 3 для $\varepsilon=\dfrac{p\log f(n)}{f(n)}$, $a_i=0$, $b_i=1$, а также леммой 7:

$$\mathbf{P}\left(|e(V,\overline{V},p)-pe(V,\overline{V})|\geqslant\frac{p\log f(n)}{f(n)}e(V,\overline{V})\right)\leqslant 2e^{-2\varepsilon^2e(V,\overline{V})}\leqslant$$

$$\leqslant 2\exp\left(-C_0 n^{-r+s+1}\log^2 f(n) n^{r-s} n^r\right) \leqslant 2\exp\left(-C_0 n^{r+1}\log^2 f(n)\right).$$

Теперь получим аналогичную оценку для утверждения 2. Заметим, что

$$\sum_{v\in V} \deg(v,p) \sim \mathrm{Bin}(e(V,\overline{V}),p) + 2\,\mathrm{Bin}(e(V),p),$$

где биномиальные слагаемые в правой части независимы. Тогда мы можем оценить эту сумму с помощью неравенства Хёфдинга для суммы $e(V,\overline{V})+e(V)$ независимых слагаемых, где первые $e(V,\overline{V})$ слагаемых распределены как $\mathrm{Bern}(p)$, а остальные – как $2\,\mathrm{Bern}(p)$, $a_i = 0$, $b_i = 2$. Тогда получаем, что

$$\mathbf{P}\left(\left|\sum_{v\in V}\deg(v,p) - p\sum_{v\in V}\deg(v)\right| \geqslant \varepsilon d_n |V|\right) =$$
$$= \mathbf{P}\left(\left|\sum_{v\in V}\deg(v,p) - p\sum_{v\in V}\deg(v)\right| \geqslant \varepsilon(e(V,\overline{V})+e(V))\right) \leqslant$$
$$\leqslant 2e^{-\frac{1}{2}\varepsilon^2(e(V,\overline{V})+e(V))} \leqslant 2e^{-\frac{1}{4}\varepsilon^2 d_n|V|}.$$

Подставляя $\varepsilon = \dfrac{p\log f(n)}{f(n)}$, получаем оценку

$$\mathbf{P}\left(\left|\sum_{v\in V}\deg(v,p) - p\sum_{v\in V}\deg(v)\right| \geqslant \varepsilon d_n |V|\right) \leqslant 2\exp\left(-C_1 n^r \log^2 f(n)\right).$$

Оценим вероятность того, что хотя бы для одного из V, удовлетворяющих условию леммы, нарушилось какое-либо из неравенств. Для этого обозначим индикатор события, соответствующего нарушению хотя бы одного из неравенств для некоторого множества V, через $\mathbb{1}_V$. Из предыдущих оценок очевидно, что

$$\mathbf{P}(\mathbb{1}_V = 1) \leqslant 4e^{-C_2 n^r \log^2 f(n)}.$$

Имеем

$$\mathbf{P}\left(\exists V:\ \binom{[cn]}{r-1} \leqslant |V| \leqslant \alpha\binom{n}{r},\ \mathbb{1}_V = 1\right)$$
$$\leqslant \sum_{\frac{1}{k}\binom{n}{r}\leqslant|V|\leqslant\frac{k-1}{k}\binom{n}{r}} \binom{\binom{n}{r}}{|V|} 4e^{-C_2 n^r\log^2 f(n)} \leqslant \binom{n}{r}\binom{\binom{n}{r}}{\frac{1}{2}\binom{n}{r}} 4e^{-C_2 n^r\log^2 f(n)}.$$

Воспользуемся неравенством $\binom{2a}{a} < 2^{2a}$ для дальнейшей оценки вероятности:

$$\binom{n}{r}\binom{\binom{n}{r}}{\frac{1}{2}\binom{n}{r}} 4e^{-C_2 n^r\log^2 f(n)} < 2^n 2^{\binom{n}{r}} 4e^{-C_2 n^r\log^2 f(n)} \leqslant$$
$$\leqslant \exp\left(n\ln 2 + \binom{n}{r}\ln 2 + 2\ln 2 - C_2 n^r \log^2 f(n)\right) \leqslant e^{-C_3 n^r\log^2 f(n)} \leqslant e^{-Cn}.$$

