Министерство образования и науки Российской Федерации

Государственное образовательное учреждение
высшего профессионального образования
«Московский физико-технический институт (государственный университет)»
Факультет инноваций и высоких технологий
Кафедра анализа данных

Магистерская диссертация

Тема: **Название моей работы (TODO)**

Направление: 010400
Прикладные математика и информатика

Выполнил:
студент 093 группы ____________ Попов М.В.

Научный руководитель:
д.физ.-мат.н., проф.(todo) ____________ Ромащенко А.Е.

г. Москва 2016

Содержание

Введение

(ToDo) Актуальность, новизна, краткая выжимка.

Коммуникационная сложность

1.1 Постановка задачи

Мы будем рассматривать задачи следующего вида: пусть имеется два человека, которые хотят совместно вычислить значение некоторой функции от двух переменных $f(x, y)$. По традиции мы будем называть первого участника игры Алисой, а второго Бобом. Сложность у этой задачи в том, что Алиса знает только значение аргумента x, а Боб значение аргумента y. Алиса и Боб могут обмениваться сообщениями по каналу связи. Требуется вычислить значение $f(x, y)$, переслав по каналу связи минимальное количество информации.

Мы предполагаем, что Алиса и Боб заранее (до того, как им станут известны значения x и y) договариваются о коммуникационном протоколе — о наборе соглашений, какие именно данные и в каком порядке они будут пересылать друг другу при тех или иных значениях x и y.

Опишем теперь всю задачу более формально. Пусть имеются конечные множества X, Y, Z и задана некоторая функция $f : X \times Y \to Z$.

Определение. *Коммуникационным протоколом для вычисления некоторой функции $f : X \times Y \to Z$ называется ориентированное двоичное дерево со следующей разметкой на вершинах и ребрах:*

- *каждая нелистовая вершина помечена буквой A или B;*
  - *у вершин с пометкой A определена функция $g_i : X \to \{0, 1\}$;*
  - *у вершин с пометкой B определена функция $f_j : Y \to \{0, 1\}$;*
- *каждой листовой вершине сопоставлен элемент множеста Z;*
- *каждое ребро помечено 0 или 1.*

Пусть Алиса и Боб договорились, что будут действовать по некоторому протоколу $\mathcal{P}$. Затем Алиса получила $x \in X$, а Боб получил $y \in Y$.

Поместим фишку в корневую вершину нашего протокола $\mathcal{P}$ и будем перемещать ее вниз по дереву, последовательно удаляясь от корня, пока она не попадём в один из листьев. Перемещение фишки выполняется следующим образом. Если текущая вершина помечена буквой A это значит, что сейчас очередь Алисы. Она применяет функцию g_i текущей вершины к своему значению x. Алиса отправляет по каналу связи бит равный $g_i(x)$ и перемещает фишку по ребру, помеченному как $g_i(x)$. Боб получает отправленный бит и понимает куда была сдвинута фишка. Для вершин помеченных буквой B поступают аналогично. Когда фишка попадает в лист дерева, записанное там значение $z \in Z$ объявляется результатом выполнения протокола.

Мы говорим, что протокол $\mathcal{P}$ вычисляет функцию $f : X \times Y \to Z$, если для любого $x \in X$ и любого $y \in Y$ при движении из корня по пути, соответствующему заданным x и y, мы попадаем в лист, помеченный $z = f(x, y)$.

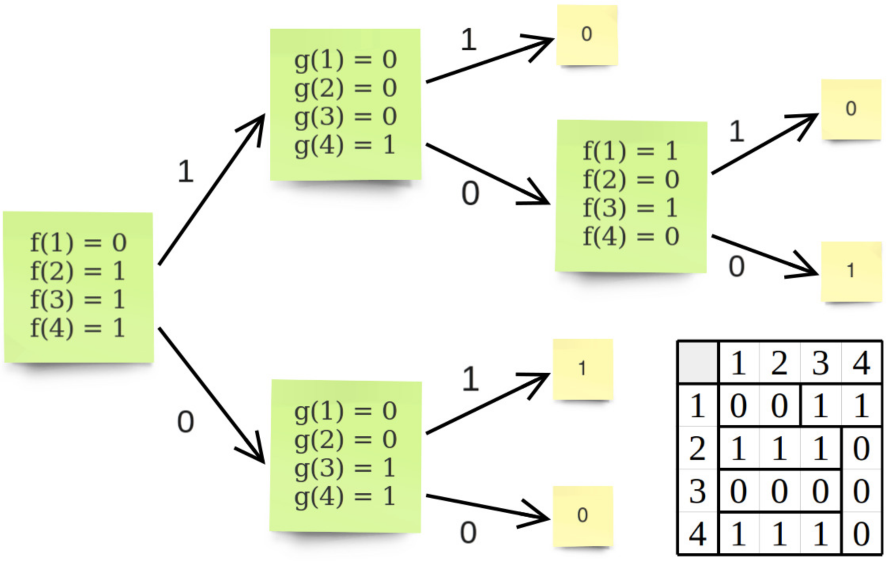
Определение. *Сложностью коммуникационного протокола называется его глубина. Коммуникационной сложностью функции f называется минимальная сложность протокола, вычисляющего f. Мы будем обозначать её $CC(f)$.*

