Лекция 11

Уменьшение вероятности ошибки алгоритма из BPP с использованием небольшого количества случайных битов

(Конспект: А. Куликов)

Целью данной лекции будет построение алгоритма, уменьшающего вероятность ошибки алгоритма (обозначим его буквой Λ) из класса **BPP** (а точнее, алгоритма, распознающего язык из **BPP**) до 2^{-k} и использующего всего $O(r+k)$ случайных битов, где r – количество случайных битов, используемых алгоритмом Λ.

Говоря неформально, наша конструкция выглядит следующим образом. Сначала мы зададим граф, каждой вершине которого будет соответствовать r-битовая строчка (эти строчки мы и будем использовать в алгоритме Λ в качестве случайной строки). (Итого в нашем графе имеется $n = 2^r$ вершин.) Используя настоящие случайные биты, выберем одну из вершин этого графа и дадим соответствующую ей строчку алгоритму Λ. Далее, опять же используя случайные биты (но лишь $O(1)$ штук), перейдем из начальной вершины по ребрам графа в другую вершину, вследствие чего получим новую псевдослучайную строчку, которую и подадим Λ. Из этой вершины опять перейдем в другую и т. д. — всего обойдем $O(k)$ вершин. Ответ, как и обычно, дадим «большинством голосов».

Теперь перейдем к формальным определениям. Приведем необходимый нам для последующего определения факт из курса алгебры:

Факт 11.1. *Пусть $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n$ – собственные числа матрицы смежности связного двудольного*[1] *d-регулярного графа. Тогда*

(1) $\lambda_1 = d = -\lambda_n$,

(2) $\lambda_1 > \lambda_2$,

(3) собственный вектор ортонормального собственного базиса, соответствующий собственному числу λ_1, имеет вид $(1, ..., 1) \cdot 2^{-n/2}$.

Определение 11.1. (Бесконечное) семейство (различных между собой) двудольных d-регулярных графов называется семейством ε-расширителей, если для любого графа из этого семейства $\lambda_1 - \lambda_2 \geq \varepsilon$.

Это не слишком традиционное определение, но оно в определенном смысле эквивалентно другим определениям расширителей.

Приведем теперь конструкцию семейства графов-расширителей. Степень каждой вершины каждого из этих графов будет равняться 7, сам же граф будет содержать ровно по m^2 вершин в каждой доле. Проиндексируем все вершины каждой из долей парами чисел (x, y), где $x, y \in [0..m-1]$. Вершину (x, y) левой доли соединим с вершинами $(x, y), (x, 2x+y), (x, 2x+y+1), (x, 2x+y+2), (x+2y, y), (x+2y+1, y), (x+2y+2, y)$ правой доли (естественно, все числа в индексах вершин здесь берутся по модулю m). Доказывать, что построенное семейство графов является семейством расширителей, мы не будем. (Отметим, что само семейство строится детерминированно, причем строить сами графы нам, на самом деле, не нужно: ведь мы можем «вычислить» всех соседей любой вершины по ее индексу (x, y).)

Мы устроим случайное блуждание на этом графе. Зададим матрицу перехода марковской цепи таким образом: $B_0 = E/2 + A/14$, где E — единичная матрица, а A — матрица смежности графа размера 2^r из построенного семейства графов (возможно, здесь число r будет отличаться на 1 от его начального значения, ведь все графы в построенном семействе имеют размер $2m^2$, но на оценку это не повлияет). Такая матрица перехода задает следующие действия: с вероятностью $1/2$ мы остаемся в текущей вершине и с вероятностью $1/2$ идем (равновероятно) в одного из соседей этой вершины. Ясно, что на один переход нам потребуется не более четырех случайных битов. Всего же переходов от одной псевдослучайной строчки к другой будем делать β штук, где β — некоторая фиксированная константа, удовлетворяющая нервенству $\lambda_2^{\beta} < 1/10$, где

[1] Здесь и далее под двудольным графом мы будем понимать лишь графы, содержащие равные количества вершин в обеих долях.

λ_2 – второе собственное число матрицы B_0. После этого будем запускать на полученной случайной строчке алгоритм Λ и совершать следующие β переходов. Всего мы используем $k' = O(k)$ случайных строчек, проделав для этого $\beta k' = O(k)$ переходов. *Отныне и до конца лекции нас будет интересовать матрица $B = B_0^\beta$ перехода за β шагов (заметим, что у этой матрицы такие собственные числа: $\lambda_1 = 1$, $\lambda_2 < \frac{1}{10}$, остальные — меньше, но неотрицательны; первый собственный вектор — $2^{-r/2}(1, \ldots, 1)$).*

Итак, алгоритм описан, в нем действительно используется лишь $O(r + k)$ случайных битов. Докажем его корректность.