Для завершения доказательства осталось применить предложение 1. ▲

Лемма 9. *Пусть $r \geqslant 2s$, $p = p(n) = f(n)n^{-\frac{r-s-1}{2}}$, где $f(n)\to\infty$, $k\in\mathbb{N}$. Тогда почти наверное выполняется свойство*

$$\exists N:\ \forall n > N,\ \forall V\subset V_G,\ \frac{1}{k}\binom{n}{r} \leqslant |V| \leqslant \frac{k-1}{k}\binom{n}{r}:$$

1. $\dfrac{e(V,\overline{V},p)}{e(G_n,p)} = (1+o(1))\dfrac{e(V,\overline{V})}{e(G_n)}$;

2. $\dfrac{\Bigl(\sum\limits_{v\in V}\deg(v,p)\Bigr)^2}{4e^2(G_n,p)} = (1+o(1))\dfrac{|V|^2}{|V_G|^2}$.

Доказательство дословно совпадает с доказательством леммы 6. ▲

Также нам понадобится определить построение из $\mathcal{A}$ множеств $\mathcal{A}'$ и $\mathcal{A}_{\text{big}}$, которое использовалось в работе [19].

Конструкция 1. Пусть $k \geqslant 2$, $k \in \mathbb{N}$. Определим алгоритм построения множеств $\mathcal{A}'$ и $\mathcal{A}_{\text{big}}$ по разбиению $\mathcal{A} = \{A_1, A_2, \ldots, A_m\}$. Нам потребуется вспомогательное множество A_{merged}, исходно оно пусто.

Проходим по всем $i \in \{1, 2, \ldots, m\}$:

- Если $|A_i| > \frac{1}{k}\binom{n}{r}$, то добавляем A_i в A';
- Если $|A_i| \leqslant \frac{1}{k}\binom{n}{r}$:
  - Если $|A_i| + |A_{\text{merged}}| < \frac{1}{k}\binom{n}{r}$, то $A_{\text{merged}} := A_{\text{merged}} \cup A_i$;
  - Если $|A_i| + |A_{\text{merged}}| \geqslant \frac{1}{k}\binom{n}{r}$, то текущее A_{merged} добавляется в $\mathcal{A}'$ и $A_{\text{merged}} := A_i$.

После итерирования по всем $A_i \in \mathcal{A}$, смотрим на множество A_{merged}:

- Если $|A_{\text{merged}}| > \frac{1}{k}\binom{n}{r}$, то A_{merged} добавляется в $\mathcal{A}'$;
- Если $|A_{\text{merged}}| \leqslant \frac{1}{k}\binom{n}{r}$, то A_{merged} игнорируется.

Множество $\mathcal{A}_{\text{big}}$ легко определить по $\mathcal{A}'$:

$$\mathcal{A}_{\text{big}} := \left\{A :\ A \in \mathcal{A}',\ |A| > \frac{2}{k}\binom{n}{r}\right\}.$$

Приведем часть доказательства теоремы 1 из работы [19] в виде леммы.

Лемма 10. *Пусть r, s таковы, что $r \geqslant \left\lfloor \frac{s}{2} \right\rfloor$. Пусть также $k \geqslant 2$, $k \in \mathbb{N}$, и $\varepsilon > 0$. Возьмем разбиение $\mathcal{A}$ и построим по нему множества $\mathcal{A}'$ и $\mathcal{A}_{\text{big}}$ с помощью конструкции* 1. *Тогда*

$$\frac{1}{2}\sum_{A\in\mathcal{A}'}\frac{e(V_A,\overline{V_A})}{e(G)} + \sum_{A\in\mathcal{A}_{\text{big}}}\left(\frac{|V_A|}{\binom{n}{r}}\right)^2 \geqslant \frac{\binom{\lceil r/2\rceil}{s}}{2\binom{r}{s}}(1-\varepsilon)\left(1-\frac{1}{k}\right) - \frac{1}{k}.$$

Наконец, для доказательства теоремы 8 нам понадобится описать разбиение, оценка модулярности которого и послужила основой доказательства теоремы 3.