1.2 Одноцветные комбинаторные прямоугольники

Определение. *Множество $S \subset X \times Y$ называется комбинаторным прямоугольником (или просто прямоугольным множеством), если существуют такие $A \subset X$ и $B \subset Y$, что $S = A \times B$.*

Пусть $\mathcal{P}$ некоторый коммуникационный протокол для вычисления функции $f : X \times Y \to Z$ и l один из листьев протокола. Определим S_l, как множество пар $(x, y) \in X \times Y$ таких, что на входе (x, y) Алиса и Боб, следуя протоколу $\mathcal{P}$, приходят в лист l.

Утверждение. *Для всякого коммуникационного протокола $\mathcal{P}$ и для всякого листа l множество S_l является комбинаторным прямоугольником.*



	1	2	3	4
1	0	0	1	1
2	1	1	1	0
3	0	0	0	0
4	1	1	1	0

Рис. 1: Пример протокола и разбиения таблицы значений.

Доказательсво этого утверждения можно прочитать например в [1]. В итоге мы получаем, что коммуникационный протокол для вычисления функции f задаёт разбиение $X \times Y$ - таблицы значений f на прямоугольные множества, соответствующие листьям. Поскольку каждому листу протокола приписано одно значение функции f, эти прямоугольные множества являются одноцветными, то есть во всех точках такого прямоугольного множества функция f принимает одно и то же значение. Например, для $X = Y = \{1, 2, 3, 4\}$, $Z = \{0, 1\}$ и протокола $\mathcal{P}$ (рис 1.) получаем разбиение на 5 одноцветных прямоугольных множества.

Подведем промежуточные итоги: всякий протокол с l листьями (вычисляющий функцию f) задаёт разбиение таблицы значений f на l одноцветных прямоугольных множества. Значит, чтобы доказать, что коммуникационная сложность $CC(f)$ не меньше n, достаточно показать, что таблицу значений невозможно разбить на менее, чем 2^n одноцветных прямоугольных множества.

1.3 Графовая интерпретация

Давайте теперь посмотрим на другое представление множества значений функции f. Рассмотрим полный двудольный граф $G = (X, Y, E)$, ребра которого раскрашены в $|Z|$ цветов. Вершины левой доли соот-

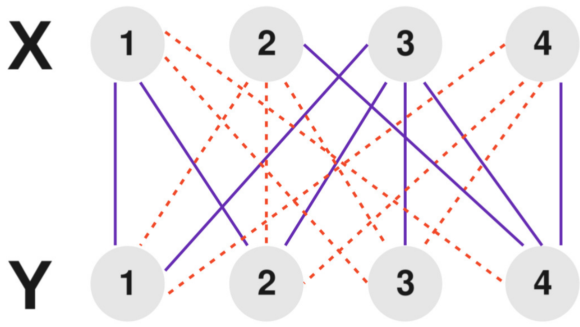


Рис. 2: Графовая интерпретация: синие – 0, красные – 1.

ветствуют элементам множества X, вершины правой доли - элементам множества Y. Ребро $(x, y) \in X \times Y$ имеет цвет $z \in Z$, если $f(x, y) = z$.

Из определения комбинаторного прямоугольника видно, что в графовой интерпертации он является ничем иным, как полным двудольным подграфом. А разбиение таблицы значений f на одноцветные прямоугольные множества это разбиение нашего полного двудольного графа G на одноцветные непересекающиеся биклики (полные двудольные подграфы). Для нашего примера графовую интерпретацию можно посмотреть на рис.2.

