

О ЧИСЛЕ РЁБЕР ОДНОРОДНОГО ГИПЕРГРАФА С ДИАПАЗОНОМ РАЗРЕШЁННЫХ ПЕРЕСЕЧЕНИЙ

© 2017 г. А. В. Бобу¹, А. Э. Куприянов¹, А. М. Райгородский^{1,2,3,*}

Представлено академиком РАН В.В. Козловым 19.01.2017 г.

Поступило 27.02.2017 г.

Настоящая работа посвящена исследованию величины $p(n, k, t_1, t_2)$ — максимально возможного числа рёбер в k -однородном гиперграфе, обладающем тем свойством, что мощности попарных пересечений рёбер лежат в отрезке $[t_1, t_2]$. Указываются ранее известные верхние и нижние оценки данной величины. Приводятся новые оценки величины $p(n, k, t_1, t_2)$, изучается их соотношение с уже известными результатами. Для некоторых значений параметров явно найдены значения исследуемой величины.

DOI: 10.7868/S0869565217220017

Напомним, что гиперграфом называется пара $H = (V, E)$, где V — некоторое конечное множество, а E — совокупность подмножеств множества V . Множество V обычно называют множеством вершин, а совокупность E — множеством рёбер гиперграфа. Гиперграф называется k -однородным, если в каждом рёбре содержится ровно k вершин.

В комбинаторике имеется ряд классических экстремальных задач, связанных с изучением максимального числа рёбер в однородном гиперграфе с разрешёнными или запрещёнными пересечениями рёбер. Мы рассмотрим задачи о максимальном числе рёбер в гиперграфе в ситуации, когда имеется ограничение на минимальное и максимальное пересечение рёбер, а также при условии лишь одного запрещённого пересечения. Речь идёт о величинах

$$f(n, k, t) = \max \{m: \exists H = (V, E), |V| = n, |E| = m,$$

$$\forall A \in E \quad |A| = k,$$

$$\forall A, B \in E \quad |A \cap B| \geq t\};$$

$$h(n, k, t) = \max \{m: \exists H = (V, E), |V| = n, |E| = m,$$

$$\forall A \in E \quad |A| = k,$$

$$\forall A, B \in E \quad |A \cap B| \leq t\};$$

¹ Московский государственный университет им. М.В. Ломоносова

² Московский физико-технический институт (государственный университет), Долгопрудный Московской обл.

³ Институт математики и информатики Бурятского государственного университета, Улан-Удэ

*E-mail: mraigor@yandex.ru

$$m(n, k, t) = \max \{m: \exists H = (V, E), |V| = n, |E| = m,$$

$$\forall A \in E \quad |A| = k,$$

$$\forall A, B \in E \quad |A \cap B| \neq t\}.$$

Задача об отыскании величины $f(n, k, t)$ полностью решена Р. Алсведе и Л. Хачатрянном (см. [1]). Работа над второй проблемой всё ещё далека от завершения: имеется довольно простая нижняя оценка Р. Варшавова и Е. Гилберта (см. [2, 3]) и ряд блестящих верхних границ, среди которых можно отметить оценки Левенштейна (см. [4]) и границу линейного программирования (см. [5]). Наконец, получение нижних оценок $m(n, k, t)$ значительно продвинуло комбинаторную геометрию ([6, 7]), а поиск верхних оценок этой величины привёл к созданию классического линейно-алгебраического метода Франкла и Уилсона (см. [8]). Впоследствии появилось большое количество работ, посвящённых изучению различных ситуаций относительно параметров k и t при $n \rightarrow \infty$ (см. [7, 9–13]).

Перейдём к формулировке нашей задачи. Она будет состоять в отыскании величины $p(n, k, t_1, t_2)$ — максимального числа рёбер в k -однородном гиперграфе на n вершинах, все рёбра которого пересекаются не менее чем по t_1 элементам и не более чем по t_2 элементам, $0 \leq t_1 \leq t_2 \leq k < n$:

$$p(n, k, t_1, t_2) = \max \{m: \exists H = (V, E), |V| = n, |E| = m,$$

$$\forall A \in E \quad |A| = k,$$

$$\forall A, B \in E \quad t_1 \leq |A \cap B| \leq t_2\}.$$

Можно сформулировать данную проблему в терминах теории кодирования: разрешение определённого диапазона пересечений $[t_1, t_2]$ эквивалентно условию, что все хэмминговы расстояния между кодовыми словами веса k заключены в отрезке $[2(k - t_2), 2(k - t_1)]$. Кроме того, можно описать нашу задачу на языке теории дистанционных графов. А именно, рассмотрим граф $G(n, k, t_1, t_2) = (V(n, k), E(n, k, t_1, t_2))$, у которого

$$\begin{aligned} V(n, k) &= \{ \mathbf{x} = (x_1, \dots, x_n) : x_i \in \{0, 1\}, \\ &\quad x_1 + \dots + x_n = k \}, \\ E(n, k, t_1, t_2) &= \{ \{ \mathbf{x}, \mathbf{y} \} : (\mathbf{x}, \mathbf{y}) \notin [t_1, t_2] \}. \end{aligned}$$

Напомним, что для произвольного графа G числом независимости $\alpha(G)$ называется максимальная мощность множества вершин, никакие две из которых не соединены ребром. В такой постановке становится очевидно, что наша задача эквивалентна поиску величины $\alpha(G(n, k, t_1, t_2))$, поскольку последняя просто совпадает с искомой $p(n, k, t_1, t_2)$. Наконец, дополнительной мотивацией для изучения поставленной проблемы служат недавние работы о хроматическом числе пространства с запрещёнными равнобедренными треугольниками (см. [14, 15]).