Конструкция 2. Рассмотрим следующее разбиение графа $G(n, r, s)$:

$$A_i := \{v \in G(n,r,s) :\ i \in v,\ \forall j > i :\ j \notin v\}$$

для всех i из $\{r, r+1, \ldots, n\}$. Нетрудно видеть, что $|A_i| = \binom{i-1}{r-1}$, а

$$e(A_i) = \frac{1}{2}\binom{i-1}{r-1}\binom{r-1}{s-1}\binom{i-r}{r-s}.$$

Тогда желаемая конструкция имеет вид

$$\mathcal{A} = \left\{ A_n, \ldots, A_{[cn]+1}, \bigcup_{i=r}^{[cn]} A_i \right\}, \quad c = \left(\frac{r-s}{r}\right)^{\frac{1}{s}}.$$

Помимо этого, в доказательстве теоремы 8 будет использована лемма 6.

3.3. Доказательство теорем 6 и 7.

Доказательство теоремы 6. В доказательстве мы можем рассматривать только достаточно большие n, поэтому без ограничения общности считаем, что

$$2\binom{n}{r-1} > \alpha\binom{n}{r}.$$

Рассмотрим функцию

$$\alpha(x) = x\frac{x-\varepsilon}{1-\varepsilon} < x^2.$$

В силу непрерывности этой функции, существует такое β, что $\alpha(\beta) \in (0,1)$, где $1-\alpha^2(\beta) < f(\varepsilon)$.

Отметим, что при данных α, β выполняется соотношение

$$r \geqslant -\frac{1}{\ln(1-\varepsilon)}s^2 + 2s - 1 = -\frac{1}{\ln\left(\dfrac{1-\beta}{1-\alpha/\beta}\right)}s^2 + 2s - 1,$$

поэтому мы можем использовать лемму 6 с данными α и β.

В силу монотонности степенного штрафа по множеству мы знаем, что для почти всех последовательностей $G_n^p \in G_p(n,r,s)$ разбиения, содержащие множества размера хотя бы $\alpha\binom{n}{r} - 2\binom{n}{r-1}$, обладают модулярностью

$$q^*(G_n^p) < 1 - \frac{\alpha^2\binom{n}{r}^2}{\binom{n}{r}^2}(1+o(1)) = 1-\alpha^2(1+o(1)),$$

не превосходящей требуемой оценки.

Таким образом, без ограничения общности можно считать, что оптимальное разбиение всех графов $G \in G_p(n,r,s)$ содержит только множества с размером меньше чем $\alpha\binom{n}{r} - 2\binom{n}{r-1}$.

Рассмотрим оптимальное разбиение $\mathcal{A}$ графа $G_n^p \in G_p(n,r,s)$. Опишем процесс преобразования $\mathcal{A}$ в новое разбиение $\mathcal{A}'$:

- Пока есть хотя бы две части разбиения $\mathcal{A}$ размера меньше $\binom{n}{r-1}$, склеиваем их;
- Если в какой-то момент у нас осталась только одна часть разбиения, имеющая размер меньше $\binom{n}{r-1}$, то приклеиваем ее к какой-то из оставшихся частей.

Отметим, что после выполнения алгоритма для каждого из полученных множеств будет выполняться условие

$$\binom{n}{r-1} \leqslant |U| \leqslant \alpha\binom{n}{r}.$$

Оценим, на какую величину могла уменьшиться модулярность при таком процессе. Заметим, что реберный вклад нестрого возрастал, поэтому достаточно оценить изменение размера степенного штрафа.

