СПБ, Computer Science Club, 28–19 марта, 4–5 апреля 2009.

Конспект лекций по курсу Экспандеры.[1] А. Ромащенко.

Disclaimer: Автор понимает, что данный текст содержит непонятные места, опечатки и просто ошибки. Все замечания и предложения по доработке конспекта будут с благодарностью учтены. 06/06/2009
email: andrei.romashchenko[AT]gmail.com

1 Определения и несколько примеров применений

В этой главе мы дадим определение экспандеров (расширяющих графов) и приведем несколько иллюстранций использования таких графов.

Определение. Двудольный граф $G = (L, R, E)$ (L и R — левая и правая доли графов, E — множество рёбер) называется (n, m, d)-экспандером (расширяющим графом), если $|L| = n$, $|R| = m$, степень всех вершин в левой доле L равна d, и выполняются следующие свойства *расширения*:

(1) Для любого множества $S \subset L$, $|S| \le \frac{n}{1000d}$ множество соседей (соседи S лежат в R) достаточно велико: $|\Gamma(S)| > \frac{7}{8}d|S|$.

(2) Для любого множества $S \subset L$, такого что $|S| \le \frac{n}{2}$, множество соседей немного больше самого множества S, а именно, $|\Gamma(S)| > \frac{9}{8}|S|$.

Таким образом, экспандер расширяет малые множества из L почти в d раз, а относительно большие множества — хотя бы на небольшую долю. Не следует воспринимать это определение догматически: в приложениях может быть полезно рассмотреть экспандер, в котором константы отличны от $1/(1000d)$, $7/8$ и $9/8$. Мы выбрали такие значения параметров, которые позволяют легко доказать теорему существования и показать несколько примеров применений.

Теорема. Существуют такие n_0, d_0 такие, что для для всех $n \ge n_0, d \ge d_0$ и $m \ge \frac{3}{4}n$ существует (n, m, d)-экспандер.

Доказательство: Выберем граф случайно. Это значит, что для каждой вершины в L мы случайно и независимо выбираем d соседей в R (таким образом, разрешаются параллельные рёбра). Покажем, что с большой вероятностью такой граф оказывается экспандером.

$$\text{Prob[первое свойство экспандера графа нарушено]} \le \sum_{S,T} \left(\frac{|T|}{m}\right)^{sd},$$

где суммирование происходит по всем множествам $S \subset L$ размера не более $n/(1000d)$ и по всем множествам $T \subset R$ размера $(7/8)d|S|$. Оценим данную

[1] При составлении этого конспекта использованы записи, сделанные А. Шенем по лекциям автора.

сумму сверху:

$$\sum_{s=1}^{n/(1000d)} C_n^s \cdot C_m^{\frac{7}{8}ds} \left(\frac{7ds}{8m}\right)^{sd} \le \sum_{s=1}^{n/(1000d)} (ne/s)^s \left(\frac{me}{\frac{7}{8}ds}\right)^{\frac{7}{8}ds} \left(\frac{7ds}{8m}\right)^{sd} \le$$
$$\le \sum_{s=1}^{n/(1000d)} \left[\frac{ne}{s} \cdot \left(\frac{8me}{7ds}\right)^{\frac{7}{8}d} \cdot (\frac{7ds}{8m})^d\right]^s \qquad (1)$$

Мы утверждаем, что данная сумма не превосходит 1/10. Чтобы доказать это, мы оценим эту сумму сверху геометрической прогрессией со знаменателем 1/20. Таким образом, нам нужно показать, что для каждого s произведение в квадратных скобках в (1) не превосходит 1/20:

$$\frac{ne}{s} \cdot \left(\frac{8me}{7ds}\right)^{\frac{7}{8}d} \cdot \left(\frac{7ds}{8m}\right)^d = de^{1+\frac{7}{8}d} \left(\frac{7ds}{8m}\right)^{\frac{d}{8}-1} \le$$
$$\le O(1) \cdot de^{\frac{7}{8}d} \left(\frac{\frac{7}{1000}n}{8 \cdot \frac{3}{4}n}\right)^{\frac{1}{8}d-1} < 1/20$$

(последнее неравенство имеет место для достаточно больших d). Обращаем внимание: мы использовали условие $m \ge 3n/4$.

$$\text{Prob[второе свойство экспандера нарушено]} \le \sum_{S,T} \left(\frac{|T|}{m}\right)^{sd}, \qquad (2)$$

(здесь мы суммируем по всем множествам $S \subset L$ таким, что $|S| \le n/2$ и $T \subset R$ таким, что $|T| = \frac{9}{8}|S|$). Мы оценим (2) сверху:

$$\text{Prob[второе экспандера нарушено]} \le \sum_{s=1}^{n/2} C_n^s \cdot C_m^{\frac{9}{8}s} \cdot \left(\frac{9s/8}{m}\right)^{sd} \le$$
$$\sum_{s=1}^{n/2} \left[(ne/s) \cdot (me/(9s/8)) \cdot \left(\frac{9s/8}{m}\right)^d\right]^s \le 1/10 \qquad (3)$$

Чтобы получить последнее неравенство, мы оцениваем сумму (3) геометрической прогрессией. Для этого достаточно доказать равномерную оценку на выражение в квадратных скобках:

$$(ne/s) \cdot (me/(9s/8)) \cdot \left(\frac{9s/8}{m}\right)^d \le O(1) \cdot \left(\frac{9n/16}{3n/4}\right)^{d-2} = O(1) \cdot (3/4)^{d-2} \le 1/20$$

(здесь мы снова использовали, что $m \ge 3/4n$). Теорема доказана.

Литература: [HLW].

Первый пример применения экспандеров: коды, исправлящие ошибки.

Покажем, как с помощью расширяющего графа посторить линейный код, позволяющий исправлять ошибки в доле $\delta = 1/(2000d)$ битов. Чтобы задать линейный код с длиной кодового слова n, достаточно описать его проверочную матрицу H (слово $x \in \{0,1\}^n$ является кодовым словом если и только если $Hx^t = 0$). Другими словами, нужно задать систему линейных уравнений для переменных $x_1, \ldots, x_n$; решения этой системы и будут кодовыми словами.

Зафиксируем некоторый $(n, 3n/4, d)$-экспандер G. Сопоставим переменные $x_1, \ldots, x_n$ вершинам в левой доле графа. Вершинам из правой доли будут соответствовать уравнения. А именно, каждой вершине v из правой доли G мы сопоставляем уравнение

$$x_{i_1} + \ldots + x_{i_s} = 0 \ (mod\ 2),$$

где $x_{i_1}, \ldots, x_{i_s}$ – это список вершин, соединённых рёбрами с v.

Число уравнений равно $3n/4$, так что размерность пространства решений не меньше $n/4$. Это значит, что в нашем коде будет не менее $2^{n/4}$ кодовых слов.

Остаётся доказать, что данный код действительно исправляет ошибки. Для этого нужно проверить, что расстояние между любыми кодовыми словами не может быть меньше $n/(1000d)$ (на самом деле для расстояния этого кода можно доказать более сильную оценку; но мы ограничимся самым простым рассуждением). Для линейного кода нужное нам условие эквивалентно тому, что в каждом ненулевом кодовом слове должно быть не менее $n/(1000d)$ единиц.

Предположим противное: пусть есть кодовое слово $x_1 \ldots x_n$ (решение системы линейных уравнений, соответствующих графу G), в котором менее $n/(1000d)$ единиц. Обозначим S множество вершин из левой доли графа, соответствующих единицам в данной битовой последовательности. Поскольку $|S| < n/(1000d)$, можно применить первое свойство экспандера: число соседей S достаточно велико

$$|\Gamma(S)| \geq \frac{7d|S|}{8}$$

Из S выходит ровно $d|S|$ вершин. Подсчитаем среднее (по всем вершинам $v \in \Gamma(S)$) число ребер, которое приходит из S в v. Очевидно, это число не превосходит

$$\frac{d|S|}{7d|S|/8} = 8/7 < 2$$

Это значит, что хотя бы у одной вершины v из правой доли есть *ровно один* сосед из S. Но в таком случае уравнение, соответствующее v, не выполненяется на наборе $x_1 \ldots x_n$. Значит, набор битов с менее чем $n/(1000d)$ единицами не может быть кодовым словом.

Для данного кода имеется быстрый алгоритм декодирования. Пусть дан набор битов $y = (y_1, \ldots, y_n)$, который отличается от от некоторого слова $x = (x_1, \ldots, x_n)$ менее чем в $n/2000d$ битов. Мы можем найти x следующим образом. Параллельно для каждого бита $i = 1, \ldots, n$ посчитаем число равных единице контрольных сумм (равенств, задаваемых проверочной матрицей), в которые этот бит входит (если все контрольные суммы равны нулю, то набор битов является кодовым словом); если это число больше $d/2$, меняем данный бит на противоположный. Повторяем эту процедуру коррекции, пока все контрольные суммы не обнулятся.

Покажем, что на каждом шаге описанной процедуры хэмминговское расстояние между кодовым словом x и декодируемым словом y уменьшается в $\Omega(1)$ раз. Таким образом, через $O(\log n)$ итераций все ошибки будут исправлены (т.е. y совпадёт с x, и декодирование будет закончено).

Для простоты обозначений рассмотрми случай $x = (0, \ldots, 0)$. Мы предполагаем, что число единиц в векторе y не превосходит $n/2000d$. Нам нужно показать, что на каждом шаге описанной процедуры число единиц будет уменьшаться не менее, чем в константу раз. Пусть S – множество номеров битов y, в которых стоят единицы; представим S в виде объединения: $S = S_0 \cup S_1$, где в S_0 – позиции, которые обнулятся на следующем шаге; S_1 – позиции, в которых сохранятся единицы. Наконец, обозначим T множество номеров тех позиций, в которых в начале стояли нули, а после применения процедуры коррекции возникла единица. Наша задача показать, что

$$|S_1| + |T| \leq c|S|$$

для некоторой константы $c < 1$ (не зависящей ни от n, ни от размера S).

Сначала покажем, что T содержит менее $|S|/2$ вершин. Предположим противное; выберем из T *ровно* $|S|/2$ вершин и назовём это множество T'. Применим к $S \cup T$ свойство расширения экспандера:

$$\frac{7}{8}d(|S| + |T'|) < |\Gamma(S \cup T')| \leq d|S| + \frac{d}{2}|T'|$$

(второе неравенство следует из того, что у кажой вершины S может быть не более d различных соседей, а у каждой вершины T — не более $d/2$ соседей *не являющихся при этом соседями* S). Сравнивая левую и правую части неравенства, получаем $|T'| < |S|/3$, что противоречит нашему предположению.

Далее, разобъем соседей (из правой доли графа) вершин S на два класса: множество U, у которых имеется ровно по одному соседу в S, и множество V, у которые есть хотя бы два соседа в S. Посчитаем число рёбер между S и $U \cup V$ двумя способами: как число рёбер, выходящих из S, и как число рёбер, выходящих из U и V в S:

$$d|S| \geq |U| + 2|V|$$

(в каждую вершину V входит не менее двух рёбер из S). Таким образом,

$|V| \le (d|S| - |U|)/2$ Применяем к S свойство экспандера:

$$\frac{7}{8}d|S| \le |\Gamma(S)| \le |U| + |V| \le \frac{d|S|}{2} - \frac{|U|}{2}$$

Следовательно, $|U| \ge \frac{3}{4}d|S|$. Разумеется, число рёбер, ведущих из S в U, тоже не меньше $\frac{3}{4}d|S|$. Таким образом, число рёбер из S в V не больше $\frac{1}{4}d|S|$.

Заметим, что не менее половины соседей каждой вершины из S_1 лежат в V. Следовательно,

$$\frac{d}{2}|S_1| \le |E(S_1, V)| \le \frac{1}{4}d|S|,$$

и $|S_1| \le \frac{1}{2}|S|$. Таким образом,

$$|S_1| + |T| \le (1/2 + 1/3)|S|$$

т.е. число 'ошибочных' битов y после операции коррекции уменьшается по крайней мере в $5/6$ раза.