1. ИЗВЕСТНЫЕ ОЦЕНКИ

Перейдём к известным оценкам величины $p(n, k, t_1, t_2)$. Начнём с нижних оценок. Для этого нам потребуется сформулировать теорему Алсведе–Хачатряна, которая даёт точное значение величины $f(n, k, t)$, и границу Варшавова–Гилберта, которая оценивает снизу величину $h(n, k, t)$.

Теорема 1. Пусть $2k - t \leq n$ и для некоторого $r \in \mathbb{N} \cup \{0\}$

$$(k - t + 1) \left(2 + \frac{t-1}{r+1} \right) \leq n < (k - t + 1) \left(2 + \frac{t-1}{r} \right),$$

причём при $r = 0$ формально полагаем, что $\frac{t-1}{r} = \infty$. Тогда

$$f(n, k, t) = \sum_{i=r}^{\min\{2r, k-t\}} C_{t+2r}^{t+i} C_{n-t-2r}^{k-t-i}.$$

Теорема 2. Пусть $2k - t \leq n$. Тогда

$$h(n, k, t) \geq \frac{C_n^k}{\sum_{i=t+1}^k C_k^i C_{n-k}^{k-i}} - 1. \quad (1)$$

На сочетании этих двух результатов основана идея нижней оценки величины $p(n, k, t_1, t_2)$, впервые доказанной в работе [14].

Теорема 3. Пусть $k \leq \frac{n}{2}$. Тогда

$$p(n, k, t_1, t_2) \geq \frac{C_{t_1+2r}^{t_1+r} C_{n-t_1-2r}^{k-t_1-r}}{\sum_{i=t_2+1}^k \sum_{j=\max\{t_1, t_1+r+i-k\}}^{\min\{t_1+r, i\}} C_{t_1+r}^j C_r^{t_1+r-j} C_{k-t_1-r}^{i-j} C_{n-k-r}^{k-t_1-r-i+j}} - 1.$$

Перейдём к верхним оценкам. Для начала отметим, что существуют два соотношения:

$$\begin{aligned} p(n, k, t_1, t_2) &\leq p(n, k, t_1, k) = f(n, k, t_1), \\ p(n, k, t_1, t_2) &\leq p(n, k, 0, t_2) = h(n, k, t_2). \end{aligned}$$

Если значение величины $f(n, k, t)$ в точности известно, выполняется следующее утверждение.

Теорема 4. Пусть $2k - t_1 \leq n$ и для некоторого $r \in \mathbb{N} \cup \{0\}$

$$(k - t_1 + 1) \left(2 + \frac{t_1-1}{r+1} \right) \leq n < (k - t_1 + 1) \left(2 + \frac{t_1-1}{r} \right).$$

Тогда

$$p(n, k, t_1, t_2) \leq \sum_{i=r}^{\min\{2r, k-t_1\}} C_{t_1+2r}^{t_1+i} C_{n-t_1-2r}^{k-t_1-i}.$$

Обратимся теперь ко второму соотношению. Для того чтобы его применить, нам потребуются некоторые результаты теории кодирования. Первый из них – граница Левенштейна – доказан в работе [4] и звучит следующим образом.

Теорема 5. Пусть $k \leq \frac{n}{2}$ и

$$\begin{aligned} 0 \leq i - j \leq 2(n - k), \quad 0 \leq i + j \leq 2k, \\ \frac{(i-j)^2}{4(k-j)} + \frac{(i+j)^2}{4(n-k+j)} + k - t_2 - i > 0, \end{aligned}$$

тогда справедливо следующее неравенство:

$$p(n, k, t_1, t_2) \leq h(n, k, t_2) \leq \frac{C_n^{k-j}}{C_k^{(i+j)/2} C_{n-k}^{(i-j)/2}} \times$$

$$\times \left[\frac{k - t_2}{\frac{(i-j)^2}{4(k-j)} + \frac{(i+j)^2}{4(n-k+j)} + k - t_2 - i} \right].$$

Важной оценкой является и граница линейного программирования, полученная в работе [5].