Определение. *Бикликовым числом $bcc(G)$ двудольного графа G будем называть наименьшее число биклик, которыми можно покрыть все ребра графа G (Биклики могут пересекаться).*

Для каждого $z \in Z$ определим двудольный граф $G_z = (X, Y, E_z)$, как граф, получающийся из G выкидыванием всех ребер цвета отличного от z, иначе говоря $E_z = \{(x, y) \in X \times Y \mid f(x, y) = z\}$.

Величины $bbc(G_z)$ дают некоторую нижнюю оценку на величину минимального покрытия непересекающимися бикликами, поэтому:

$$2^{CC(f)} \geq \sum_{z \in Z} bcc(G_z)$$

Замечание. *На самом деле величины $bbc(G_z)$ тесно связаны с недетерминированной коммуникационной сложностью $NCC(f)$. Для произво-*

вльного множества Z верно:

$$NCC(f) \leq \lceil log_2(\sum_{z \in Z} bcc(G_z)) \rceil + 1$$

а для $Z = \{0, 1\}$:

$$NCC(f) = \lceil log_2(bcc(G_1)) \rceil$$

Подробнее про это можно прочитать, например в [2].

В итоге мы получили мощный иструмент для доказательства нижних оценок коммуникационной сложности. К сожалению задача нахождения величины $bcc(G)$ является PSPACE-полной [3], а точное значение известно только для очень скудного класса графов (например для "Crown graphs"), поэтому напрямую мы не можем использовать эту оценку. В следующей главе я рассмотрю несколько методов, позволяющих для произвольного двудольного графа оценивать снизу величину $bcc(G)$.

Оценивание $bcc(G)$

В этой главе я опишу три различных метода оценивания бикликового покрытия:

- метод трудного множества ("Fooling Set");
- метод Куликова-Юкны;
- метод энтропийных неравенств.

Первые два метода работают для произвольных графов (необязательно двудольных), а третий применим к большому классу двудольных графов.

2.1 Метод трудного множества

Данный метод тесно связан с одноцветными прямоугольными множествами. Классическое определение трудного множества выглядит следующим образом:

Определение. *Для функции* $f : X \times Y \to Z$ *и элемента* $z \in Z$ *будем называть множество* $S_z \subset X \times Y$ *трудным (в англоязычной литературе fooling set), если верно:*

- *для всякой пары* $(x, y) \in S_z$ *имеем* $f(x, y) = z$*;*
- *для любых двух несовпадающих пар* $(x, y) \in S_z$ *и* $(x', y') \in S_Z$ *имеем* $f(x, y') \neq z$ *или* $f(x', y) \neq z$*.*

Нас будет интересовать немного более общее определение трудного множества (графовая интерпретация):

Определение. *Пусть* $G = (V, E)$ *произвольный неориентированный граф. Будем называть подмножество ребер* $S \subseteq E$ *трудным, если для любых двух различных ребер* $(x, y) \in S$ *и* $(x', y') \in S$ *имеем* $(x, y') \notin E$ *или* $(x', y) \notin E$*.*

Замечание. *Классическое определение получается из графового, применением к двудольному графу* $G_z = (X, Y, E_z)$*, который строится по функции* $f : X \times Y \to Z$*.*

Теорема. *Для произвольного неориентированного графа* $G = (V, E)$*, если подмножество ребер* $S \subseteq E$ *является трудным, то* $bcc(G) \geq |S|$*.*

Доказательство. *Достаточно доказать, что два ребра, лежащие одновременно в одном трудном множестве, не могут попасть в одну биклику. Пусть не так, значит существуют два ребра* $(x, y) \in B \cap S$ *и* $(x', y') \in B \cap S$*, где* B *- биклика, а* S *- трудное подмножество ребер. Но тогда ребра* (x, y') *и* (x', y) *также принадлежат биклике* B*, а значит лежат и в нашем множестве ребер* E*. Противоречие.* ■

Замечание. *На практике нахождение максимального по мощности трудного множества применяют редко, потому что эта задача является* $PSPACE$*-полной [3]. Часто рассматривают "нечестный" метод, а именно доказывают, что определенного размера трудное множество обязательно найдется.*

Теорема. *Для произвольного неориентированного графа* $G = (V, E)$ *обозначим за* $v(G)$ *- размер максимального паросочетания, а за* $cl(G)$

такое максимальное число r, что в нашем графе содержится биклика $K_{r,r}$. Тогда среди ребер этого паросочетания можно найти трудное множество размера $\left\lceil \frac{v(G)}{cl(G)} \right\rceil$.