Позднее мы вернёмся к задаче о кодах, исправляющих ошибки, и рассмотрим более сложную и эффективную конструкцию кода, порождаемого экспандером.

Отметим, что для практического применения нам в данном случае нужна *явная* конструкция экспандера – нам необходим алгоритм, который по заданному n за время $\mathrm{poly}(n)$ строит $(n, 3n/4, d)$-экспандер.

Литература: [HLW].

Второй пример применения экспандеров: увеличение вероятности успеха в алгоритмах с датчиком случайных чисел.

Определение Язык L принадлежит сложностному классу RP, если существует полиномиальный алгоритм A такой что

1. для $x \in L$ для всех $r \in \{0,1\}^{poly(n)}$ $A(x, r) = 1$
2. для $x \notin L$ не более чем для $1/2000$ всех $r \in \{0,1\}^{\mathrm{poly}(n)}$ может выполняться $A(x, r) = 1$

Покажем, что для любого $\varepsilon > 0$ полиномиальный вероятностный алгоритм A можно модифицировать таким образом, чтобы вероятность ошибки уменьшилась до ε, а число используемых случайных битов не изменится.

Пусть исходный алгоритм использует $k = k(n)$ случайных битов для вычислений на входах длины n. Зафиксируем $(2^k, 2^k, d)$-экспандер G, где $d > 1/\varepsilon$. Новый алгоритм действует следующим образом: выбирается случайная вершина v из левой доли графа (для этого требуется k случайных битов); затем исходный алгоритм A последовательно запускается на всех d наборах случайных битов, соответствующих соседям вершины v. Если все

полученные ответы равны 1, новый алгоритм также возвращает единицу; в противном случае возвращается ноль.

Покажем, что у нового алгоритма вероятность ошибки не превосходит $1/d$. В самом деле, обозначим $B = B(x)$ множество таких вершин w из правой доли графа, которые соответствуют неверному ответу старго алгоритма на входе x; аналогично, обозначим $S = S(x)$ множество таких вершин v из левой доли графа, которые для которых новый алгоритм даёт неверный ответ на входе x. Очевидно, S состоит из вершин, все соседи которых лежат в B.

Предположим, что S содержит не менее $n/(1000d)$ вершин. Выберем среди них ровно $n/(1000d)$ вершин и назовём это множество S'. По свойству экспандера, имеем

$$|\Gamma(S')| \geq \frac{7}{8} d \frac{n}{1000d} = 7n/8000 > n/2000$$

Это противоречит тому, что все соседи S' лежат в B.

В данном случае нам нужна *явная* в более сильном (чем в первом примере) смысле конструкция экспандера. Размер графа экспоненциально растёт с увеличенем k, и нам необходим алгоритм, который по заданному номеру вершины v (из левой доли графа) за время $\mathrm{poly}(k)$ находит список номеров всех соседей этой вершины (в правой доле графа).

Литература: [HLW].

Третий пример применения экспандеров: хранение множества со сверхбыстрым запросом элементов.

Мы организуем хранение m-элементного множества $S \subset \{1, \ldots, n\}$ в виде описания X, состоящего из $O(m \log n)$ битов. При этом проверка принадлежности $a \in S$ будет производиться чрезвычайно быстро. А именно, мы построим такой вероятностный алгоритм, который по любому входу a запрашивает из X один (sic!) бит; если этот бит оказывается равным единице, то алгоритм отвечает, что a является элементом S; в противном случае алгоритм говорит, что a множеству не принадлежит. При этом для каждого $a \in \{1, \ldots, n\}$ алгоритм ошибается с вероятностью не более $1/3$.

Чтобы постоить нужное нам хранилище X, мы сначала зафиксируем некотрый экспандер, у которого левая доля L состоит из n вершин, правая R из $k = O(m \log n)$ вершин, степень всех вершин левой доли одинакова и равна некоторому d, и для каждого $A \subset L$ размера не более $2m$

$$|\Gamma(A)| \geq \frac{7}{8} d|A|$$

X будет состоять в разметке вершин правой доли нулями и единицами. Эту разметку нужно выбрать таким образом, чтобы у каждой вершины из S не менее $2/3$ соседей были помечены единицей, а у каждой вершины не из S не менее $2/3$ соседей были помечены нулями.

Вероятностный алгоритм для проверки $a \in S$ работает очевидным образом: выбирается случайная вершина из $\Gamma(a)$ и запрашивается пометка данной вершины (в правой доле графа).

Остаётся объяснить, как построить нужную нам разметку правой доли графа. Будем строить её последовательными приближениями. Сначала пометим всех соседей всех вершин из S единицами, а все остальные вершины – нулями. На данной разметке наш алгоритм с вероятностью 1 возвращает правильный ответ для всех $a \in S$. Однако для a не из S проверояный алгоритм может работать неверно. Обозначим T множество всех таких вершин вне S, у которых более $d/3$ соседей помечены единицей. Поменяем разметку: пометим всех соседей T нулём. Теперь разметка может стать плохой для части вершин из S. Обозначим S' множество всех таких вершин из S, у которых более $d/3$ соседей помечены нулями. Далее, поменяем разметку у всех соседей S' на единицы. После этого может вновь возникнуть множество 'неправильных' вершин вне S, и т.д.

Чтобы доказать, что данный процесс в конце концов сойдётся, нужно показать, что на каждом шаге число 'проблемных' вершин уменьшается в константу раз. Поскольку все шаги аналогичны, достаточно разобрать самый первый: докажем, что T в константу раз меньше, чем S.

Мы воспользуемся тем, что для $S \cup T$ выполнено свойство расширения:

$$(7/8)d(|S| + |T|) \leq |\Gamma(S \cup T)| \leq d|S| + (2/3)d|T|$$

Откуда получаем $|T| \leq 3/5|S|$.

Упражнение: Экспандер с какими параметрами нам нужен в данной конструкции?

Литература: [BMRV].

2 Алгебраические экспандеры.

Нашей задачей является построение 'явной' конструкции экспандеров. Мы ограничимся экспандерами специального вида. Во-первых, мы будем рассматривать *однородные* экспандеры, т.е. такие графы, у которых число вершин в левой и правой доле совпадало. Во-вторых, будем требовать, чтобы граф был симметричен. Это значит, что вершины графа в левой и правой долях можно занумеровать таким образом, что если в графе есть ребро, соединяющее i-ую вершину из левой доли с j-ой вершиной из правой, то должно быть и симметричное ему ребро, ведущее из j-ой вершины левой доли в i-ую вершину правой.

Для однородных и симметричных экспандеров двудольная конструкция графа представляется излишним усложенением. В самом деле, мы можем 'склеить' соответствующие вершины левой и правой доли графа. При этом мы получим неориентированный граф с n вершинами степени d (из каждой вершины выходит d рёбер; допускаются кратные рёбра и петли). Такой граф описывается *матрицей смежности* A, в которой a_{ij} равно числу рёбер,

соединяющих вершины i и j. Эта матрица симметрична; сумма чисел в любой её строке или столбце равна d.

Различные свойства графа удобно описывать в терминах этой матрицы:

- (i, j)-й элемент матрицы A^k есть число путей длины k, идущих из вершины i в вершину j;
- если разделить матрицу A на d, то получится матрица, у которой сумма любой строки и любого столбца равна 1. Умножение на эту матрицу описывает случайное блуждание: если p — вектор-столбец, состоящий из вероятностей, описывающих распределение по вершинам в какой-то момент, то (Ap) — распределение через один шаг случайного блуждания (мы выбираем случайную вершину согласно распределению p и переходим к её соседу, выбрав случайно одно из d рёбер).

Последнее наблюдение показывает, что случайное блуждание по графу (из текущей вершины мы равновероятно переходим по одному из рёбер в другую, и так далее) связано со степенями матрицы A/d: чем ближе эти степени к матрице равномерного перемешивания (в которой все элементы равны $1/n$), тем более равномерно распределен результат случайного блуждания. Изучать степени матрицы естественно в собственном базисе. Заметим, что в терминах собственных чисел матрицы A удобно выражать некоторые комбинаторные свойства графа. Сделаем несколько простых наблюдений:

- матрица A симметрична и потому имеет ортогональный собственный базис над вещественным полем, с вещественными собственными значениями;
- поскольку сумма всех чисел в каждой строке равна d, вектор $(1, 1, \ldots, 1)$ (соответствующий столбец) является собственным вектором и имеет собственное значение d;
- все собственные значения не превосходят d по модулю: поскольку суммы элементов во всех строках матрицы равны d, то максимум модулей собственного вектора при умножении на A увеличивается не более чем в d раз;
- если граф состоит из нескольких компонент связности, то имеется несколько собственных векторов с собственным значением d (для вершин одних компонентах связности берём единицы, для других нули);
- напротив, если граф связен, то собственный вектор со значением d единственный: возьмём максимальную по модулю координату этого вектора (вершину графа), она равна среднему по соседям, и потому во всех соседях должно быть то же значение; то же верно для соседей соседей, и т.д.;

- для двудольного графа имеется собственный вектор со значением $-d$: надо в одной доле взять единицы, а в другой минус единицы;
- если имеется собственный вектор со значением $-d$, то граф имеет двудольную связную компоненту (возьмём максимальную по модулю координату, в её соседях будет то же число с противоположным знаком, и так далее: связная компонента этой вершины делится на две доли).

Литература: [AB, HLW].

Геометрические свойства алгебраического экспандера

Определение. Регулярный граф степени d с n вершинами, у которого все собственные числа кроме одного по абсолютной величине не превосходят αd, будем называть алгебраическим (n, d, α)-экспандером.

Данное определение оправдано, поскольку оценки на собственные числа влекут за собой комбинаторные свойства графа, в той или иной форме означающие "хорошее перемешивание".

Пусть граф G имеет n вершин, степень d и второе по абсолютной величине собственное число αd. Имеет место следующее утверждения:

Лемма о перемешивании [expander mixing lemma] Пусть A и B — произвольные (возможно, пересекающиеся) множества вершин графа. Тогда число $|E(A, B)|$ ребер, ведущих из A в B удовлетворяет такой оценке:

$$\left||E(A,B)| - \frac{d \cdot |A| \cdot |B|}{n}\right| \leq \alpha d \sqrt{|A| \cdot |B|}$$

Доказательство: Обозначим 1_A и 1_B харакетристические векторы множеств A и B (i-ая координата соотвествующего вектора равен единице, если i-ая вершина графа принадлежит A или B соответственно; в противном случае значение координаты равно нулю).

Пусть $e_1, \ldots, e_n$ — ортонормированный собственный базис матрицы M заданного экспандера, а $\lambda_1, \ldots, \lambda_n$ – соответствующие собственные числа. Мы считать, что собственные числа упорядочены по убывания абсолютной величины:

$$|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_n|$$

При этом

$$e_1 = \frac{1}{\sqrt{n}}(1, 1, \ldots, 1),$$

а $\lambda_1 = d$, и $|\lambda_i| \leq \alpha d$ для $i > 1$. Разложим векторы 1_A и 1_B по собственному базиса: $1_A = \sum a_i e_i$, $1_B = \sum b_i e_i$. Если M — матрица графа, то

$$|E(A,B)| = 1_A M 1_B = (\sum_{i=1}^{n} a_i e_i) M (\sum_{i=1}^{n} b_i e_i)$$

Выделим первый член этой суммы:

$$|E(A,B)| = d\frac{|A|}{\sqrt{n}} \cdot \frac{|B|}{\sqrt{n}} + \sum_{i=2}^{n} \lambda_i a_i b_i$$

Таким образом,

$$\left| |E(A,B)| - \frac{d \cdot |A| \cdot |B|}{n} \right| = |\sum_{i=2}^{n} \lambda_i a_i b_i| \leq \alpha d |\sum_{i=1}^{n} a_i b_i| \leq \alpha d \cdot ||1_A||_2 \cdot ||1_B||_2 = \alpha d \cdot \sqrt{|A||B|}$$

и лемма доказана.

Теорема о рёберном расширении. Пусть граф G имеет n вершин, степень d и второе собственное число αd. Тогда для любого множества A из не более чем $n/2$ элементов не менее $\frac{(1-\alpha)d}{2}|A|$ рёбер ведут вовне A.