Сначала разберем случай, когда процесс закончился после склеивания двух маленьких частей. Рассмотрим все множества $\mathcal{A}_{\mathrm{new}} := \{A \in \mathcal{A}' : A \notin \mathcal{A}\}$, которые появились в ходе нашего процесса. Очевидно, размер любого множества из $\mathcal{A}_{\mathrm{new}}$ не больше чем $2\binom{n}{r-1}$, что дает возможность применить для них лемму 6. Тогда степенной штраф асимптотически почти наверное увеличился не более чем на

$$\sum_{A \in \mathcal{A}_{\mathrm{new}}} \frac{\Big(\sum\limits_{v \in A} \deg(v,p)\Big)^2}{4e^2(G,p)} = (1+o(1)) \sum_{A \in \mathcal{A}_{\mathrm{new}}} \frac{|A|^2}{|V_G|^2} \leqslant (1+o(1)) \frac{2\binom{n}{r-1}}{\binom{n}{r}} = O(n^{-1}).$$

Теперь посмотрим на изменение степенного штрафа, когда процесс закончился доклеиванием одного маленького множества (обозначим его через B) к какому-то из оставшихся множеств C. В данном случае для всех $A \in \mathcal{A}_{\mathrm{new}} \setminus \{B \cup C\}$ выполнено $|A| \leqslant 2\binom{n}{r-1}$, а значит, и нужная оценка. Осталось провести оценку для $B \cup C$:

$$\frac{\Big(\sum\limits_{v \in B\cup C} \deg(v,p)\Big)^2}{4e^2(G,p)} - \frac{\Big(\sum\limits_{v \in B} \deg(v,p)\Big)^2}{4e^2(G,p)} - \frac{\Big(\sum\limits_{v \in C} \deg(v,p)\Big)^2}{4e^2(G,p)} \leqslant$$
$$\leqslant \frac{\Big(\sum\limits_{v \in B\cup C} \deg(v,p)\Big)^2}{4e^2(G,p)} - \frac{\Big(\sum\limits_{v \in C} \deg(v,p)\Big)^2}{4e^2(G,p)} = (1+o(1)) \left(\frac{(|B|+|C|)^2}{|V_G|^2} - \frac{|C|^2}{|V_G|^2} \right) =$$
$$= (1+o(1)) \frac{|B|(|B|+2|C|)}{|V_G|^2} \leqslant 2(1+o(1)) \frac{|B|}{|V_G|} = O(n^{-1}).$$

Итак, удалось доказать, что для почти всех последовательностей $G_n^p \in G_p(n,r,s)$ существуют разбиения $\mathcal{A}'_n$ графов G_n^p, удовлетворяющее условию

$$q^*(G_n^p) - q(\mathcal{A}'_n) = O(n^{-1}), \quad \forall U \in \mathcal{A}'_n : \binom{n}{r-1} \leqslant |U| \leqslant \alpha \binom{n}{r}.$$

Для таких разбиений в силу леммы 6 имеем

$$q(\mathcal{A}'_n) = 1 - \frac{1}{2} \sum_{A \in \mathcal{A}'_n} \frac{e(A, \overline{A}, p)}{e(G,p)} - \sum_{A \in \mathcal{A}'_n} \frac{\Big(\sum\limits_{v \in A} \deg(v)\Big)^2}{4e^2(G,p)} =$$
$$= (1+o(1)) \left(1 - \frac{1}{2} \sum_{A \in \mathcal{A}'_n} \frac{e(A, \overline{A})}{e(G)} - \sum_{A \in \mathcal{A}'_n} \frac{|A|^2}{|V_G|^2} \right) =$$
$$= (1+o(1)) \left(\sum_{A \in \mathcal{A}'_n} \frac{e(A)}{e(G)} - \sum_{A \in \mathcal{A}'_n} \frac{|A|^2}{|V_G|^2} \right),$$