Теорема 6. Введём функции

$$H_2(x) = -x \log_2 x - (1-x) \log_2 (1-x);$$

$$g(x) = H_2 \left(\frac{1 - \sqrt{1-x}}{2} \right).$$

Пусть $0 \leq t'_1 \leq t'_2 \leq k' \leq \frac{1}{2}$, тогда

$$\overline{\lim}_{n \rightarrow \infty} \frac{\log_2 p(n, k'n, t'_1 n, t'_2 n)}{n} \leq \overline{\lim}_{n \rightarrow \infty} \frac{\log_2 h(n, k'n, t'_2 n)}{n} \leq$$

$$\leq \begin{cases} 0, & \text{если } t'_2 \leq k'^2, \\ g(u^2), & \text{если } t'_2 > k'^2, \end{cases}$$

где $u = -2(k' - t'_2) + 2\sqrt{t'_2(t'_2 - 2k' + 1)}$.

Наконец, с помощью простой модификации теоремы Франкла–Уилсона, несложно получить следующую границу.

Теорема 7. Справедливо следующее неравенство:

$$p(n, k, t_1, t_2) \leq \sum_{q=0}^{t_2-t_1} C_n^q. \quad (2)$$

2. ФОРМУЛИРОВКИ РЕЗУЛЬТАТОВ

Нами были получены новые нижняя и верхняя оценки величины $p(n, k, t_1, t_2)$.

Теорема 8. Пусть $k \leq \frac{n}{2}$ и $0 \leq \tau \leq t_1$, $0 \leq r \leq k - \tau$ — целые числа. Тогда

$$p(n, k, t_1, t_2) \geq \frac{C_{\tau+2r}^{\tau+r} C_{n-\tau-2r}^{k-\tau-r}}{C_1 + C_2} - 1,$$

где

$$C_1 = \sum_{i=\tau}^{t_1-1} \sum_{j=\max\{\tau, \tau+r+i-k\}}^{\min\{\tau+r, i\}} C_{\tau+r}^j C_r^{\tau+r-j} C_{k-\tau-r}^{i-j} C_{n-k-r}^{k-\tau-r-i+j},$$

$$C_2 =$$

$$= \sum_{i=t_2+1}^k \sum_{j=\max\{\tau, \tau+r+i-k\}}^{\min\{\tau+r, i\}} C_{\tau+r}^j C_r^{\tau+r-j} C_{k-\tau-r}^{i-j} C_{n-k-r}^{k-\tau-r-i+j},$$

причём при $\tau = t_1$ формально полагаем $C_1 = 0$.

Идейно доказательство устроено следующим образом. На первом шаге мы берём конструкцию из доказательства теоремы 1 так, чтобы попарные пересечения её представителей были не меньше τ . Затем с помощью алгоритма теоремы 2 мы удаляем некоторые множества таким образом, чтобы пересечения оставшихся лежали в отрезке $[t_1, t_2]$.

Сформулируем новую верхнюю оценку.

Теорема 9. Пусть $0 \leq r < k - t_2$ — целое число. Введём следующие обозначения:

$$n_0 = n - t_1 - 2r; \quad k'_0 = \frac{k - t_1 - r}{n - t_1 - 2r}; \quad t'_0 = \frac{t_2 - t_1}{n - t_1 - 2r};$$

$$H_2(x) = -x \log_2 x - (1-x) \log_2 (1-x);$$

$$g(x) = H_2 \left(\frac{1 - \sqrt{1-x}}{2} \right);$$

$$u = -2(k'_0 - t'_0) + 2\sqrt{t'_0(t'_0 - 2k'_0 + 1)};$$

$$C_1 = \overline{\lim}_{n \rightarrow \infty} \frac{1}{n} \log_2 \frac{C_n^{t_1+2r}}{C_k^{t_1+r} C_{n-k}^r},$$

$$C_2 = \begin{cases} \frac{n_0}{n} g(u^2), & \text{если } t'_0 > k_0'^2, \\ 0, & \text{если } t'_0 \leq k_0'^2. \end{cases}$$

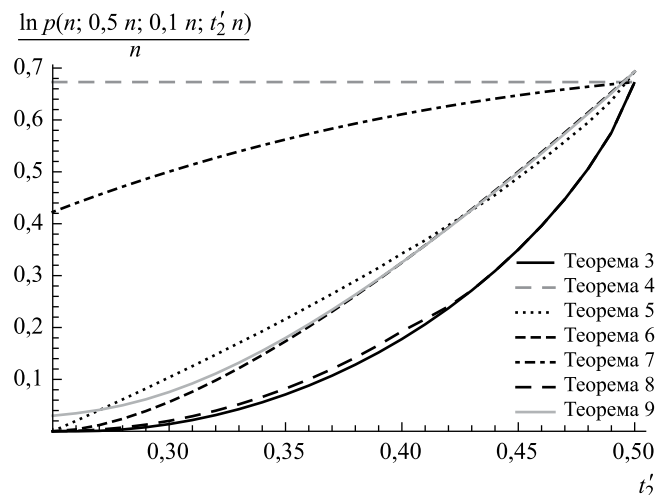


Рис. 1. Сравнение оценок теорем 3–9 при $k' = 0,5$ и $t'_1 = 0,1$.