Доказательство: Второе собственное число матрицы графа M можно вычислить как максимум отношения

$$\frac{|f^t M f|}{||f||^2}$$

по всем векторам с нулевой суммой координат (т.е. максимум по всем векторам из ортогонального дополнения к первому собственному вектору $e_1 = (1, 1, \ldots, 1)$). Для доказательства теоремы нам нужно проверить, что максимум этого отношения не меньше $d - 2h_v(G)$, где $h_v(G)$ — коэффициент рёберного расширения:

$$h_v(G) \min_{A \subset V,\ |A| \leq n/2} \frac{|E(A, \bar{A})|}{|A|}$$

Пусть A есть множество вершин графа (размера не более $n/2$). Обозначим 1_A и $1_{\bar{A}}$ характеристические векторы самого множества A и его дополнения. Рассмотрим вектор

$$f = |\bar{A}|1_A - |A|1_{\bar{A}}$$

сумма координат которого равна нулю. Его норма

$$||f||^2 = |\bar{A}|^2 \cdot |A| + |A|^2 \cdot |\bar{A}| = |A| \cdot |\bar{A}| \cdot n$$

Далее,

$$f^t M f = 2|E(A,A)| \cdot |\bar{A}|^2 + 2|E(\bar{A},\bar{A})| \cdot |A|^2 - 2|E(A,\bar{A})| \cdot |A| \cdot |\bar{A}|$$

Нетрудно посчитать, что данное выражение равно

$$dn|A| \cdot |\bar{A}| - |E(A,\bar{A})|n^2$$

Следовательно, второе собственное число M не может быть меньше

$$\frac{dn|A| \cdot |\bar{A}| - |E(A,\bar{A})|n^2}{n|A||\bar{A}|} = d - \frac{n|E(A,\bar{A})|}{|A||\bar{A}|} \geq d - \frac{2|E(A,\bar{A})|}{|A|}$$

и теорема доказана.

Следствие: Диаметр всякого (n, d, α)-экспандера равен $O(\log n)$

Доказательство: Заметим, что у множества $A \subset V$ размера не более $n/2$ должно быть не менее $\frac{(1-\alpha)}{2}|A|$ соседей (поскольку в каждую из соседних вершин может входить не более d рёбер из A).

Пусть x и y — две вершины экспандера. Рассмотрим 'шары' с центрами в этих точках, т.е. множества вершин графа, удалённые от x и y соотвтественно на расстояние не более r. Пока эти шары содержат менее $n/2$ вершин, каждое увеличении радиуса r на единицу увеличивает число точек в шарах не менее чем в $(1-\alpha)/2$ раз. Следовательно, при $r = O(\log n)$ размеры шаров превысят $n/2$, и они пересекутся.

Отметим, что верхняя число α определяет константу в O-большом.

Теорема. Для любого d в d-регулярных графах с n вершинами второе по абсолютной величине собственное число ограничено снизу: $\lambda \geq 2\sqrt{d-1} - o(1)$ при $n \to \infty$.

Доказательство: Обозначим $\lambda(G)$ второе по абсолютное величине собственное число d-регулярного графа G. Чтобы вычислить $\lambda(G)$, рассмотрим степень графа (l-ая степень графа G есть граф с тем же множеством вершин; ребрами же становятся пути длины l в исходном графе). Матрица l-ой степени графа есть l-ая степень матрицы исходного графа; собственные числа при этом тоже возводятся в l-ую степень.

Мы рассмотрим некоторую чётную степень графа $l = 2k$. Напомним, что для оценки второго собственного значение симметричной матрицы, надо ограничить соответствующую ей квадратичную форму на ортогональное дополнение к первому собственному вектору $e = (1, \ldots, 1)$ и взять её максимум на единичном шаре. Таким образом,

$$\lambda(G)^{2k} = \lambda(G^{2k}) \geq \frac{f^t M^{2k} f}{||f||^2}$$

для любого вектора f с нулевой суммой координат. В качестве f мы берём вектор вида

$$f = (0, 0, \ldots, 0, 1, 0, \ldots, 0, -1, 0, \ldots)$$

У этого вектора ровно две ненулевые координаты (i-ая и j-ая). Вершины i и j мы выбираем так, чтобы расстояние между ними было максимальным возможным (т.е., равно диаметру графа G).

Для выбранного вектора f имеем

$f^t M^{2k} f =$ [число $(2k)$-путей из i в i]+[число $(2k)$-путей из j в j]−[число $(2k)$-путей из i в j]

Теперь выбираем $k = \lfloor \frac{diameter(G)-1}{2} \rfloor$; поскольку расстояние между i и j больше $2k$, число $(2k)$-путей из i в j равно нулю. Остаётся оценить число циклов с началами и концами в i и в j. Мы оценим число циклов в графе G снизу через число циклов в дереве степени d (такие циклы соответствуют

циклам в G, которые можно "стянуть" в вершину по рёбрам графа). Таким образом,

$\lambda^{2k} \geq$ числа путей длины $2k$ с началом и концом в корне в *дереве степени* d

А число циклов в регулярном дереве равно

$$[k\text{-ое число Каталана}] \cdot (d-1)^k = \frac{C_{2k}^k}{k+1}(d-1)^k = \frac{2^{2k}}{\mathrm{poly}(k)} \cdot (d-1)^k$$

Действительно, $2k$ рёбер цикла в дереве делятся на шаги, на которых мы удаляемся от корня, и шаги, на которых мы приближаемся к корню; каждый раз, когда мы удалаяемся от корня, мы выбираем одно ребро из $(d-1)$; число способов разделить $2k$ шагов на шаги, на которых мы приближаемся к корню, и шаги, на которых мы удаляемся от корня (число правильных скобочных структур) равно числу Каталана. Таким образом,

$$\lambda \geq 2\sqrt{d-1}\left(\frac{1}{\mathrm{poly}(k)}\right)^{1/2k}$$

Остаётся заметить, что k (диаметр графа, деленный пополам) стремится к бесконечности при росте числа вершин графа n.

Литература: [HLW, AS].

Алгебраические экспандеры: существование

Мы хотим доказать, что бывают графы, у которых второе по абсолютной величине собственное число мало по сравнению с d. Мы докажем, что при определённом соотношении между числом вершин и числом рёбер почти все однородные графы (при некотором распределении вероятностей) обладают таким свойством.

Опишем распределение вероятностей. Будем считать, что n (число вершин) чётно. Тогда на n вершинах существуют графы степени 1 (вершины разбиваются на пары, соединённые ребром). Будем считать все такие паросочетания равновероятными и складывать d случайно выбранных графов степени 1 (при этом мы учитываем кратность параллельных рёбер, если таковые возникнут). Получим некоторое распределение вероятностей на графах степени d с n вершинами.

Мы обозначаем собственные числа полученного графа λ_i и считаем, что

$$d = |\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_n|.$$

Второе собственное число λ_2 будем оценивать так. При возведении матрицы в степень (мы выберем десятую степень) все собственные числа возводятся в ту же степень и след матрицы станет равным

$$\lambda_1^{10} + \lambda_2^{10} + \ldots + \lambda_n^{10}.$$

Первое слагаемое равно d^{10}; если для какой-то матрицы вся сумма близка к d^{10}, то все остальные слагаемы малы. А существование такой матрицы будет доказано, если мы убедимся что *среднее* значение следа матрицы A^{10} (когда A выбирается случайно описанным выше способом) близко к d^{10}.

След A^{10} равен сумме диагональных элементов, поэтому его среднее значение равно среднему значению одного элемента, умноженному на n. А среднее значение одного элемента равно среднему числу путей длины 10, начинающихся и кончающихся в данной вершине. Так что нам надо доказать, что среднее число таких путей чуть больше d^{10}/n.

Подсчёт удобно интепретировать в терминах вероятностей. Будем считать, что помимо d паросочетаний $P_1, \ldots, P_d$ (каждое из которых выбирается независимо, причём все паросочетания равновероятны) мы отдельно (и независимо) выбираем 10 чисел $\omega = (\omega_1, \ldots, \omega_{10})$ от 1 до d. После этого мы (для фиксированной вершины графа) строим путь длины 10, выходящий из этой вершины. На первом шаге он идёт вдоль паросочетания P_{ω_1}, на втором — вдоль P_{ω_2}, и так далее. Нас интересует вероятность того, что после 10 шагов мы вернёмся в исходную точку: нужно доказать, что она равна $1/n(1 + o(1))$.

Поменяем порядок усреднения: если усреднять сначала по выбору ω_i, то получается число петель длины 10 (делённое на d^{10}), которое затем можно усреднять по выбору P_i. Мы же оцениваем в другом порядке: сначала для данного набора ω_i мы усредняем по всем графам, и лишь потом усредняем по наборам.

Все наборы $\omega_1, \ldots, \omega_{10}$ делятся на три категории:

- гарантированно приводящие в исходную точку (независимо от выбора P_d); более точно, к этой категории относятся наборы, в которые после сокращений подряд идущих равных чисел ничего не остаётся;
- наборы, состоящие из десяти разных чисел.
- наборы, которые сокращаются не полностью, но в которых присутствуют равные числа (мы идём несколько раз по одному и тому же паросочетанию, но не обязательно подряд).

Для каждого типа мы оцениваем количество таких наборов, а также (для каждого набора) вероятность замкнутого пути при случайном выборе паросочетаний.

- Количество наборов первого типа не превосходит $O(d^5)$. В самом деле, есть некоторое (фиксированное, так как число 10 фиксировано) число способов сокращения, и для каждого способа имеется не более d^5 способов его реализации (пять сокращающихся пар). Вероятность замкнутого пути равна 1.
- Наборы без повторений составляют большинство (при достаточно больших d) из общего числа d^{10}, и вероятность замкнуться на последнем

шаге (при любом выборе предыдущих паросочетаний, а потому и в целом) есть $1/(n-1) = 1/n(1+o(1))$. (Мы пользуемся лишь тем, что последнее число в наборе ранее не встречалось.)

- Количество наборов второго типа есть $O(d^9)$, где константа в O-обозначении соответствует числу возможных пар позиций, где происходит совпадение (то есть $10 \cdot 9/2$).

  Докажем, что вероятность вернуться в исходную точку для такого набора есть $O(1/n)$. Разобьём это событие на случаи в зависимости от того, когда путь в первый раз возвращается в уже пройденную вершину и того, какой эта вершина была по счёту. Случаев снова будет не больше $10 \cdot 9/2$, так что достаточно рассмотреть вероятность одного из них.

  В момент перед назначенным возвращением в уже пройденную вершину уже фиксированы некоторые рёбра некоторых паросочетаний (те, что использованы в пути), а следующее ребро (по которому мы должны попасть в уже посещённую вершину) ещё не фиксировано. Поэтому для его конца остаётся не менее $n-10$ вариантов, и вероятность выбрать один из них не больше $1/(n-10) = O(1/n)$. (Говоря формально, мы разбиваем событие на части, соответствующие использованным до последнего шага перед возвращением рёбрам, и в каждой части оцениваем условную вероятность.)

Осталось сложить все вероятности. Получаем

$$O\left(\frac{1}{d^5}\right) \cdot 1 + O\left(\frac{1}{d}\right) \cdot O\left(\frac{1}{n}\right) + 1 \cdot \frac{1}{n}(1+o(1)).$$

Если $n \sim d^4$, то последний член основной, и потому среднее значение следа есть

$$n \cdot d^{10} \cdot \frac{1}{n}(1+o(1)) = d^{10}(1+o(1).$$

Следовательно, существуют и даже образуют большинство графы, у которых след десятой степени матрицы близок к d^{10} и потому все собственные числа (кроме первого) суть $o(d)$. Таким образом, доказана

Теорема. Пусть $\gamma > 0$ — произвольное число. Тогда для достаточно больших d существует граф с $n = d^4$ вершинами степени d, у которого все собственные числа, кроме первого d, не превосходят γd по модулю.

Упраженение. Докажите аналогичное утверждение для $n = d^8$.