На самом деле намного проще доказать, что $bcc(G) \geq \left\lceil \frac{v(G)}{cl(G)} \right\rceil$ без участия трудного множества. Этот факт сразу следует из того, что любая биклика $K_{r,s}$ содержит как максимум $min\{r, s\}$ ребер максимального паросочетания. Сама теорема будет доказана в одной из последующих глав.

2.2 Метод Куликова-Юкны

Следующий метод был впервые описан в статье [4] и работает он для произвольного неориентированного графа.

Теорема. *Для произвольного неориентированного графа $G = (V, E)$ верно:*

$$bcc(G) \geq \left\lceil \frac{v(G)^2}{|E|} \right\rceil$$

Доказательство. *Пусть $M \subseteq E$ - это максимальное паросочетание, тогда рассмотрим бикликовое покрытие, на котором достигается минимум $E = B_1 \cup B_2 \cup \ldots \cup B_{bcc(G)}$. Определим отображение $g : M \to \{1, \ \ldots, \ bcc(G)\}$, как $g(e) = min\{i \mid e \in B_i\}$ и пусть $M_i = \{e \in M \mid g(e) = i\}$. Иначе говоря M_i содержит только те ребра максимального паросочетания M, которые покрываются бикликой B_i впервый раз.*

Пусть $F_i \subseteq B_i$ биклика, индуцированная вершинами ребер из M_i. Пусть $F = F_1 \sqcup F_2 \sqcup \ldots \sqcup F_{bcc(G)}$ (биклики F_i не пересекаются по построению).

Очевидно, что F_i - биклика размера $r_i \times r_i$, где $r_i = |M_i|$. Получаем следующие соотношения:

$$r_1 + r_2 + \ldots + r_{bcc(G)} = |M| = v(G)$$

и

$$r_1^2 + r_2^2 + \ldots + r_{bcc(G)}^2 = |F|$$

Из неравенства Коши-Буняковского получаем

$$v(G)^2 = (r_1+r_2+\ldots+r_{bcc(G)})^2 \le bcc(G)\cdot(r_1^2+r_2^2+\ldots+r_{bcc(G)}^2) = bcc(G)\cdot|F|$$

А так как $F \subseteq E$, *то* $v(G)^2 \le bcc(G)\cdot|F| \le bcc(G)\cdot|E|$. ■

На некоторых графах данная оценка превосходит $\left\lceil \frac{v(G)}{cl(G)} \right\rceil$, а на некоторых уступает:

- пусть двудольный граф $G = (L, R, E)$ состоит из совершенного паросочетания размера $n = |L| = |R|$ и еще некоторого константного числа непересекающихся биклик $K_{r,r}$. К тому же, пусть $r = \Theta(\sqrt{n})$, тогда
$$\frac{v(G)^2}{|E|} = \frac{n^2}{cr^2+n} = \Theta(n) \gtrsim \Theta(\sqrt{n}) = \frac{n}{r} = \frac{v(G)}{cl(G)}$$
- рассмотрим двудольный граф Леви, построенный при помощи конечной проективной плоскости порядка $p \in \mathbb{P}$. В каждой доле этого графа содержится $n = p^2+p+1$ вершин, причем степень каждой $p+1$. Этот граф не содержит $K_{2,2}$ (любые две прямые пересекаются максимум в одной точке). А так как в регулярных двудольных графах обязательно найдется совершенное паросочетание, то

$$\frac{v(G)^2}{|E|} = \frac{(p^2+p+1)^2}{(p^2+p+1)(p+1)} = \Theta(\sqrt{n}) \lesssim \Theta(n) = \frac{p^2+p+1}{1} = \frac{v(G)}{cl(G)}$$

2.3 Метод энтропийных неравенств

Следующий метод оценивания бикликового покрытия был описан в статье [5] как результат применения энтропийного неравенства:

$$H(A|B,X) + H(A|B,Y) \le H(A|B)$$

К сожалению, это неравенство выполняется не для произвольного совместного распределения случайных величин A, B, X, Y и соответственно на двудольный граф будет накладываться дополнительное условие (*).

Теорема. *Пусть ребра двудольного графа* $G = (L, R, E)$ *раскрашены следующим образом:*

(*) для произвольной биклики $C \subseteq E$ и для произвольной пары ребер (x, y') и (x', y) из C, покрашеных в цвет a, цвет ребер (x, y) и (x', y') тоже a.