после чего доказательство дословно повторяет рассуждения теоремы 5. ▲

Доказательство теоремы 7. Пойдем по пути доказательства теоремы 1 с небольшими изменениями:

$$q(\mathcal{A}) = \sum_{A\in\mathcal{A}} \frac{e(A,p)}{e(G,p)} - \sum_{A\in\mathcal{A}} \frac{\Bigl(\sum\limits_{v\in V_A} \deg(v,p)\Bigr)^2}{4e^2(G,p)} =$$

$$= 1 - \frac{1}{2}\sum_{A\in\mathcal{A}} \frac{e(V_A,\overline{V_A},p)}{e(G,p)} - \sum_{A\in\mathcal{A}} \frac{\Bigl(\sum\limits_{v\in V_A} \deg(v,p)\Bigr)^2}{4e^2(G,p)}.$$

Строим по разбиению $\mathcal{A}$ множества $\mathcal{A}'$ и $\mathcal{A}_{\rm big}$ по алгоритму из конструкции 1. Для них выполнены следующие условия:

$$\forall A\in\mathcal{A}' : |A| > \frac{1}{k}\binom{n}{r}, \quad \forall A\in\mathcal{A}_{\rm big} : |A| \geqslant \frac{2}{k}\binom{n}{r}.$$

Для случайного графа по-прежнему верно, что

$$q(\mathcal{A}) \leqslant 1 - \frac{1}{2}\sum_{A\in\mathcal{A}'} \frac{e(V_A,\overline{V_A},p)}{e(G,p)} - \sum_{A\in\mathcal{A}_{\rm big}} \frac{\Bigl(\sum\limits_{v\in V_A} \deg(v,p)\Bigr)^2}{4e^2(G,p)}.$$

Воспользуемся леммой 6 и получим

$$q(\mathcal{A}) \leqslant 1 - \frac{1}{2}(1+o(1))\sum_{A\in\mathcal{A}'} \frac{e(V_A,\overline{V_A})}{e(G)} - (1+o(1))\sum_{A\in\mathcal{A}_{\rm big}} \frac{\Bigl(\sum\limits_{v\in V_A} \deg(v)\Bigr)^2}{4e^2(G)} =$$

$$= 1 - (1+o(1))\left(\frac{1}{2}\sum_{A\in\mathcal{A}'} \frac{e(V_A,\overline{V_A})}{e(G)} + \sum_{A\in\mathcal{A}_{\rm big}} \left(\frac{|V_A|}{\binom{n}{r}}\right)^2\right).$$

Дальше по лемме 10 воспользуемся оценкой на суммы в скобках:

$$q(\mathcal{A}) \leqslant 1 - (1+o(1))\left(\frac{\binom{\lceil r/2\rceil}{s}}{2\binom{r}{s}}(1-\varepsilon)\Bigl(1-\frac{1}{k}\Bigr) - \frac{1}{k}\right).$$

Устремляя n и k к бесконечности и пользуясь тем, что $\varepsilon > 0$, получаем

$$q(\mathcal{A}) \leqslant 1 - \frac{\binom{\lceil r/2\rceil}{s}}{2\binom{r}{s}}. \quad \blacktriangle$$

3.4. Доказательство теоремы 8. Рассмотрим разбиение $\mathcal{A}$ из конструкции 2. Нетрудно видеть, что при $[cn] \geqslant r$ каждый элемент этого разбиения имеет размер не меньше чем $\binom{[cn]}{r-1}$ и не больше чем $\binom{[cn]}{r} \leqslant c^r\binom{n}{r}$. Выберем $\varepsilon \in (0,1)$ таким, что $\ln(1-\varepsilon) < -s^2$, тогда

$$\frac{-1}{\ln(1-\varepsilon)} < \frac{1}{s^2}.$$

Теперь обратим внимание на функцию $\alpha(x) = x\dfrac{x-\varepsilon}{1-\varepsilon}$. Как было доказано ранее, $\alpha(x) < x^2, x \in (0,1)$. Заметим, что $\alpha(x)$ непрерывна на $\mathbb{R}$, а $\alpha(1) = 1$, поэтому существует $\beta \in (0,1)$, для которого $\alpha(\beta) = c^r$.