Замечание. Аналогичное рассуждение позволяет построить графы с фиксированным d и возрастающим n, у которых второе собственное число мало по сравнению с d. Для этого следует рассматривать не фиксированную (десятую или какую-либо ещё) степень матрицы, а степень порядка $k = \log_{d/2}(n^2)$. Таким образом можно доказать, что у большинства графов второе собственное число имеет порядок $O(d^{3/4})$ (Broder–Shamir). Намного более сложные рассуждения (Joel Friedman) позволяют доказать, что

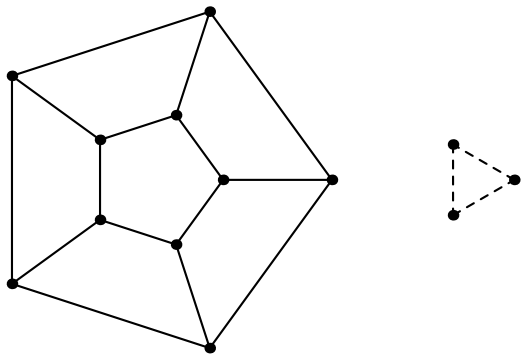
у большинства d-регулярных графов второе собственное число близко к $2\sqrt{d-1}$.

Литература: [HLW].

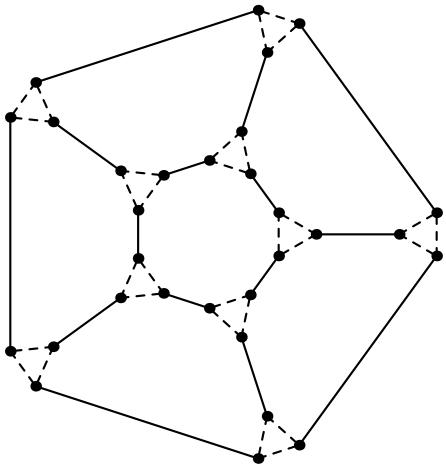
3 Явные конструкции графов: подстановочное и зигзаг-произведение

Долгое время явные конструкции экспандеров были довольно сложными. Оценки собственных чисел использовали преобразование Фурье или сложные конструкции из алгебры и теории представлений. Однако позже был найден довольно простой способ получать такие графы с помощью особых "произведений" графов. Эти произведения позволяют "собирать" большие алгебраические экспандеры из маленьких блоков (а подходящего вида маленькие блоки, которые и сами должны быть эспандерами, мы можем найти перебором.)

Пусть даны два графа $G(n, D)$ и $H(D, d)$. Запись в скобках указывает параметры: число вершин и степень каждой вершины (одинаковую для всех вершин). Пусть при этом число вершин второго графа равно степени первого (как на рисунке).



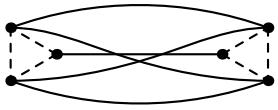
Мы определим три различных *произведения* для этих графов. Для этого каждую вершину первого графа заменим маленькой копией второго графа, прикрепив рёбра первого графа к вершинам второго. (В маленьком графе как раз нужное число вершин.) Внимание: в прикреплении есть произвол — конкретный выбор соответствия в каждой вершине не играет роли. Получится граф с nD вершинами и рёбрами двух типов — большими из первого графа и малыми из второго. (На рисунке — сплошные и пунктирные линии.)



Простое подстановочное произведение (replacement product): Построенный выше граф в точности и есть простое подстановочное произведение G и H. В полученном графе nD вершин (n 'галактик' по D вершин в каждой); степень каждой вершине равна $d+1$ (из каждой вершины выходит одно сплошное и d пунктирных рёбер).
Сбалансированное подстановочное произведение (balanced replacement product): Отличие состоит лишь в том, что мы берём каждое сплошное ребро с кратностью d. таким образом, в полученном графе-произведении из каждой вершины выходит $2d$ рёбер: d сплошных и d пунктирных.
Зигзаг-произведение (zig-zag product): Вершины у зигзаг-произведения будут те же, но рёбра совсем другие. Мы берём все пути дины 3 вида

пунктир – сплошной – пунктир

и объявляем их рёбрами зигзаг-произведения. Другими словами, каждое сплошное ребро порождает d^2 рёбер зигзаг-произведения (соединяющих пары пунктир-соседей концов сплошного ребра), как показано на рисунке (рёбра зигзаг-произведения кривые):



Легко видеть, что все вершины зигзаг-произведения имеют степень d^2 (каждое из двух пунктирных рёбер можно выбрать d способами).

Опишем матрицу графа, полученного в результате зигзаг-произведения. Для этого мы рассмотрим две матрицы размера $nD \times nD$ (координаты соответствуют вершинам графа). Первая матрица $\tilde{H}$ есть матрица графа с рёбрами из пунктирных линий, вторая матрица $\tilde{G}$ соответстует графу с теми же вершинами, но с рёбрами из сплошных линий. (Обозначения показывают, что эти матрицы происходят из соответственно первого и второго

графов, участвующих в зигзаг-произведении). В $\tilde{H}$ каждый столбец и каждая строка содержат d единиц, а в $\tilde{G}$ — только одну единицу. Отметим, что $\tilde{G}$ задаёт перестановку на множестве вершин, и её норма равна единице. Ясно, что матрицей зигзаг-произведения будет произведение матриц $\tilde{H}\tilde{G}\tilde{H}$.

Далее мы докажем оценки для второго собственного числа зигзаг-произведения и сбалансированного подстановочного произведения.

Зигзаг-произведение: оценка второго собственного числа

Докажем, что зигзаг-произведение двух графов, у которых малы вторые (по абсолютной величине) собственные числа, тоже имеет небольшое второе собственное число.

Теорема. Зигзаг-произведение алгебраического (n, D, α)-экспандера G и алгебраического (D, d, β)-экспандера H является алгебраическим экспандером с параметрами $(nD, d^2, \leq \alpha + 2\beta + \beta^2)$.

Замечание. На самом деле можно улучшить оценку для второго собственного числа до $\alpha+\beta$, но мы ограничимся доказательством более слабого утверждения.

Доказательство Чтобы оценить второе собственное значение симметричной матрицы, надо ограничить квадратичную форму на ортогональное дополнение к первому собственному вектору $e_0 = (1, \ldots, 1)$ и взять её максимум на единичном шаре. Другими словами, модуль второго собственного числа матрицы $\tilde{H}\tilde{G}\tilde{H}$ есть максимум выражения

$$|f^\top \tilde{H}\tilde{G}\tilde{H} f|$$

по всем векторам f длины nD, имеющим единичную длину и ортогональных $_0$ (то есть имеющих нулевую сумму координат). Чтобы оценить это выражение, разложим f в сумму $f = f_\parallel + f_\perp$ следующим образом: координаты $f_\parallel$ одинаковы в каждой из 'галактик' (копий графа H), а для $f_\perp$ на каждой копии графа H сумма координат равна нулю. Чтобы оценить значение квадратичной формы на произвольном векторе $f \perp e_0$, мы отдельно изуим действие матриц $\tilde{G}$ и $\tilde{H}$ на векторы $f_\parallel$ и $f_\perp$:

(a) $\tilde{H}g = d \cdot g$, поскольку внутри каждой копии H веса (координаты) вектора g одинаковы, и каждый из них распространяется в d соседей в той же галактике.

(b) $\|\tilde{H}h\| \leq \beta d \cdot \|h\|$, поскольку вектор h в каждой копии H ортогонален первому собственному вектору матрицы графа H, а все остальные собственные векторы этой матрицы не превосходят (по модулю) βd. (Если в каждой компоненте норма оператора не превосходит βd, то это верно и для всего оператора.)

(c) $|g^\top \tilde{G} g| \leq \alpha \|g\|^2$; в самом деле, квадратичная форма в левой части содержит один ненулевой член для каждого ребра, позаимствованного

из графа G (теперь это ребро соединяет две вершины из разных галактик). Поэтому выражение внутри знака модуля из левой части равно $(\hat{g})^\top \hat{G}\hat{g}$, где $\hat{G}$ — матрица смежности графа G, а $\hat{g}$ получается из g склеиванием равных значений в каждой компоненте. При этом сумма координат в $\hat{g}$ (как и в исходном векторе f) равна нулю. Поэтому данное выражение (по предположению о графе G) оценивается как $\alpha m\|\hat{g}\|^2$, что равно как раз $\alpha\|g\|^2$.

Теперь мы можем оценить искомое выражение:

$$|f^\top \tilde{H}\tilde{G}\tilde{H}f| = |(g+h)^\top \tilde{H}\tilde{G}\tilde{H}(g+h)| \le$$
$$\le |g^\top \tilde{H}\tilde{G}\tilde{H}g| + 2|g^\top \tilde{H}\tilde{G}\tilde{H}h| + |h^\top \tilde{H}\tilde{G}\tilde{H}h|.$$

Первое из трёх слагаемых равно $d^2|g^\top \tilde{G}g|$ по свойству (a) и потому не превосходит $\alpha d^2\|g\|^2$ по (b). Второе слагаемое не превосходит $2\cdot(\beta d)\cdot d\cdot\|g\|\cdot\|h\|$ (матрица $\tilde{G}$ есть матрица перестановки и сохраняет норму). Наконец, третье слагаемое не превосходит $(\beta d)^2\|h\|^2$ по аналогичным причинам. В этих оценках можно заменить $\|g\|$ и $\|h\|$ на $\|f\|$ и получить

$$(\alpha + 2\beta + \beta^2)\|f\|^2,$$

что и требовалось. Теорема доказана.

Упражнение: Улучшите оценку до $(\alpha + \beta + \beta^2)\|f\|^2$.

Эффективное построение экспандеров

Построим последовательность явно заданных графов одной и той же степени с растущим числом вершин, имеющих малые собственные числа. Основная идея: возводя матрицу в квадрат, мы не меняем число вершин графа и уменьшаем (возводим в квадрат) нормализованное (т.е. делённое на степень графа) второе собственное число. Зато мы увеличиваем (тоже возводим в квадрат) степень вершины. Но это можно скомпенсировать зигзаг-умножением на фиксированный граф H.

Более подробно, фиксируем граф $H(d^4, d, 1/10)$ для некоторого d (для достаточно больши**х** d такой граф, как мы видели, существует). Затем построим последовательность графов $G_0, G_1, \ldots$, положив

- $G_0 = H^2$. Параметры: $(d^4, d^2, 1/100)$.
- G_{n+1} есть зигзаг-произведение G_n^2 и H. По индукции доказывается, что экспандер G_n имеет параметры $(d^{4n+4}, d^2, \le 1/2)$. В самом деле, после возведения в квадрат получаем $(d^{4n+4}, d^4, 1/4)$, а умножение на $H(d^4, d, 1/10)$ даёт число вершин d^{4n+8}, степень каждой вершины d^2 и третий параметр
$$\frac{1}{4} + 2\frac{1}{10} + \frac{1}{10^2} < \frac{1}{2},$$
что завершает доказательство.

В каком смысле мы получаем явно заданную последовательность графов? Эти графы можно выписать за полиномиальное время от числа вершин. Но в приложениях нам может понадобиться явно заданная последовательность в более сильном смысле: кодируя вершины последовательностями битов, можно требовать полиномиальной вычислимости функции *вращения* $\langle x, i\rangle \mapsto N(x, i)$, где x — код вершины, i — порядковый номер ребра, $N(x, i)$ — код второго конца этого ребра. В этом случае время вычисления должно быть полиномиально относительно длины кода вершины, то есть логарифма от числа вершин – это существенно более сильное свойство.

Как происходит вычисление функции N в построенной последовательности? Номер ребра представляет собой пару чисел, каждое от 1 до d. Вершина графа G_{n+1} представляет собой пару: одна вершина G_n^2 (=вершина G_n) и одна вершина H. Движение по ребру: сначала идём по ребру H, попадаем в какую-то вершину H (в диапазоне $1 \dots d^4$), воспринимаем её как пару чисел в диапазоне $1 \dots d^2$, рекурсивно идём по двум рёбрам графа G_n и затем делаем ещё ход в H. Таким образом, вычисление N для графа G_{n+1} использует два вызова аналогичного вычисления для G_n, что приводит к экспоненциальной оценке по n, и полиномиальной вычислимости функции N не получается.

