Конспект[1] к лекции 8. (Санкт-Петербург, 16 апреля 2017 г.)

20 Экстракторы случайности

В этом параграфе мы обсудим конструкции, позволяющие улучшить «качество» битов, получаемых на выходе случайного генератора. Представим себе, что у нас имеется физическое устройство, по требованию выдающее последовательность из n случайных битов. В идеале нам бы хотелось, чтобы все n битов были независимы и равномерно распределены (т.е., чтобы каждый из 2^n потенциально возможных наборов значений появлялся с вероятностью $1/2^n$). Однако на практике распределение на выходе генератора может немного отличаться от идеального. Значения битов могут быть немного зависимы, вероятности нуля и единицы в каждой из позиций могут немного смещены относительно $1/2$, и т.д. Для некоторых приложений даже небольшое ухудшение качества случайных битов может оказаться критическим. Скажем, если для каждого из n битов, выдаваемых генератором, вероятности появления нуля и единицы равны 0.499 и 0.501 соответственно (вместо ожидаемого симметричного распределения с вероятностями $1/2$), то в последовательности «случайных» битов длиной в неколько миллионов частоты нулей и единиц будут довольно заметно отличаться от ожидаемых 50%. В результате некоторые вероятностные алгоритмы, хорошо работающие на «идеальных» случайных битах, могут давать неожиданно большую вероятность ошибки при использовании «смещённых» случайных битов.

Тем не менее, интуитивно кажется ясным, что даже немного «смещённые» случайные биты содержат в себе большую неопределённость. Нельзя ли каким-то образом выделить полезную случайность и превратить «смещённые» случайные биты в «почти идеальные»? Мы покажем, что в некотором смысле это оказывается возможным. Чтобы уточнить данное утверждение, нам придётся дать некоторые формальные определения. Мы должны пояснить, какие именно распределения вероятностей считаются «не очень сильно смещёнными», и какие распределения можно называть «почти идеальными».

Определение 20.1 *Говорят, что* min-*энтропия случайной величины X равна k, если*

$$\max_a \mathrm{Prob}[X = a] = 2^{-k}$$

(*максимум берётся среди всех значений a случайной велиичны X*).

Именно распределения с большой min-энтропией мы будем считать «несильно смещённым» и «достаточно качественными». Рассмотрим несколько несложных примеров:

[1] В этой лекции мы изучили одну из конструкций экстрактора (см. параграф 20), а также начали обсуждение генераторов Нисана. О генераторах Нисана см. конспект следующей лекции.

- Если случайная величина X равномерно распределена на $\{0,1\}^n$, то min-энтропия X равна n (каждое из 2^n возможных значений имеет вероятность $1/2^n$). Отметим, что у *любого* распределения на $\{0,1\}^n$ min-энтропия не превосходит n, так что равномерное распределение достигает экстремального значения min-энтропии.

- Если случайная величина X равномерно распределена на множестве всех n-битных строк, содержащих чётное число единиц, то min-энтропия X равна $n-1$ (имеется 2^{n-1} возможных значений, и вероятность каждого из них равна $1/2^{-(n-1)}$).

- Если $X = (X^{(1)} \dots X^{(n)})$ есть последовательность независимых одинаково распределённых двоичных случайных величин, и для каждого $i = 1, \dots, n$
$$\begin{cases} \text{Prob}[X^{(i)} = 0] = \frac{1}{2} - \delta, \\ \text{Prob}[X^{(i)} = 1] = \frac{1}{2} + \delta, \end{cases}$$
то min-энтропия X равна
$$\log\left(\frac{1}{(\frac{1}{2}+\delta)^k}\right) = -k\log\left(\frac{1}{2}+\delta\right) = k(1-\Theta(\delta))$$
(среди всех возможных значений максимальную вероятность имеет последовательность $11\dots1$ — последовательность из одних единиц).

Далее мы уточним, какие распредления естественно считать «почти идеальными» (для всевозможных практических применений).

Определение 20.2 *Распределение вероятностей $X = (X_1, \dots, X_n)$ на $\{0,1\}^n$ считается ε-близим к равномерному, если для любого $A \subset \{0,1\}^n$*
$$\left|\text{Prob}[X \in A] - \frac{|A|}{2^n}\right| \leqslant \varepsilon.$$

Упражнение 20.1 *Предположим, что некоторый вероятностный алгоритм $\mathcal{A}$ при использовании набора из n «идеальных» (независимых и равномерно распределённых) случайных битов на каждом входе ошибается с вероятностью не более δ. Докажите, что при использовании неидеального, но ε-близого к равномерному источника случайности, данный алгоритм будет ошибаться с вероятностью не более $\varepsilon + \delta$.*

Упражнение 20.2 *Покажите, что ε-близость к равномерному распределению означает, что распределение X удалено от равномерного распределения на расстояние не более 2ε в смысле l_1-нормы. Другими словами, распредление $X = (X_1, \dots, X_n)$ является ε-близим к равномерному, если и только если*
$$\sum_{(i_1,\dots,i_n)\in\{0,1\}^n} \left|\text{Prob}[X = (i_1, \dots, i_n)] - \frac{1}{2^n}\right| \leqslant 2\varepsilon.$$

Далее мы дадим определения *экстрактора случайности* — схемы, позволяющей превратить набор неидеальных случайных битов с большой min-энтропией в «почти идеальные» биты. При этом экстрактору потребуется в качестве «затравки» небольшое число по-настоящему идеальных случайных битов.