Пусть также на ребрах этого графа задано произвольное вероятностное распределение. Определим случайные величины (X, Y, A) следующим образом:

- $X = [\textit{левый конец ребра}]$,
- $Y = [\textit{правый конец ребра}]$,
- $A = [\textit{цвет ребра}]$.

Тогда выполняется неравенство:

$$bcc(G) \geq 2^{\frac{1}{2}(H(A|X)+H(A|Y)-H(A))}$$

Пример. *Определим двудольный граф $G_{n,k} = (L, R, E)$ следующим образом:*

- *L и R всевозможные k-элементные подмножества $\{1,\ 2,\ \ldots,\ n\}$,*
- *$E \subseteq L \times R$ состоит из пар непересекающихся множеств.*

Определим цвет ребра $(x, y) \in E$ как $x \sqcup y$ и пусть на ребрах задано равномерное распределение. Условие (*) выполнено, потому что любые два одноцветных ребра не могут лежать в одной биклике. А так как $H(A|X) = H(A|Y) = \log_2 \binom{n-k}{k}$ и $H(A) = \log_2 \binom{n}{2k}$, то

$$bcc(G_{n,k}) \geq \sqrt{\binom{n-k}{k}^2 \Big/ \binom{n}{2k}}$$

Если $n \gg k$, то $\binom{n-k}{k}^2 / \binom{n}{2k}$ близко к $\binom{2k}{k} \approx 2^{2k}$ и мы получаем нижнюю оценку $bcc(G_{n,k}) \geq 2^k$.

Геометрические конфигурации

В этой главе мы приведем класс двудольных графов, построенных при помощи некоторой геометрической конфигурации Г. Далее мы увидим, что к этим двудольным графам применимы все наши оценки и поэтому, изменяя Γ, мы можем сравнить какие методы работают лучше, а какие хуже.

3.1 Описание двудольного графа

Определение. *Геометрической конфигурацией* Γ *(Partial Linear Space) будем называть конечное множество прямых* A *и конечное множество точек* V *на них, что выполняются следующие две аксиомы:*

- *Любые две точки лежат как максимум на одной прямой.*
- *На каждой прямой лежит хотя бы две точки.*

Определение. *Проективной плоскостью с параметрами* (p_γ, l_π) *называется геометрическая конфигурация, состоящая из* p *точек и* l *прямых, причем через каждую точку проходит ровно* γ *прямых и на каждой прямой лежит ровно* π *точек.*

Пусть у нас имеется некоторая геометрическая конфигурация $\Gamma = (V, A)$, тогда определим двудольный граф $G_{n,\Gamma} = (L, R, E)$ следующим образом:

- $L = R = V^n$
- $E = \{(x, y) \in L \times R \mid \forall i : x_i \neq y_i$ и лежат на одной прямой из $A\}$

3.2 Оценки для проективных плоскостей

Для произвольной проективной плоскости Γ с параметрами (p_γ, l_π) найдем какие оценки на $bcc(G_{n,\Gamma})$ дают наши методы:

– Метод трудного множества:

Лемма. *Если в* Γ *имеется цикл нечетный длины, то в* $G_{1,\Gamma}$ *можно найти трудное множество размера 3.*

Доказательство. *Рассмотрим нечетный цикл минимальной длины* $\{v_1,\ v_2,\ \dots,\ v_{2k+1}\}$*, где* $k \geq 1$*. Заметим, что прямые могут проходить только через соседние точки этого цикла, иначе бы мы* *нашли нечетный цикл меньшей длины. Тогда если $k > 1$, то множество ребер $\{(v_1, v_2),\ (v_2, v_3),\ (v_3, v_4)\}$ образует трудное множество, а если $k = 1$, то $\{(v_1, v_2),\ (v_2, v_3),\ (v_3, v_1)\}$ образует трудное множество.* $\square$

Замечание. *Если на какой-нибудь прямой лежит по крайне мере три точки, то мы уже имеем цикл длины 3.*

Если нечетных циклов в Γ нет, то мы получаем геометрическую конфигурацию аналогичную двудольному графу. Если этот двудольный граф полный, то наибольшее трудное множество имеет размер 2, а если неполный, то мы можем найти трудное множество размера 3.