Однако можно модифицировать конструкцию, используя не предыдущий граф G_n, а граф с половинным индексом. Для начала выберем граф H с параметрами $(d^8, d, 1/10)$, а затем построим последовательность

$$
\begin{aligned}
&G_0\colon (1, d^2, 1/2)\\
&G_1\colon (d^8, d^2, 1/2)\\
&G_2\colon (d^{16}, d^2, 1/2)\\
&\quad\dots\\
&G_n\colon (d^{8n}, d^2, 1/2)\\
&\quad\dots
\end{aligned}
$$

Начальные графы G_0 и G_1 построить легко (G_0 — это граф, состоящий из единственной вершины и d^2 петель, граф G_1 можно получить из H размножением рёбер в d раз, что не меняет собственных чисел). Затем можно воспользоваться рекуррентной формулой

$$G_n = (G_{\lfloor (n-1)/2\rfloor} \otimes G_{\lceil (n-1)/2\rceil})^2 ⓩ H,$$

где $\otimes$ обозначает тензорное произведение, а ⓩ — зигзаг-произведение. Тензорное произведение в скобках имеет параметры $(d^{8(n-1)}, d^4, 1/2)$, после возведения в квадрат получается $(d^{8(n-1)}, d^8, 1/4)$ и после зигзаг произведение $(d^{8n}, d^2, 1/2)$ (в силу того же вычисления с $1/4$ и $1/10$, что и раньше).

Преимущество новой конструкции в том, что при вычислении функции ‘вращения’ N два рекурсивных вызова относятся к половинным значениям n; глубина рекурсии теперь стала логарифмической по n, и общее время вычисления полиномиально по n.

Оценка собственных чисел для сбалансированного подстановочного произведения

Докажем немного другую оценку на собственные числа: будем оценивать не их малость второго собственного числа, а его отделённость от единицы. Для разнообразия будем иметь дело с подстановочным произведением (хотя аналогичная оценка возможна и для зигзаг-произведения).

Теорема. Пусть графы G и H являются алгебраическими экспандерами с параметрами $(n, D, 1-\varepsilon)$ и $(D, d, 1-\delta)$ соответственно. Тогда их сбалансированное подстановочное произведение $G\textcircled{r}H$ имеет параметры $(nD, 2d, 1-\varepsilon\delta^2/24)$.

Доказательство. Удобно описывать происходящее в терминах блужданий (соответствующих нормализованным матрицам, полученных делением матрицы смежности на степень графа). Блуждание по взвешенному произведению является полусуммой двух блужданий: 'локального', где мы движемся внутри одной 'галактики' в соответствии с матрицей графа H, и глобального, где мы движемся по рёбрам графа G (а выбор ребра определяется текущей H-координатой: вершин в графе H как раз столько, сколько рёбер в G, и мы предполагаем, что фиксировано какое-то соответствие). Таким образом, матрицу блуждания можно записать как

$$U = \frac{1}{2}\hat{G} + \frac{1}{2}\hat{H},$$

где $\hat{G}$ и $\hat{H}$ — матрицы перехода по 'локальным' и 'глобальным' рёбрам. Чтобы оценить второе собственное число U, достаточно оценить второе собственное число $U^3 = (\frac{1}{2}\hat{G} + \frac{1}{2}\hat{H})^3$ и доказать, что оно не больше $1 - \varepsilon\delta^2/8$ (а затем воспользоваться неравенством Бернулли).

В разложении для U^3 будет восемь членов. Все эти члены имеют два инвариантных подпространства: одномерное — векторы, у которых все координаты равны (все восемь членов на этом подпространстве единичные, и при каждом стоит коэффициент $1/8$), и ортогональное к нему (векторы, сумма координат которых равна нулю), где максимальное собственное значение (и тем самым норма ограничения на это подпространство) и есть интересующий нас параметр. Если бы мы доказали, что для одного из этих восьми произведений второе собственное число не больше $1-\varepsilon\delta^2$, то это бы гарантировало, что для U^3 это второе собственное число не больше $1-\varepsilon\delta^2/8$, поскольку у оставшихся семи произведений норма не больше 1.

Какое из восьми слагаемых выбрать? Кажется, что наилучшие шансы на перемешивание у $\hat{H}\hat{G}\hat{H}$ (сначала перемешиваем внутри галактики с помощью графа H, потом идём по ребру большого графа, потом перемешиваем внутри другой галактики — как в зизаг-произведении). Если бы перемешивание внутри галактики было полным (переход в случайную точку галактики), то такой переход был бы переходом в случайную вершину случайной соседней галактики. Соответствующее преобразование является тензорным произведением G и полного перемешивания, и потому имеет второе собственное число $1-\varepsilon$, как у G.

Это много лучше, чем нам надо ($1-\varepsilon$ вместо $1-\varepsilon\delta^2/8$), что не удивительно: мы самовольно заменили перемешивание вдоль H полным перемешиванием. Поскольку переход по ребрам H не является полным перемешиванием, нам придётся усложнить рассуждение (при этом вместо $1-\varepsilon$ мы получим для второго собственного числа более скромную оценку $1-\varepsilon\delta^2/8$).

Лемма. Пусть S — матрица блуждания по некоторому графу (первое собственное число равно 1), и все остальные собственные числа по модулю не превосходят $1-\delta$. Тогда S можно представить в виде $(1-\delta)A+\delta B$, где A — матрица с нормой не больше 1, а B — матрица полного перемешивания (все матричные элементы равны 1/(число вершин)).

Доказательство: вычитая из S матрицу δB, мы уменьшаем первое собственное число (единицу) на δ, а остальные не трогаем, так что все собственные числа становятся не больше $1-\delta$ по модулю.

Теперь мы можем повторить примерно те же рассуждения, что и выше, но уже с тремя слагаемыми. Применим лемму к блужданию по графу H и разложим его в сумму $(1-\delta)A+\delta B$. Это разложение можно провести в каждой галактике и получить разложение $\hat H=(1-\delta)\hat A+\delta\hat B$, где $\hat A$ — некоторый оператор с нормой не больше 1, а B — то самое полное перемешивание внутри галактик, о котором мы говорили выше.

Повторяем прежнее рассуждение, но в разложении

$$U=\frac{1}{2}\hat G+\frac{1-\delta}{2}\hat A+\frac{\delta}{2}\hat B$$

будет уже не 8, а 27 слагаемых. В одном из слагаемых (в $\hat B\hat G\hat B$) второе собственное значение не превосходит $1-\varepsilon$, а остальные представляют собой операторы с нормой не больше 1 с некоторыми скалярными коэффициентами (сумма коэффициентов при этих 27 слагаемых равна единице). Поэтому второе собственное значение U не больше $1-\varepsilon\delta^2/8$, что и требовалось доказать.

Ниже мы воспользуемся конструкцией подстановочного произведения и доказанной только что теоремой, чтобы получить теорему Рейнгольда. А сейчас мы применим полученные оценки чтобы описать ещё одну явную конструкцию алгебраических экспандеров. Прежде всего мы выберем алгебраический экспандер H с параметрами $(d^{50},d/2,1/100)$, а также алгебраические экспандеры G_1 и G_2 с параметрами $(d^{100},d,<1/2)$ и $(d^{200},d,<1/2)$ соответственно (такие графы существуют для всех достаточно больших d). Далее построим последовательность

$$G_n=(G_{\lfloor(n-1)/2\rfloor}\otimes G_{\lceil(n-1)/2\rceil})^{50}\textcircled{r}H,$$

Упражнение: Проверьте, что G_n является алгебраическим экспандером с параметрами $(d^{100n},d,<1-1/50)$; функция вращения для такого графа G_n вычисляется за полиномиальное от n время.

Ещё один вариант явного построения экспандеров

Все явные последовательности экспандеров, которые мы строили, формировались вокруг 'затравочного' экспандера H с подходящими параметрами, который мы находили перебором (длина перебора не зависела от n, так что мы имели право использовать его в полиномиальном по n алгоритме). Сейчас мы опишем алгебраическую конструкцию, которая позволяет строить нужный нам 'затравочный' граф без перебора.

Пусть q – простое число. Рассмотрим граф AP_q, вершинами которого являются все пары $(a,b) \in \mathbb{Z}^2$, а рёбрами соединены такие вершины (a,b), (c,d), что

$$ac = b + d \ mod \ q$$

Полезно представлять себе пару (a,b) точкой аффинной плоскости над $\mathbb{Z}_q$, а (c,d) – прямой с уравнением $y = cx - d$, которая проходит через эту точку.

Таком образом, граф состоит из q^2 вершин, и степень каждой вершины равна q. Покажем, что второе по абсолютной величине собственное число графа равно $\sqrt{q}$.

Можно заранее догадаться, что данный граф обладает хорошими свойствами перемешивания. В самом деле, вторая степень этого графа (соответствующая блужданию по графу AP_q по путям длины два) очень близка к полному перемешиванию. Поэтому удобно произвести спектральный анализ не для самого AP_q, а для его квадрата.

Обозначим A матрицу графа AP_q. Будем считать, что вершины (a,b) нумеруются сначала по первой, а потом по второй координате. Таким образом, матрица A состоит из q^2 квадратных блоков размера $q \times q$; в каждом таком блоке (i-ом по горизонтали, j-ом по вертикали) ребра соответствуют переходу из вершин вида $(i,*)$ в вершины $(j,*)$.

Матрицу A^2 легко выписать в явном виде. Действительно, A^2 описывает пути длины 2 на AP_q. Если $i \neq j$, то есть ровно один такой путь из (i,k) в (j,l) (поскольку на плоскости есть ровно одна прямая, которая проходит через точки (i,k) и (j,l)). Если $k \neq l$, то из (i,k) в (i,l) нет путей длины два (мы не рассматриваем вертикальные прямые на плоскости). Наконец, из для каждой вершины (i,k) имеется q циклов длины два.

Таким образом, матрица A^2 имеет вид

$$\begin{pmatrix} qI & J & J & \dots & J \\ J & qI & J & \dots & J \\ \dots & \dots & \dots & \dots & \dots \\ J & J & J & \dots & qI \end{pmatrix}$$

где I — диагональная единичная матрица $q \times q$, а J_q — матрица $q \times q$, в которой на всех местах стоят единицы.

В тензорных обозначениях это можно записать так:

$$A^2 = I_{q\times q} \otimes (qI_{q\times q}) + (J_{q\times q} - I_{q\times q}) \otimes J_{q\times q}$$

У матрицы $I_{q\times q}$ все собственные числа равны единице; у $J_{q\times q}$ есть собственное число 1 кратности один и собственное число 0 кратности $(q-1)$. Простой подсчёт показывает, что у A^2 спектр состоит из чисел q^2 (кратность 1), 0 (кратность $(q-1)$) и q (кратность $q(q-1)$). Следовательно, у самой матрицы A второе собственное число равно $\sqrt{q}$.

Зафиксируем простое число q и рассмотрим следующую последовательность графов:

$$\begin{aligned} AP^1 &= AP_q \ \otimes \ AP_i \\ AP^{k+1} &= AP^k \ \textcircled{z} \ AP_q \end{aligned}$$

По свойству зигзаг-произведения, AP^k является алгебраическим экспандером с параметрами $(q^{2(k+1)}, q^2, O(\frac{i}{\sqrt{q}}))$. Таком образом, при $k = 7$ (для достаточно больших q) мы получаем граф, который можно брать в качестве графа H в нашей основной конструкции явно заданных экспандеров.

4 Графы Кэли

В этом разделе мы определим графы Кэли и приведём несколько простейших примеров спектрального анализа таких графов.

Пусть G — произвольная группа, а $S \subset G$ — симметричное множество элементов группы (если $h \in S$, то $h^{-1} \in S$). Графом Кэли (G, S) называется граф, вершинами которого являются все элементы группы G; вершины a и b соединяются ребром, если $a = bh$ для некоторого $h \in S$ (поскольку S симметрично, данное определение корректно).