Для удобства будем считать, что алгоритм Λ ошибается с вероятностью, меньшей 1/100. Введем следующие обозначения: $\bar{W}$ — матрица размера $2^r \times 2^r$, в которой 1 записаны лишь на диагонали в местах, соответствующих r-битовым строкам, на которых алгоритм Λ ошибается; $W = E - \bar{W}$ (в ней 1 стоят в «верных» местах); $p_0 = 2^{-r}(1, \ldots, 1)$ — начальная строка состояния нашей Марковской цепи (данная строка показывает, что начальную вершину мы выбираем равновероятно). Тогда pB — распределение вероятности после одного перехода из распределения p, $||p\bar{W}||_1$ – вероятность быть в «хорошей» (такой, на которой Λ дает верный ответ) строке, $||pW||_1$ – в «плохой». Для доказательства корректности построенного нами алгоритма достаточно показать, что вероятность того, что большинство строчек, выданных алгоритму Λ, будут «плохими», мала.

Пусть $S_i = W$, если i-я поданная на вход Λ строчка «хорошая», и $S_i = \bar{W}$ в противном случае. Тогда верна следующая лемма.

Лемма 11.1. *$P\{S_1, \ldots, S_{k'}\} = ||p_0 BS_1 \ldots BS_{k'}||_1$ (здесь под S_i-ым в левой части равенства мы понимаем соответствующее этой матрице событие «хорошести» i-ой входной строчки).*

Напомним, что нам нужно показать, что если среди S_i-ых много «плохих» событий, то вероятность появления такой последовательности мала. Покажем мы это, воспользовавшись следующей леммой.

Лемма 11.2. $\forall p$

1. $||pBW||_2 \le ||p||_2$
2. $||pB\bar{W}||_2 \le \frac{1}{5} \cdot ||p||_2$

Доказательство. 1. $||pBW||_2 \le ||pB||_2 = ||\sum c_i \lambda_i e_i||_2 = \sqrt{\sum c_i^2 \lambda_i^2} \le \sqrt{\sum c_i^2} = ||p||_2$. Здесь $\{e_i\}$ — ортонормированный базис из собственных векторов матрицы B; последнее неравенство верно в силу того, что собственные числа B заключены в отрезке $[-1; 1]$.

2. Разложим вектор p в сумму: $p = x + y$, где $x = c_1e_1, y = \sum_{i\geq 2} c_ie_i$. Тогда $||pB\bar{W}||_2 \leq ||xB\bar{W}||_2 + ||yB\bar{W}||_2$. Оценим отдельно каждое слагаемое.

   - $||xB\bar{W}||_2 = ||c_1\lambda_1e_1\bar{W}||_2 = ||c_1e_1\bar{W}||_2 \leq c_1/10 \leq \frac{1}{10}||p||_2$. Предпоследнее неравенство здесь выполнено по причине того, что в $\bar{W}$ достаточно «мало» 1 (т.к. алгоритм Λ ошибается с вероятностью, меньшей 1/100), а все компоненты e_1 равны между собой.
   - $||yB\bar{W}||_2 = ||\sum_{i\geq 2} c_i\lambda_ie_i\bar{W}||_2 \leq ||\sum_{i\geq 2} c_i\lambda_ie_i||_2 = \sqrt{\sum_{i\geq 2} c_i^2\lambda_i^2} \leq \sqrt{\sum_{i\geq 2} c_i^2\lambda_2^2} = \lambda_2||p||_2 \leq \frac{1}{10}||p||_2$.

□

Итак, вероятность появления конкретной «плохой» последовательности оценивается следующим образом:

$$\begin{aligned} P\{S_1, \ldots, S_{k'}\} &= ||p_0BS_1 \ldots BS_{k'}||_1 \\ &\leq 2^{r/2}||p_0BS_1 \ldots BS_{k'}||_2 \\ &\leq 2^{r/2} \cdot 5^{-k'/2} \cdot 2^{-r} \cdot 2^{r/2} \\ &= 5^{-k'/2}. \end{aligned}$$

Тогда вероятность того, что построенный нами алгоритм получит плохую последовательность, не превосходит $(2 \cdot 5^{-1/2})^{k'}$ (здесь мы воспользовались тем, что количество плохих последовательностей не превосходит количества последовательностей вообще, равного $2^{k'}$). Выбрав $k' = 7k$, получаем искомую вероятность ошибки $\leq 2^{-k}$.