Лемма. *Если в $G_{1,\Gamma}$ существует трудное множество размера k, то в $G_{n,\Gamma}$ существует трудное множество размера k^n*

Доказательство. *Докажем вначале, что если в графе G_1 имеется трудное множество размера n_1, а в графе G_2 – трудное множество размера n_2, тогда в $G_1 \otimes G_2$ можно найти трудное множество размера $n_1 n_2$. (где $\otimes$ - произведение Кронекера). Пусть $\{v_{i,j}\}$ трудное множество в графе G_1, тогда в каждой подматрице $v_{i,j} \cdot G_2$ матрицы графа $G_1 \otimes G_2$ рассмотрим клетки, соответствующие трудному множеству графа G_2. Всего мы получили $n_1 n_2$ клеток, образующие трудное множество графа $G_1 \otimes G_2$ по построению.*

Вернемся к доказательству леммы. Так как матрица графа $G_{n,\Gamma}$, есть не что иное, как Кронекерово произведение n матриц графа $G_{1,\Gamma}$, то мы можем найти трудное множество размера k^n. $\square$

В итоге мы получили, что если Γ является аналогом полного двудольного графа, то

$$bcc(G_{n,\Gamma}) \geq 2^n$$

иначе

$$bcc(G_{n,\Gamma}) \geq 3^n$$

– Метод Куликова-Юкны:

Так как Γ имеет параметры (p_γ, l_π), то каждая вершина графа $G_{1,\Gamma}$ соединена с $\gamma\cdot(\pi-1)$ другими, а значит всего ребер $\gamma\cdot(\pi-1)\cdot p$. Тогда в графе $G_{n,\Gamma}$ всего ребер $\gamma^n\cdot(\pi-1)^n\cdot p^n$. Так как у нас однородный двудольный граф, то у нас имеется совершенное паросочетание, а значит $v(G_{n,\Gamma}) = p^n$. В итоге получаем оценку:

$$bcc(G_{n,\Gamma}) \geq \frac{p^{2n}}{\gamma^n\cdot(\pi-1)^n\cdot p^n} = \left(\frac{p}{\gamma\cdot(\pi-1)}\right)^n$$

– Метод энтропийных неравенств:

Определим раскраску ребер нашего графа $G_{n,\Gamma} = (L, R, E)$: сопоставим каждой прямой конфигурации Γ свой цвет, тогда цвет ребра $(x, y) \in E$ равен n-мерному вектору цветов прямых проходящих через x_i и y_i.

Проверим свойство (*): пусть (x, y') и (x', y) одного цвета и лежат в одной биклике C, значит для любого i точки x_i, y'_i, x'_i, y_i лежат на одной прямой (некоторые точки могут совпадать), но тогда очевидно, что ребро (x, y) такого же цвета.

Пусть на ребрах графа задано равномерное распределение, тогда $H(A) = \log_2 l^n = n\cdot\log_2 l$ и $H(A|X) = H(A|Y) = \log_2\gamma^n = n\cdot\log_2\gamma$. В итоге получаем оценку:

$$bcc(G_{n,\Gamma}) \geq 2^{n\cdot\log_2\gamma-\frac{n}{2}\cdot\log_2 l} = \left(\frac{\gamma}{\sqrt{l}}\right)^n$$

3.3 Сравнение методов оценивания

Рассмотрим какие оценки получаются на известных геометрических конфигурациях. Симметричные конфигурации ($p = l$ и $\gamma = \pi$) будем обозначать сокращенно (p_γ).