Пример 0. G– произвольная группа, $S = G$. Графом Кэли (G, S) будет полный графы с $|G|$ вершинами (с петлями).

Пример 1. $G = \mathbb{Z}_n$, $S = \{1, -1\}$. Графом Кэли (G, S) будет цикл длины n.

Пример 2. $G = \mathbb{Z}_2^k$; S состоит из естественных образующих группы: $S = \{e_i = (0, 0, \ldots, 0, 1, 0, \ldots, 0),\ i = 1, \ldots, n\}$ (у e_i единица стоит в позиции номер i; остальные координаты нулевые). Графом Кэли (G, S) будет граф рёбер n-мерного гиперкуба.

Для спектрального анализа графа Кэли полезно рассмотреть неприводимые представления группы G. Для конечных абелевых групп нужно изучить характеры G:

Определение. Характерами группы G называют гомоморфизмы в мультипликативную группу комплексных чисел $\xi : G \to \mathbb{C}^*$

Теорема Пусть $G = \{a_1, \ldots, a_n\}$ — конечная абелева группа, $S \subset G$ — её симметричное подмножество, ξ — один из характеров группы. Тогда вектор $(\xi(a_1), \ldots, \xi(a_n))$ является собственным для матрицы группы Кэли (G, S); соответствующее этому вектору собственное число равно

$$\sum_{h \in S} \xi(h)$$

Доказательство: Подействуем на вектор $(\xi(a_1), \ldots, \xi(a_n))$ матрицей M графа Кэли. Вычислим значение в i-ой координате полученного в результате вектора. Понятно, что там должна стоять сумма величин $\xi(a_j)$ по всем a_j, которые соединены ребром с a_i. Это значит, что a_j получается из a_i умножением на некоторый элемент из S. Таким образом, i-ая координата $M(\xi(a_1), \ldots, \xi(a_n))^t$ равна

$$\sum_{h \in S} \xi(a_i h) = \xi(a_i) \sum_{h \in S} \xi(h)$$

Тем самым, теорема доказана.

Возвращение к примеру 1. Характер группы $\mathbb{Z}_n$ однозначно определяется его значением на элементе 1 (при этом характер должен отобразить элементы группы в корни из единицы степени n). Таким образом, мы получаем n линейно независимых характеров ξ_k, определяемых условием

$$\xi_k(1) = e^{2\pi k/n}$$

Соответственно, собственные числа графа равны $\lambda_k = \xi_k(1) + \xi_k(-1) = 2\cos(2\pi k/n)$, $k = 0, 1, \ldots, n-1$.

Возвращение к примеру 2. Характеры группы $\mathbb{Z}_2^n$ однозначно определяются значениями на образующих e_i ($\xi(e_i)$ может быть равно 1 или -1). Таким

образом, мы имеем 2^n линейно независимых характеров $\xi_{b_1 \dots b_n}$ ($b_1 \dots b_n$ – строка битов):

$$\xi_{b_1 \dots b_n}(a_1, \dots, a_n) = \prod_{i=1}^{n} (-1)^{a_i b_i}$$

Собственные числа графа Кэли равны

$$\lambda_{b_1 \dots b_n} = [\text{число нулей в строке } b_1 \dots b_n] - [\text{число единиц в строке } b_1 \dots b_n]$$

т.е. собственными числами будут значения $n, n-2, n-4, \dots, -n$ (кратности собственных чисел будут равны соответствующему биномиальному коэффициенту).

Графы Рамануджана.

d-регулярный граф называется графом Рамануджана, если его второе по модулю собственное число не превосходит $2\sqrt{d-1}$. Любоцкий, Сарнак, Филлипс и Маргулис указали явную конструкцию графов Кэли, являющихся графами Рамануджана. Опишем эту конструкцию.

Пусть p и q простые числа, $p = 1 \ mod \ 4$ и $q = 1 \ mod \ 4$. В качестве группы G возьмём $PGL(2, \mathbb{Z}/q\mathbb{Z})$, т.е. невырожденные матрицы 2×2 над полем вычетов по модулю q, профакторизованные по отношению пропорциональности (с обычной операцией матричного умножения).

Далее мы зададим в этой группе симметричное множество S. Выберем такое целое i, что $i^2 = -1 \ mod \ q$. Можно доказать, что тогда имеется ровно $(p+1)$ целочисленное решение уравнения

$$a_0^2 + a_1^2 + a_2^2 + a_3^2 = p$$

такое, что a_0 положительно и нечётно, а a_1, a_2, a_3 чётны. Каждой такой четвёрке сопоставим матрицу

$$A = \begin{pmatrix} a_0 + ia_1 & a_2 + ia_3 \\ -a_2 + ia_3 & a_0 - ia_1 \end{pmatrix}$$

Эти матрицы и образуют множество S.

Нетрудно понять, что граф Кэли (G, S) состоит из $\Theta(q^3)$ вершин, и степень каждой вершины равна $(p+1)$. Свойства данного графа зависят от соотношения p и q. Рассмотрим случай, когда p является квадратичным вычетом по модулю q. Тогда полученный граф Кэли состоит из двух связных компонент (поскольку все матрицы из S лежат в подгруппе G индекса два — подгруппе матриц, определитель которых является квадратичным вычетом). Обозначим $X^{p,q}$ связную компоненту полученного графа. Можно доказать, что у $X^{p,q}$ второе по абсолютной величине собственное число не превосходит $2\sqrt{p}$, т.е. мы получили граф Рамануджана. Однако доказательсто этого факта непросто и использует серьёзную алгебру. Полное доказательство этой теоремы можно найти в [Sar].

5 Некоторые применения экспандеров в computer science

Применение экспандеров 1: Алгоритм Рейнгольда

Ранее мы видели, что зигзаг и подстановочное произведение позволяют ‘собирать’ из маленьких экспандеров (маленькие экспандеры можно найти перебором)

сколь угодно большие экспандеры с ограниченной степенью и достаточно малым вторым собственным числом. Теперь мы рассмотрим ещё одно замечательное одно применение этих операций. Мы докажем теорему Рейнгольда (Omer Reingold) о дерандомизации одного из самых знаменитых вероятностных алгоритмов – алгоритма проверки s-t-связности в неориентированном графе (задача UPATH) с логарифмической памятью.

Описание задачи UPATH: задан неориентированный граф $G = (V, E)$, в котором выделены две вершины $s, t \in V$. Требуется выяснить, есть ли в графе путь из вершины s в вершину t.

Теорема Задача UPATH может быть решена вероятностным алгоритмом с логарифмической памятью.

Вероятностный алгоритм для решения задачи UPATH устроен очень просто: нужно сделать $N = \mathrm{poly}(|V|)$ (выбор полинома мы утоним чуть позже) шагов случайного блуждания по графу, начав с вершины s. Если за N шагов нам удастся побывать в вершине t, мы точно знаем, что в графе есть путь из s в t. В противном случае мы полагаем, что такого пути нет.

В каждый момент работы алгоритма нам требуется помнить номер текущего шага блуждания (от 1 до N) и номер вершины, в которой мы в данным момент находимся. Для хранения этой инофрмации достаточно памяти размера $O(\log |V|)$.

Ясно, что если пути из s в t нет, то алгоритм выдаст правильный ответ. Остаётся оценить вероятность другой ошибки: путь из s в t существует, но за N шагов блуждания мы его не обнаружим. Без ограничения общности можно считать, что граф регулярен и недвудолен (мы всегда можем добиться этого, добавив в граф некоторое количество петель). Далее покажем, что при случайном блуждании по связному однородному и недвудольному графу распределение вероятностей на вершинах быстро приближается к однородному. Ключевое свойство графа:

Лемма В d-регулярном однородном и недвудольном графе с n вершинами щель между первым и вторым по абсолютной величине собственными числами не может быть меньше $1/\mathrm{poly}(n)$, т.е.

$$\lambda/d \geq 1 - \Theta(1/n^c)$$

для некоторой константы c (не зависящей ни от n, ни от d).

Упражнение: докажите эту лемму.

С помошью леммы легко оценить корректность работы нашего алгоритма. Обозначим $\bar{p}(i)$ распределение веротяностей на вершинах после i шагов случайного блуждания по графу (распределение $\bar{p}(0)$ сосредоточено в единственной вершине s). Пусть обозначим равномерное распределение $\bar{u} = (\frac{1}{n}, \ldots \frac{1}{n})$ на вершинах компоненты связности s, и разложим $\bar{p}(i)$ в сумму $\bar{u}$ и некоторого вектора из его ортогонального дополнения:

$$\bar{p}(i) = \bar{u} + \bar{q}(i),$$

где сумма координат вектора $\bar{q}(i)$ равна нулю. Если M — нормализованная матрица графа, то $\bar{q}(i+1) = M\bar{q}(i)$. На подпространстве векторов с нулевой суммой норма линейного оператора M равна (нормализованному) второму собственному числу графа; по лемме это число не может быть больше $1 - \Theta(1/n^c)$, где n есть число вершин в компоненте связности вершины t. Следовательно, на каждом шаге норма $\bar{q}(i)$ уменьшается по крайней мере в $(1-\Theta(1/n^c))$ раз, и через $\mathrm{poly}(n)$ шагов распределение $\bar{p}(i)$ станет очень близко к равномерному (на компоненте связности графа). Таким образом, если s и t принадлежат одной компоненте связности, то

вероятность попасть через $\mathrm{poly}(n)$ шагов в вершину t будет близка к $1/n$. Если же увеличить число шагов ещё в полином раз, то веротяность хотя бы раз побывать в t станет близка к единице.

Рейнгольд придумал, как дерандомизовать алгоритм блуждания на графе без значительного увеличения используемой памяти:

Теорема Задача UPATH может быть решена *детерминированным* алгоритмом с логарифмической памятью.

Прежде чем доказывать теорему, заметим, что мы уже умеем решать на логарифмической памяти задачу UPATH для $(n, d, 0.99)$-экспандеров. В самом деле, мы знаем, что диаметр такого экспандера равен $O(\log n)$. Мы можем перебрать все пути длины $C \log n$ с началом в вершине s и проверить, ведёт ли хотя бы один из них в t; такая проверка очевидно требует лишь логарифмической памяти (и полиномиального времени).

Чтобы решить задачу для произвольного графа G, мы превратим его в экспандер с помощью произведения. Для этой цели годится и зигзаг-произведение, и сбалансированное подстановочное произведение. В оригинальной работе Рейнгольда применялось загзаг-произведение (этот способ даёт лучшие оценки для мультипликативной константы в оценке размера памяти алгоритма). Мы приведём немного другое рассжудение, воспользовавшись сбалансированным подстановочным произведением (для этого у нас уже доказана нужная оценка для собственных чисел подстановочного призведения).

Доказательство теоремы: Мы предполагаем, что нам задан (в виде оракула) неориентированный граф G с n вершинами, без петель и параллельных рёбер. Далее мы построим на основе G несколько 'воображаемых' графов; мы сможем моделировать блуждание по каждому из этих вообраемых графов с помощью исходного оракула и дополнительной памяти размера $O(\log n)$.

Воображаемый граф G': заменим в исходном графе каждую вершину степени $deg > 3$ на цикл длины deg; рёбра, входившие ранее в данную вершину мы по одному присоединим к вершинам этого цикла. Таким образом, в графе G' степень каждой вершины не превосходит 3. Обозначим через n' число вершин в G' (это число не превосходит $\mathrm{poly}(n)$).

Воображаемый граф G'': Добавим к каждой из вершин G_1 нужное число петель так, чтобы получился d-регулярный граф (константу d мы выберем так, чтобы существовал алгебраический $(d^{50}, d/2, < 0.01)$-экспандер H).

Воображаемые графы $G_i : G_0 = G''$; каждый следующий граф G_{i+1} определяется рекурсивно:

$$G_{i+1} = (G_i ⓩ H)^{50}$$

При этом каждый граф G_i будет экспандером с параметрами

$$(n' \cdot d^{50i}, d^{50}, < 1 - \varepsilon_i)$$

Если ε_i достаточно мало, то число ε_{i+1} получается из ε_i умножением сначала на $(0.99)^2/24$ (свойство подстановочного произведения), а затем умножением на 50 (для малых x имеем $(1-x)^{50} \approx 1 - 50x$). Таком образом, $\varepsilon_{i+1} \approx 2\varepsilon_i$.