Определение 20.3 *Экстрактором с параметрами* $(n, m, k, t, \varepsilon)$ *называется отображение*

$$Ext \ : \ \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^m$$

такое, что для всякого распределения $X = (X_1, \ldots, X_n)$ *с* min*-энтропией* $\geqslant k$ *и для равномерно распределенной (независимой с* X*) случайной величины* $Y = (Y_1, \ldots, Y_t)$ *получаемое распределение значений* $Ext(X, Y)$ *оказывается* ε*-близким к равномерному.*

Теорема 20.1 *Для бесконечно многих* n*, для всех* $k \leqslant n$ *и всех* $\varepsilon < 1$ *существует полиномиально вычислимый экстрактор* Ext *с параметрами* $(n, m = n, k, t = O(n - k + \log \frac{1}{\varepsilon}), \varepsilon)$.

Замечание: Предлагаемое ниже доказательство использует спектральный экспандер на 2^n вершинах. Используя конструкцию из лекции 4, мы можем построить такие экспандеры не просто для *бесконечно многих* n, а для некоторой последовательности чисел n, образующих геометрическую прогрессию (см. замечание в конце главы 4).

Доказательство теоремы. Прежде всего опишем конструкцию экстрактора $Ext(X, Y)$. Зафиксируем некоторый спектральный экспандер с параметрами

$$(2^n, d = O(1), \gamma < 1).$$

Отображение Ext будет устроено следующим образом. С помощью случайной величины X (распределённой на наборах из n нулей и единиц, с min-энтропией не менее k) выбираем случайную вершину экспандера. Далее с помощью случайной величины Y организуем случайное блуждание на экспандере, начинающееся в выбранной вершине: t битов «идеально распределённой» величины Y позволяют нам выбрать случайный путь длины

$$T = t/(\log d)$$

по рёбрам графа. Последнюю вершину этого пути (точнее, её n-битный индекс) мы и возвращаем в качестве значения $Ext(X, Y)$.

Покажем, что получаемое в результате распределение достаточно близко к равномерному. Для начала измерим «близость» с помощью l_2-нормы. Обозначим $\mathbf{p}^{(i)}$ распределение вероятностей на вершинах графа после i-го шага блуждания. В частности, $\mathbf{p}^{(0)}$ совпадает с исходным распределением X, а $\mathbf{p}^{(T)}$ есть итоговое распределение $Ext(X, Y)$. Далее, обозначим $\hat{M}$ нормализованную матрицу экспандера.

По определению спектрального экспандера, операторная норма $\hat{M}$ на подпространстве векторов, ортогональных собственному вектору $(1, 1, \ldots, 1)$, равна γ. Это значит, что при каждом умножении на $\hat{M}$ норма разности $\|p^{(i)} - (\frac{1}{2^n}, \frac{1}{2^n}, \ldots, \frac{1}{2^n})\|$ уменьшается не менее, чем в γ раз. Таким образом,

$$\|p^{(T)} - (\frac{1}{2^n}, \ldots, \frac{1}{2^n})\| \leqslant \gamma^T \cdot \|p^{(0)} - (\frac{1}{2^n}, \ldots, \frac{1}{2^n})\| \leqslant \gamma^T \cdot \left(\|p^{(0)}\| + \|(\frac{1}{2^n}, \ldots, \frac{1}{2^n})\| \right).$$

При этом $\|(\frac{1}{2^n}, \ldots, \frac{1}{2^n})\| = 2^{-n/2}$, а

$$\|p^{(0)}\| = \sqrt{\sum_{i=1}^{n} \left(p_i^{(0)}\right)^2} \leqslant \sqrt{\frac{1}{2^k} \cdot \sum_{i=1}^{n} p_i^{(0)}} = 2^{-k/2}$$

(здесь мы воспользовались тем, что исходное распределение X на вершинах графа имеет min-энтропию k). Следовательно,

$$\|p^{(T)} - (\frac{1}{2^n}, \ldots, \frac{1}{2^n})\| \leqslant \gamma^T \cdot (2^{-k/2} + 2^{-n/2}).$$

Остаётся сравнить l_0 и l_1 нормы:

$$\left\| p^{(T)} - (\frac{1}{2^n}, \ldots, \frac{1}{2^n}) \right\|_0 \leqslant \sqrt{2^n} \cdot \left\| p^{(T)} - (\frac{1}{2^n}, \ldots, \frac{1}{2^n}) \right\| \leqslant \gamma^T \cdot 2^{n/2} (2^{-k/2} + 2^{-n/2}).$$

Теперь подберём такое T, чтобы правая часть неравенства оказалась меньше ε. Легко проверить, что достаточно взять $T = O(n - k + \log \frac{1}{\varepsilon})$. Таким образом, теорема доказана.