Убедимся, что r, s удовлетворяют условиям теоремы 4 с выбранными α и β. Действительно, так как $\ln\Big(\dfrac{1-\beta}{1-\alpha/\beta}\Big) = \ln(1-\varepsilon)$, имеем

$$r \geqslant 2s = 2s + \frac{s^2}{s^2} - 1 > -\frac{s^2}{\ln(1-\varepsilon)} + 2s - 1 = -\frac{s^2}{\ln\Big(\dfrac{1-\beta}{1-\alpha/\beta}\Big)} + 2s - 1.$$

Применяя лемму 6, получаем, что модулярность разбиения $\mathcal{A}$ как разбиения $G_p(n,r,s)$ почти наверное равна $(1+o(1))q(\mathcal{A})$, где $q(\mathcal{A})$ – модулярность разбиения $\mathcal{A}$ как разбиения $G(n,r,s)$. Наконец, применяя теорему 3, получаем требуемое утверждение. ▲

Авторы выражают благодарность проф. А.М. Райгородскому за постановку задачи и обсуждение полученных результатов.

СПИСОК ЛИТЕРАТУРЫ

1. *Newman M.E.J., Girvan M.* Finding and Evaluating Community Structure in Networks // Phys. Rev. E. 2004. V. 69. № 2. P. 026113 (15 pp.). `https://doi.org/10.1103/PhysRevE.69.026113`
2. *Lancichinetti A., Fortunato S.* Limits of Modularity Maximization in Community Detection // Phys. Rev. E. 2011. V. 84. № 6. P. 066122 (8 pp.). `https://doi.org/10.1103/PhysRevE.84.066122`
3. *Miasnikof P., Prokhorenkova L., Shestopaloff A.Y., Raigorodskii A.* A Statistical Test of Heterogeneous Subgraph Densities to Assess Clusterability // Learning and Intelligent Optimization (13th Int. Conf. LION'13. Chania, Crete, Greece. May 27–31, 2019. Revised Selected Papers). Lect. Notes Comput. Sci. V. 11968. Cham: Springer, 2000. P. 17–29. `https://doi.org/10.1007/978-3-030-38629-0_2`
4. *Newman M.E.J.* Fast Algorithm for Detecting Community Structure in Networks // Phys. Rev. E. 2004. V. 69. № 6. P. 066133 (5 pp.). `https://doi.org/10.1103/PhysRevE.69.066133`
5. *Ostroumova Prokhorenkova L.* General Results on Preferential Attachment and Clustering Coefficient // Optim. Lett. 2017. V. 11. № 2. P. 279–298. `https://doi.org/10.1007/s11590-016-1030-8`
6. *Porter M.A., Onnela J.-P., Mucha P.J.* Communities in Networks // Notices Amer. Math. Soc. 2009. V. 56. № 9. P. 1082–1097. Available at `https://www.ams.org/notices/200909/rtx090901082p.pdf`.
7. *Brandes U., Delling D., Gaertler M., Görke R., Hoefer M., Nikoloski Z., Wagner D.* On Finding Graph Clusterings with Maximum Modularity // Graph-Theoretic Concepts in Computer Science (33rd Int. Workshop WG'2007. Dornburg, Germany. June 21–23, 2007. Revised Papers). Lect. Notes Comput. Sci. V. 4769. Berlin: Springer, 2007. P. 121–132. `https://doi.org/10.1007/978-3-540-74839-7_12`
8. *De Montgolfier F., Soto M., Viennot L.* Asymptotic Modularity of Some Graph Classes // Algorithms and Computation (Proc. 22nd Int. Sympos. ISAAC'2011. Yokahama, Japan. Dec. 5–8, 2011). Lect. Notes Comput. Sci. V. 7074. Berlin: Springer, 2011. P. 435–444. `https://doi.org/10.1007/978-3-642-25591-5_45`
9. *Trajanovski S., Wang H., Van Mieghem P.* Maximum Modular Graphs // Eur. Phys. J. B. 2012. V. 85. № 7. Art. 244 (14 pp.) `https://doi.org/10.1140/epjb/e2012-20898-3`
10. *McDiarmid C., Skerman F.* Modularity of Regular and Treelike Graphs // J. Complex Netw. 2018. V. 6. № 4. P. 596–619. `https://doi.org/10.1093/comnet/cnx046`