Применим Лемму о втором собственном числе произвольного регулярного графа к G_0: $\varepsilon_0 \geq \Theta(1/(n')^c)$. Следовательно, для $k = O(\log n)$ граф G_i оказывается экспандером, у которого нормализованное второе собственное число не превосходит 0.99.

Вершины G_i получаются как тензорное произведение вершины графа G'' и i копий вершин графа H. Подстановочное произведение устроено так, что вопрос о существовании пути из s в t в исходном графе G эквивалентен вопросу о существовании пути в G_i из вершин, у которых первая тензорная компонента равна s, в вершины, у которых первая тензорная компонента рана t. Поскольку для $k = O(\log n)$ у графа G_i второе собственное число не превосходит 0.99, мы можем проверить данное свойство, перебрав все пути логарифмической длины.

Остаётся заметить, что моделирование блуждания по графу G_k моделируется на логарфимической памяти. В самом деле, для хранения номера вершины G_k нам нужно хранить набор из $(i + 1)$ компонент; самая первое содержит некоторый номер вершины $G_0 = G''$, а каждая следующая — номер одной из вершин H. Ребро в графе G_i есть путь длины 50 в графе $(G_i ⓩ H)$. Остаётся понять, как организовать рекурсивный вызов для моделирования одного шага по ребру $(G_i ⓩ H)$. Если ребро является локальным (соответствует переходу внутри 'галактики' вершин, являющейся копией H), то нам нужно только пересчитать координаты в i-ой компоненте в соответствии с матрицей графа H. Если же ребро соединяет вершины двух соседних галактик, мы рекурсивно вызываем операцию перехода для графа G_{i-1}; (при этом запись в i-ой компоненте текущей вершины нужно интерпретировать как номер ребра, по которому мы выходим из текущей вершины в графе G_{i-1}). Рекурсивный вызов возвращает, во-первых, номер новой вершины v графа G_{i-1} (это содержимое компонент с 0-ой по $(i - 1)$-ую), а также номер ребра графа G_i, по которому мы только что прошли, *с точки зрения вершины* v, в которую мы попали (этот номер мы запишем в i-ую компоненту текущей вершины).

Мы предлагаем читателю убедиться, что организация рекурсии требует лишь $O(1)$ ячеек памяти на каждую компоненту $i = 1, \ldots, i_{\max} = O(\log n)$. Таким образом, вся процедура работает на зоне $O(\log n)$

Литература: [AB, Rei]

Применение экспандеров 2: Коды Земора (Zémor codes)

Будем называть кодом отображение $C : \{0,1\}^k \to \{0,1\}^n$. Образы этого отображения (n-битные строки $C(x)$) называют кодовыми словами. *Расстоянием* кода C называется минимум хэмминговского расстояния между кодовыми словами:

$$dist_C = \min_{x \neq x'} \{dist(C(x), C(x'))\}$$

Будем обозначать параметры такого кода тройкой чисел $(n, k, dist)$.

Если расстояние кода равно $dist$, говорят, что код позволяет исправлять $e = \lceil \frac{dist-1}{2} \rceil$ ошибок. Это значит, что если в кодовом слове $y = C(x)$ выбрать не более e позиций и заменить соответствующие биты на противоположные, то по полученному слову y' можно однозначно восстановить исходное кодовое слово y и соответствующее $x = C^{-1}(x)$.

В практических приложениях желательно иметь коды с достаточно большой *скоростью* k/n и по возможности большим кодовым расстоянием. Нетрудно получить следующие оценки. С одной стороны, всякого $(n, k, dist)$-кода

$$2^k \leq \frac{2^n}{C_n^0 + \ldots + C_n^{\lceil \frac{dist-1}{2} \rceil}}$$

(граница Хэмминга). Если мы хотим, чтобы код позволял исправлять δn ошибок (и кодовое расстояние было больше $2\delta n$), данное неравенство даёт асимптотическую оценку $\frac{k}{n} \le (1 - h(\delta)) + o(n)$, где $h(p) = -p \log p - (1-p)\log(1-p)$. С другой стороны, нетрудно показать, что если

$$2^k < \frac{2^n}{C_n^0 + \ldots + C_n^{dist-1}}$$

то $(n, d, dist)$-код существует. Если $dist = 2\delta n + 1$, это неравенство даёт асимптотическую оценку $\frac{k}{n} \ge (1 - h(2\delta)) - o(n)$ для оптимального соотношения k и n (отношение k/n называют *скоростью кода*).

Код называется линейным, если отображение C оказывается линейным оператором $C : \mathbb{Z}_2^k \to \mathbb{Z}_2^n$. В этом случае множество кодовых слов образует линейно подпространство в $\mathbb{Z}_2^n$. Для линейного кода *кодовое расстояние* совпадает с минимальным числом единиц, которое может встретиться ненулевом в кодовом слове (поскольку побитовый XOR любых двух кодовых слов сам оказывается кодовым словом).

Отметим, что линейный код можно задать двумя способами: (а) с помощью матрицы оператора C (это будет матрица $n \times k$, столбцы которой задают базис в пространстве кодовых слов); и (б) с помощью *проверочной* матрицы $(n-k) \times n$, ядром которой является пространство кодовых слов.

Упражнение: Если

$$2^k < \frac{2^n}{C_n^0 + \ldots + C_n^{d-1}}$$

то существует линейный код с параметрами (n, k, d) (граница Варшамова–Гилберта).

Данное упражнение показывает, что коды (и даже линейные коды) с определёнными параметрами существуют. Трудность состоит в том, чтобы получить *эффективную* конструкцию такого кода (с полиномиальными алгоритмами кодирования и декодирования) для больших k и n. Далее мы приведём конструкцию Земора, позволяющую с помощью экспандеров строить коды с достаточно хорошим (хотя и заведомо неоптимальным) соотношением параметров, для которых есть быстрые алгоритмы кодирования и декодирования.

Рассмотрим сбалансированный двудольный граф $G = (L, R, E)$, у которого по m вершин в левой и правой доле, и степень каждой вершины равна d. Будем считать, что $m \times m$-матрица смежности данного графа совпадает с матрицей алгебраического экспандера с параметрами (m, d, α) для некоторого достатончо малого α.

В графе G имеется $n = dm$ рёбер. Далее мы опишем линейный код с параметрами $(n, k, dist)$ (параметры k и $dist$ мы уточним ниже). Прежде всего, каждому ребру графа припишем переменную x_i, $i = 1, \ldots, n$. Далее мы припишем каждой вершине графа некоторое семейство линейных уравнений с переменными x_i (над полем $\mathbb{Z}_2$). Вся эта совокупность уравнений и будет задавать пространство кодовых слов.

Для начала выберем *проверочную матрицу* H для линейного $(d, k', 2\delta d)$-кода с максимальным возможным k' ($k' \approx d(1 - h(2\delta))$, см. упражнение выше). Матрица H имеет размер $(d - k') \times d$ (она состоит из $\approx (dh(2\delta))$ линейно независимых d-битовых строк.) Для фиксированного d такую матрицу можно найти перебором. Далее, сопоставим каждой вершине нашего графа систему уравнений, состоящую в умножении H на вектор-столбец, составленный из x_i, приписанных рёбрам, выходящим из данной вершины.

Таким образом, каждая вершина даёт $k' \approx dh(2\delta)$ уравнений. Общее число уравнений, таким образом, составляет примерно $2mdh(2\delta)$. Это значит, что размерность пространства решений этой системы $\approx n - 2mdh(2\delta) = (1 - 2h(2\delta))n$ (таким образом, скорость кода равна $(1 - 2h(2\delta))$. Остаётся понять, сколько ошибок позволяет исправлять данный код.

Утверждение Расстояние построенного линейного кода не меньше $2\delta(2\delta - \alpha)n$.

Доказательство: Рассмотрим ненулевое кодовое слово $\bar{x} = (x_1 \dots x_n)$ с минимальным числом единиц (минимальным весом). Обозначим E' множество рёбер, соответствующих ненулевым битам этого кодового слова. Назовём S и T множества вершин из левой и правой долей графа, в которые входит хотя бы одно ребро из E'. Поскольку набор битов $\bar{x}$ является кодовым словом, он удовлетворяет уравнениям, приписанным всем вершинам графа. Но матрица H устроена так, что она может аннулировать только такой d-битовый вектор, в котором есть хотя бы $2\delta d$ единиц (либо уж *все* биты равны нули). Следовательно, в каждую из вершин S и T входит не менее $2\delta d$ рёбер E'. Таким образом, $|E'| \geq \frac{2\delta d}{2}(|S| + |T|)$ (в знаменателе стоит коэффициент 2, т.к. у ребра есть два конца — в левой и в правой доле).

Поскольку двудольный граф G построен из (m, d, α)-экспандера, мы можем применить лемму о перемешивании:

$$|E'| \leq |E(S,T)| \leq \frac{d|S||T|}{m} + \alpha d\sqrt{|S||T|}$$

Получаем

$$\delta(|S| + |T|) \leq \frac{|S||T|}{n} + \alpha\sqrt{|S|T|} \leq \frac{(|S| + |T|)^2}{4n} + \frac{\alpha}{2}(|S| + |T|)$$

(среднее геометрическое не превосходит среднего арифметического). Следовательно, $|S| + |T| \geq 2m(2\delta - \alpha)$. Вспомним, что $|E'| \geq \delta d(|S| + |T|)$. Получается, что число рёбер в E' не может быть меньше $2\delta(2\delta - \alpha)n$, и утверждение доказано.

Вывод: мы научились строить линейный код с параметрами

$$(n, (1 - 2h(2\delta))n, 2\delta(2\delta - \alpha)n)$$

Данная конструкция имеет смысл если $h(2\delta) < 1/2$; это условие выполнено для $\delta < 0.01$.

Нетрудно описать для данного кода полиномиальный алгоритм кодирования (*указание:* достаточно применить метод Гаусса и выписать общее решение системы уравнений, задающих пространство кодовых слов). Сложнее дело обстоит с процедурой декодирования искаженных кодовых слов. Оказывается, для любого $\varepsilon > 0$ можно построить быстрый алгоритм, который восстанавливает кодовое слово, в котором испорчено не более $(\frac{2\delta(2\delta-\alpha)}{2} - \varepsilon)n$ битов (на сколь угодно малую долю ε меньше, чем половина кодового расстояния). Процедура декодирования устроена следующим образом. Попеременно для всех вершин левой и правой доли графа мы производим локальную процедуру декодирования: для каждой вершины берём все входящие в неё рёбра; если значения d соответствующих переменных x_i не удовлетворяют проверочной матрице H, мы меняем их на ближайшее кодовое слов длины d (так, чтобы все контрольные суммы в данной вершине стали равны нулю). Поскольку на каждом шаге мы применяем процедуру коррекции к

вершинам в одной доле графа, каждое ребро может участвовать только в одной такой процедуре (так что коллизий не возникает). Именно здесь существенно, что выбранный нами граф двудольный.

Можно показать, что на если в самом начале пометки x_i на рёбрах были достаточно близки к некоторому кодовому слову, то на на каждом шаге число неправильных битов уменьшается в константу раз (так что через логарифмическое число итераций мы получаем кодовое слово, очищенное от ошибок). Доказательство этого факта основано на экспандерной лемме о перемешивании; подробности можно прочитать в [Zem].

Более сложные конструкции экспандерных кодов, можно найти в [Spi]; см. также обзор [Gu].

Применение экспандеров 3: Надёжные схемы из функциональных элементов

В этом разделе мы обсуждаем задачу построения надёжных схем из функциональных элементов. Мы предполагаем, что читатель знаком с понятием схемы из функциональных элементов, вычисляющей булеву функцию $f : \mathbb{B}^n \to \mathbb{B}$. Мы будем предполагать, что зафиксирован некоторый конечный *полный базис* булевых функций B, и каждой внутренней вершине схемы сопоставляется некоторая функция $g \in B$, причём арность g совпадает с входной степенью вершины (строго говоря, нужно ещё зафиксировать соответствие между входящими рёбрами и аргументами g). Входным вершинам схемы (вершинам с входной степенью 0) сопоставляются $x_1, \ldots, x_n$ (аргументы функции, которую должна вычислять схема.