Название	FS	KJ	EI	Результат
Треугольник (3_2)	$\geq 3^n$	$\left(\frac{3}{2}\right)^n$	$\left(\frac{2}{\sqrt{3}}\right)^n$	$FS > KJ > EI$
Полный четырех-сторонник $(4_3,\ 6_2)$	$\geq 3^n$	$\left(\frac{4}{3}\right)^n$	$\left(\frac{3}{\sqrt{6}}\right)^n$	$FS > KJ > EI$
K_m при $m > 4$ $\left(m_{m-1},\ \binom{m}{2}_2\right)$	$\geq 3^n$	$\left(\frac{m}{m-1}\right)^n$	$\left(\sqrt{\frac{2(m-1)}{m}}\right)^n$	$FS > EI > KJ$
$K_{m,m}$ при $m > 0$ $\left(2m_m,\ m_2^2\right)$	$\geq 2^n$	2^n	1^n	$FS = KJ > EI$
Плоскость Фано (7_3)	$\geq 3^n$	$\left(\frac{7}{6}\right)^n$	$\left(\frac{3}{\sqrt{7}}\right)^n$	$FS > KJ > EI$
Конфигурация Мёбиуса-Кантора (8_3)	$\geq 3^n$	$\left(\frac{4}{3}\right)^n$	$\left(\frac{3}{\sqrt{8}}\right)^n$	$FS > KJ > EI$
Конфигурация Дезарга (10_3)	$\geq 3^n$	$\left(\frac{5}{3}\right)^n$	$\left(\frac{3}{\sqrt{10}}\right)^n$	$FS > KJ > EI$
Конфигурация Гессе $(9_4,\ 12_3)$	$\geq 3^n$	$\left(\frac{9}{8}\right)^n$	$\left(\frac{2}{\sqrt{3}}\right)^n$	$FS > EI > KJ$
Конфигурация Шлефли $(12_5,\ 30_2)$	$\geq 3^n$	$\left(\frac{12}{5}\right)^n$	$\left(\frac{5}{\sqrt{30}}\right)^n$	$FS > KJ > EI$
Проективная плоскость $\left((m^2+m+1)_{m+1}\right)$	$\geq 3^n$	$\left(\frac{m^2+m+1}{m(m+1)}\right)^n$	$\left(\frac{m+1}{\sqrt{m^2+m+1}}\right)^n$	$FS > EI > KJ$
Конфигурация Кокса $\left((2^{m-1})_m\right)$	$\geq 3^n$	$\left(\frac{2^{m-1}}{m(m-1)}\right)^n$	$\left(\frac{m}{\sqrt{2^{m-1}}}\right)^n$	$FS > KJ > EI$

Из таблицы видно, что метод трудного множества всегда работает лучше, чем остальные. В данном случае точная оценка по методу трудного множества может превосходить величину 3^n на некоторых графах, в результате мы не можем доказать, что метод Куликова-Юкны работает всегда хуже. Но зато величины 3^n достаточно для метода энтропийных неравенств, а значит можно сформулировать следующее утверждение:

Утверждение. *Для произвольной геометрической конфигурации (p_γ, l_π) оценка, получаемая по методу трудного множества, превосходит оценку метода энтропийных неравенств. Иначе говоря*

$$3(2) \geq \frac{\gamma}{\sqrt{l}}$$

Доказательство. *Условия*

$$\begin{cases} p \cdot \gamma = l \cdot \pi, \\ p \geq \gamma \cdot (\pi - 1) + 1. \end{cases}$$

должны обязательно выполняться для того, чтобы геометрическая конфигурация была корректно определена.

Используя эти ограничения получаем

$$\frac{\gamma^2}{l} = \frac{\pi \cdot \gamma}{p} \leq \frac{p + \gamma - 1}{p} = 1 + \frac{\gamma - 1}{p} < 4$$

Что и требовалось доказать $\square$.

Теперь давайте сравним метод Куликова-Юкны и метод энтропийных неравенств. Рассмотрим два случая:

- Пусть выполняется условие $l \geq \gamma^2$, тогда

$$\gamma^2 \cdot (\pi - 1) \leq l \cdot (\pi - 1) < p \cdot \gamma \leq p \cdot \sqrt{l}$$

  То есть получаем, что $KJ > EI$.

- Пусть теперь верно $l \leq \gamma^2$, тогда

$$\gamma^2 \cdot (\pi - 1) \geq l \cdot (\pi - 1) = p \cdot \gamma - l \geq p \cdot \sqrt{l} - l$$

  Поделив обе части на $\sqrt{l} \cdot \gamma \cdot (\pi - 1)$ получаем

$$\frac{\gamma}{\sqrt{l}} \geq \frac{p}{\gamma \cdot (\pi - 1)} - \frac{\sqrt{l}}{\gamma \cdot (\pi - 1)}$$

  В итоге получаем, что с некоторой небольшой точностью $EI \gtrsim KJ$

Трудное множество vs Куликов-Юкна

В данном разделе мы докажем, что метод трудного множества всегда дает оценку лучше, чем метод Куликова-Юкны. Также рассмотрим вопрос о величине разницы в получаемых оценках: всегда ли она невелика; или может быть, что для некоторых графов оценки из этих методов будут отличаться очень сильно.