11. *Ostroumova Prokhorenkova L., Prałat P., Raigorodskii A.* Modularity in Several Random Graph Models // Electron. Notes Discrete Math. 2017. V. 61. P. 947–953. `https://doi.org/10.1016/j.endm.2017.07.058`
12. *Bollobás B.* The Isoperimetric Number of Random Regular Graphs // European J. Combin. 1988. V. 9. № 3. P. 241–244. `https://doi.org/10.1016/S0195-6698(88)80014-3`
13. *McDiarmid C., Skerman F.* Modularity of Erdős–Rényi Random Graphs // Random Structures Algorithms. 2020. V. 57. № 1. P. 211–243. `https://doi.org/10.1002/rsa.20910`
14. *Bollobás B., Narayanan B.P., Raigorodskii A.M.* On the Stability of the Erdős–Ko–Rado Theorem // J. Combin. Theory Ser. A. 2016. V. 137. P. 64–78. `https://doi.org/10.1016/j.jcta.2015.08.002`
15. *Мак-Вильямс Ф.Дж., Слоэн Н.Дж.А.* Теория кодов, исправляющих ошибки. М.: Связь, 1979.
16. *Frankl P., Wilson R.M.* Intersection Theorems with Geometric Consequences // Combinatorica. 1981. V. 1. № 4. P. 357–368. `https://doi.org/10.1007/BF02579457`
17. *Kahn J., Kalai G.* A Counterexample to Borsuk's Conjecture // Bull. Amer. Math. Soc. (N.S.). 1993. V. 29. № 1. P. 60–62. `https://doi.org/10.1090/S0273-0979-1993-00398-7`
18. *Райгородский А.М.* Вокруг гипотезы Борсука // Геометрия и механика. Современная математика. Фундаментальные направления. Т. 23. М: РУДН, 2007. С. 147–164. `http://mi.mathnet.ru/cmfd96`
19. *Ipatov M.M., Koshelev M.M., Raigorodskii A.M.* Modularity of Some Distance Graphs // European J. Combin. (submitted).
20. *Ipatov M.M.* Exact Modularity of Line Graphs of Complete Graphs // Moscow J. Comb. Number Theory. 2021. V. 10. № 1. P. 61–75. `https://doi.org/10.2140/moscow.2021.10.61`
21. *Koshelev M.M.* New Lower Bound on the Modularity of Johnson Graphs // Moscow J. Comb. Number Theory. 2021. V. 10. № 1. P. 77–82. `https://doi.org/10.2140/moscow.2021.10.77`
22. *Hoeffding W.* Probability Inequalities for Sums of Bounded Random Variables // J. Amer. Statist. Assoc. 1963. V. 58. № 301. P. 13–30. `https://doi.org/10.2307/2282952`

Деревянко Никита Михайлович
Московский физико-технический институт
(национальный исследовательский университет)
`nikitaderevyanko@gmail.com`
Кошелев Михаил Михайлович
Московский государственный университет
им. М.В. Ломоносова
`mkoshelev99@gmail.com`

Поступила в редакцию
22.06.2021
После доработки
27.11.2021
Принята к публикации
27.11.2021