Пусть задана схема из N функциональных элементов, вычисляющая некоторую функцию $f : \mathbb{B}^n \to \mathbb{B}$. Рассмотрим работу данной схемы 'с ошибками'. Будем предполагать, что каждый из функциональных элементов независимо от других элементов (и от входов схемы) с некоторой веротяностью ε 'ломается', становится 'сбойным'. Будем называть данное распределение сбоев *ε-случайным*. Мы считаем, что сбойные элементы схемы переходят во власть злонамеренного противника, который по своему произволу определяет их выходы. При этом выходы на остальных (несбойных) функциональных элементах определяются по обычным правилам.

Определение Схема из функциональных элементов (ε, δ)-надёжно вычисляет функцию f, если для любого набора входных значений, при ε-случайном выборе сбойных элементов, с вероятностью не менее $(1 - \delta)$ схема выдаёт правильное значение функции, как бы ни действовал противник.

Теорема[фон Нейманн] Для произвольного полного базиса булевых функций B, для всех достаточно малых ε найдётся $\delta = O(\varepsilon)$ такое, что всякая булева функция может быть вычислена (ε, δ)-надёжной схемой в данном базисе.

Доказательство: Прежде всего заметим, что если теорема верна для одного полного базиса, то она обязана выполняться и для любого другого базиса, быть может с другими ε и δ (поскольку элементы одного базиса можно моделировать блоками, составленными из элементов другого базиса). Без ограничения общности мы можем считать, что наш базис состоит из всех булевых функий трёх аргументов. Мы покажем, что любую обычную булеву схему можно переделать в (ε, δ)-надёжную. Доказательство проведём индукцией по глубине формулы.

Итак, пусть выход (обычной) булевой схемы вычислется применением функционального элемента b к тройке значений f_1, f_2, f_3. Каждое из значений f_1, f_2, f_3

в свою очередь вычисляются некоторыми подсхемами (быть может, пересекающимися). Глубины этих подсхем заведомо меньше, чем глубина всей схемы; поэтому мы можем считать, что для f_1, f_2, f_3 уже имеются (ε, δ)-надёжные схемы T_1, T_2, T_3. Если к выходам схем T_1, T_2, T_3 применить опрерацию b, то вероятность получить неверный ответ не превосходит $(3\delta + \varepsilon)$ (итоговый результат может оказаться неверным, если хотя бы одно из значений f_i вычислено неправильно или если сбой произошёл в самом элементе b). Назвоём построенную схему R. Чтобы уменьшить вероятность ошибки, мы изготовим три копии схемы R и применим к выходам этих трёх схем функцию большинства. Вероятность того, что и после этого мы получим ошиочный ответ, не превосходит

$$3(3\delta + \varepsilon)^2 + \varepsilon$$

(ошибка должна случиться хотя бы в двух из трех независимых копий схемы C либо в итоговом вычсилении большинства). Для малых ε и подходящего $\delta = O(\varepsilon)$ получаем

$$3(3\delta + \varepsilon)^2 + \varepsilon \leq \delta$$

и теорема доказана.

Отметим, что приведённая конструкция может экспоненциально увеличить размер схемы, хотя её глубина увеличивается лишь в константу раз.

Упражнение: Докажите, что для всех достаточно малых ε найдётся $\delta = O(\varepsilon)$ такое, что функцию большинства $majority(x_1, \ldots, x_n)$ можно вычислить (ε, δ)-надёжной схемой размера $\mathrm{poly}(n)$.

Далее мы докажем более сильный вариант теоремы фон Неймана:

Теорема Для произвольного полного базиса булевых функций B, для всех достаточно малых ε найдётся $\delta = O(\varepsilon)$ такое, что всякая булева схема C из N элементов может быть переделана (за время $\mathrm{poly}(N)$) в (ε, δ)-надёжную схему размера $O(N \log N)$.

Доказательство: Прежде чем доказывать теорему, введём определение:

Определение Двудольный граф называется (k, d, α, β)-компрессором, если

1. в левой и правой долях графа содержится по k вершин;
2. степень каждой вершины равна d;
3. пусть A — произвольное множество вершин левой доли графа, и $|A| \leq \alpha k$; обозначим B множество таких вершин правой доли графа, у которых большинство соседей принадлежат A; тогда размер B не превосходит βk.

Лемма о компрессоре Если $4\alpha(\gamma^2 + \alpha) < \beta < 1/2$, то матрица смежности алгебраического (k, d, γ)-экспандера задаёт (k, d, α, β)-компрессор (двудольный граф с $2 \times k$ вершинами также задаётся матрицей $k \times k$).

Отложим доказательство леммы и покажем, как она помогает доказать теорему. Зафиксируем некоторый параметр k (в последствии мы выберем $k = O(\log N)$). Долее, построим (k, d, α, β)-компрессор такой, что $\beta + \Gamma\varepsilon < \alpha$ (константа Γ не зависит от k и определяется соотношением числа d и базиса, над которым мы строим схему; подробнее значение Γ мы обсудим ниже).

Теперь мы преобразовываем заданную нам схему C в эквивалентную ей (ε, δ)-надёжную схему C'. Для этого мы заменим каждый функциональный элемент на некоторый блок из $O(k)$ элементов (устройство такого блока мы сейчас опишем). Если в схеме C выход элемента номер i подавался на вход элементу номер j,

то в новой схеме C' от блока номер i к блоку номер j будет идти 'кабель' из k проводов. В идеальной ситуации (когда нет ошибок) сигналы во всех проводах этого кабеля будут одинаковы; более того, это будет ровно тот сигнал, который проходил по соответствующему проводу в исходной схеме (при тех же значениях входов схемы).

Теперь опишем устройство блока, соответствующего одному из элементов схемы C. Мы объясним конструкцию на простейшем примере: пусть в C присутствовал функциональный элемент конъюнкция; наша задача — построить надёжный блок, успешно моделирующий этот функциональный элемент при 'умеренном' количестве ошибок. В этот блок будут входить $2k$ сигналов (два кабеля по k проводов). Мы сводим соответствующие провода из этих кабелей (первый с первым, второй со вторым, и т.д.) и для каждой пары вычисляем конъюнкцию. Получаем k результирующих сигналов. Затем пропускаем эти сигналы через *корректор*: это схема с k входами и k выходами; каждый выход вычисляется как *большинство* среди некоторых d входов; а правило, по которому каждому из выходов сопоставляются d входов, есть (k, d, α, β)-компрессор. Отметим, что блок реализуется схемой глубины $O(1)$ и состоит из $O(k)$ функциональных элементов (константы зависят от выбора базиса).

С помощью оценки вероятнсти больших уклонений (неравенство Чернова) легко показать, что если $k = \Omega(\log N)$, то с большой вероятностью ни в одном из N описанных блоков не случится больше $\Gamma\varepsilon k$ (число Γ определяется глубиной схемы-корректора, т.е. зависит от выбора базиса). В таком случае, если каждый из входных кабелей несет не более αk 'неправильных' сигналов (т.е. сигналов, отличных от значения в соответствующем проводе исходной схемы C), то и среди k выходных сигналов не более αk ошибочных. Действительно, перед применением *корректора* неправильные сигналы обоих входов складываются – их может стать 2α. Затем мы пропускаем сигналы через компрессор, и доля ошибок уменьшается до β. Наконец, нужно учесть ещё $O(\varepsilon k)$ новых ошибок, которые могли случиться в самом блоке. Всего на выходе имеем долю ошибок $\beta + O(\varepsilon) < \alpha$.

Чтобы закончить конструкцию, нам нужно вычленить из k-жильного кабеля на выходе последнего блока *один* сигнал с ответом. Для этого нам нужно вычислить *большинство* среди значений этих k сигналов. Это можно делать разными способами; например, можно применить 'экспоненциальную' конструкцию фон Нейманна (при вычислении фукнции большинства среди $O(\log N)$ значений данный метод даст схему размера $\mathrm{poly}(\log N)$, см. Упраженение выше).

Доказательство леммы о компрессоре: Пусть $e_1, \ldots, e_k$ — ортонормированный собственный базис матрицы M заданного (k, d, γ)-экспандера, а $\lambda_1, \ldots, \lambda_k$ – соответствующие собственные числа. Мы будем считать, что собственные числа упорядочены по убывания абсолютной величины. При этом

$$e_1 = \frac{1}{\sqrt{k}}(1, 1, \ldots, 1),$$

а $\lambda_1 = d$ (по условию леммы остальные собственные числа по модулю не превосходят γk). Пусть A – некоторое множество вершин графа, и $|A| \leq \alpha k$. Обозначим $f = (f_1, \ldots, f_k)$ характеристический вектор этого множества ($f_i = 1$ если и только если i-ая вершина графа принадлежит A). Ясно, что $||f||^2 \leq \alpha k$. Оценим норму вектора Mf.

$$||Mf||^2 = (Mf, Mf) = (f, M^2 f) = \sum_{i=1}^{k} \lambda_i^2 (f, e_i)^2 = \alpha^2 d^2 k + \sum_{i=2}^{k} \lambda_i^2 (f, e_i)^2$$

Поскольку все собственные числа кроме первого по модулю не превосходят γk, получаем

$$||Mf||^2 \leq \alpha^2 d^2 k + (\gamma d)^2 \sum_{i=2}^{k} (f, e_i)^2 \leq \alpha^2 d^2 k + (\gamma d)^2 ||f||^2 \leq (\alpha^2 d^2 + \alpha \gamma^2 d^2) k$$

Далее, для выбранного A мы рассмотрим множество B, которое состоит из всех вершин графа, у которых не менее $d/2$ соседей лежат в A. Это значит, что B состоит из таких вершин $i = 1, \ldots, n$, что в i-ой координате вектора $f' = (Mf)$ стоит число не менее $d/2$. Получаем

$$|B| \leq \sum_{i=1}^{k} \left(\frac{f'}{d/2} \right)^2 \leq \frac{4}{d^2} ||Mf||^2 \leq 4(\alpha^2 + \alpha \gamma^2) k \leq \beta k$$

Литература: [Gacs]

Список литературы

[HLW] S. Hoory, N. Linial, A. Wigderson. Expander graphs and their applications. Bulletin of the AMS, vol. 43, Number 4, Oct. 2006, pp.439–561.

[AB] S. Arora, B. Barak. Computational Complexity: A modern Approach. (Draft version is available online: `http://www.cs.princeton.edu/theory/complexity/`)

[AS] N. Alon, J. H. Spencer. The Probabilistic Method. 2nd ed. Wiley-Interscience Publication. *Русский перевод:* Н. Алон, Дж. Спенсер. Вероятностный метод. Бином. Лаборатория знаний, 2007

[BMRV] H. Buhrman, P.B. Miltersen, J. Radhakrishnan, S. Venkatesh. Are Bitvectors optimal? SIAM J. Comput., 31(6):1723–1744, 2002.

[Zem] G. Zémor. On Expander codes. IEEE Trans. on Inf. Theory. 47(2), 835–837, 2001.

[Sar] P. Sarnak. Some applications of modular forms. Cambridge University Press, 1990. *Русский перевод:* П. Сарнак. Модулярные формы и их приложения. Москва, Фазис, 1998.

[Rei] O. Reingold. Undirected st-connectivity in log-space. In Proceedings of the 37th Annual ACM Symposium on Theory of Computing, pages 376–385, 2005.

[Gu] V. Guruswami. Error-correcting Codes and Expander Graphs. SIGACT News Complexity Theory Column 45, 2004.

[Spi] D. Spielman. Constructing error-correcting codes from expander graphs. In Emerging Applications of Number Theory, IMA volumes in mathematics and its applications, volume 109, 1996.

[Gacs] P. Gács. Book chapter on reliable computation. http://www.cs.bu.edu/ gacs/papers/iv-eng.pdf