4.1 Теорема Турановского типа

Как уже видно из названия для дальнейших изысканий нам потребуется классическая теорема Турана:

Теорема. *Пусть имеется неориентированный граф* $G = (V, E)$, *где* $|V| = n$ *и число независимости равно* α. *Тогда в графе выполняется следующая оценка*

$$|E| \geq n \cdot \left[\frac{n}{\alpha}\right] - \alpha \cdot \frac{\left[\frac{n}{\alpha}\right] \cdot \left(\left[\frac{n}{\alpha}\right] + 1\right)}{2}$$

Доказательство этой теоремы можно найти в книге [6]. Используя эту теорему мы можем с легкостью доказать следующее:

Теорема. *Пусть имеется неориентированный граф* $G = (V, E)$, *тогда среди ребер максимального паросочетания можно найти трудное множество размера*

$$\left\lceil \frac{v(G)^2}{|E|} \right\rceil$$

Доказательство. *Давайте вместо графа* $G = (V, E)$ *рассмотрим граф* $\widetilde{G} = (\widetilde{V}, \widetilde{E})$, *в котором останутся только вершины из максимального паросочетания. Так как* $|E| \geq |\widetilde{E}|$, *то достаточно найти трудное множество размера*

$$\left\lceil \frac{v(G)^2}{|\widetilde{E}|} \right\rceil$$

Пусть (v_1, v_1'), (v_2, v_2'), $\ldots$, (v_m, v_m') *– ребра максимального паросочетания. Построим граф* $\widehat{G} = (\widehat{V}, \widehat{E})$ *такой, что вершин в нем ровно* m. *Обозначим вершины* $\{\widehat{v}_1,\ \widehat{v}_2,\ \ldots,\ \widehat{v}_m\}$, *причем* $\widehat{v}_i \leftrightarrow (v_i, v_i')$. *Определим множество ребер* $\widehat{E}$ *следующим образом*

$$(\widehat{v}_i, \widehat{v}_j) \in \widehat{E} \text{ если } (v_i, v_i') \notin \widetilde{E} \text{ или } (v_j, v_j') \notin \widetilde{E}$$

Очевидно, что трудное множество на ребрах максимального паросочетания соответствует клике в $\widehat{G}$ такого же размера. Пусть число независимости дополнения графа $\widehat{G}$ равно α, тогда мы можем предъявить трудное множество размера α. Используя теорему Турана для дополнения графа $\widehat{G}$ получаем

$$|\overline{\widehat{E}}| \geq m \cdot \left[\frac{m}{\alpha}\right] - \alpha \cdot \frac{\left[\frac{m}{\alpha}\right] \cdot \left(\left[\frac{m}{\alpha}\right] + 1\right)}{2} =$$

Пусть $m = k \cdot \alpha + r$, где $r < \alpha$

$$= (k \cdot \alpha + r) \cdot k - \alpha \cdot \frac{k \cdot (k+1)}{2} = \frac{\alpha \cdot k^2}{2} + r \cdot k - \frac{\alpha \cdot k}{2}$$

Так как каждое ребро из дополнения графа $\widehat{G}$ порождает два ребра в $\widetilde{G}$, а также еще имеется m ребер самого паросочетания, то получаем

$$|\widetilde{E}| \geq m + 2 \cdot \left(\frac{\alpha \cdot k^2}{2} + r \cdot k - \frac{\alpha \cdot k}{2}\right) =$$

$$= \alpha \cdot k^2 + 2r \cdot k + r \geq \alpha \cdot k^2 + 2r \cdot k + \left\lceil\frac{r^2}{\alpha}\right\rceil = \left\lceil\frac{m^2}{\alpha}\right\rceil$$

В итоге получили, что

$$|\widetilde{E}| \geq \left\lceil\frac{m^2}{\alpha}\right\rceil \Longleftrightarrow \alpha \geq \left\lceil\frac{m^2}{|\widetilde{E}|}\right\rceil = \left\lceil\frac{v(G)^2}{|\widetilde{E}|}\right\rceil \blacksquare$$

Список литературы

[1] Kushilevitz Eyal, Nisan Noam. Communication Complexity. Cambridge University press, 2006.

[2] Razborov Alexander. Communication Complexity. In: An Invitation to Mathematics: from Competitions to Research. Springer, 2011.

[3] Gruber H., Holzer M. Finding lower bounds for nondeterministic state complexity is hard. Springer, 2006.

[4] Jukna S., Kulikov A. S. On covering graphs by complete bipartite subgraphs. Discrete Math, 2009.

[5] Kaced Tarik, Romashchenko A. E., Vereshchagin N. K. Conditional Information Inequalities and Combinatorial Applications. CoRR, 2015.

[6] Оре Ойстин. Теория графов. Наука, 1968.