Курс лекций по теоретической криптографии

Тема 6. **Псевдослучайные генераторы**

Шокуров А.В.

Мотивация

Если противник неограниченный, для абсолютной стойкости шифра (совершенной секретности по Шеннону) нужен длинный случайный ключ.

Если противник моделируется эффективным алгоритмом, достаточно уметь на основе коротких случайных строк генерировать такие длинные ключи, которые никакой эффективный алгоритм *не сможет отличить* от чисто случайных.

Вспомогательные понятия

Будем рассматривать случайные величины ξ_n, ζ_n $(n \in \mathbb{N})$ со значениями в $\mathbb{B}^*$.

υ_n обозначает случайную величину, равномерно распределённую на $\mathbb{B}^n$
$(\forall x \in \mathbb{B}^n \quad \Pr\left[\upsilon_n = x\right] = \frac{1}{2^n})$.

Определение

Семейства случайных величин $\{\xi_n\}_{n\in\mathbb{N}}$ и $\{\zeta_n\}_{n\in\mathbb{N}}$ называются *вычислительно неразличимыми* (computational indistinguishable in polynomial time), если для любой п. в. м. Т. D

$$|\Pr\left[D(1^n, \xi_n) = 1\right] - \Pr\left[D(1^n, \zeta_n) = 1\right]| = \mathrm{negl}(n).$$

Определение

Семейство случайных величин $\{\xi_n\}_{n\in\mathbb{N}}$ называется *псевдослучайным* (pseudorandom), если оно вычислительно неотличимо от равномерно распределённого семейства случайных величин $\{\upsilon_{m(n)}\}_{n\in\mathbb{N}}$.

NB Для таких ξ_n и $m(n)$ с большой вероятностью $|\xi_n| = m(n)$.

Основное определение (Блюм, Микали, 1980)

Определение

Функция $g : \mathbb{B}^* \to \mathbb{B}^*$, такая, что $g(\mathbb{B}^n) \subseteq \mathbb{B}^{m(n)}$ для некоторого полинома m, называется *псевдослучайным генератором* (pseudorandom (number) generator)
или, полностью, *криптографически стойким генератором псевдослучайных последовательностей*, если

1. g — полиномиально вычислима;
2. $m(n) > n$ для всех $n \in \mathbb{N}$;
3. $\{g(\upsilon_n)\}_n$ — псевдослучайное семейство случайных величин.

NB Третье условие можно сформулировать так:
для любой п. в. м. Т. D и равномерно распределённых случайных величин υ_n

$$|\Pr\left[D(1^n, g(\upsilon_n)) = 1\right] - \Pr\left[D(1^n, \upsilon_{m(n)}) = 1\right]| = \mathrm{negl}(n).$$

NB Можно вместо полиномиально вычислимой функции g рассматривать её генератор — полиномиальную (детерминированную) машину Тьюринга G.

Эквивалентное определение (Яо, 1982)

Пусть при всех $n \in \mathbb{N}$ случайные величины ξ_n принимают значения в $\mathbb{B}^{m(n)}$.

Определение

Семейство случайных величин $\{\xi_n\}_n$ удовлетворяет условию *непредсказуемости следующего бита* (next bit unpredictability), если
для любой п. в. м. Т. P

$$\Pr_{i \in_{\mathcal{U}} \{1, \ldots, m(n)\}} \left[P(1^n, \xi_n^{[1 \ldots i-1]}) = \xi_n^{[i]} \right] \leqslant \frac{1}{2} + \mathrm{negl}(n).$$

Теорема (Yao)

Семейство случайных величин $\{\xi_n\}_n$ псевдослучайно тогда и только тогда, когда $\{\xi_n\}_n$ удовлетворяет условию непредсказуемости следующего бита.

Следствие (эквивалентное определение псевдослучайного генератора)

Функция $g : \mathbb{B}^* \to \mathbb{B}^*$, $g(\mathbb{B}^n) \subseteq \mathbb{B}^{m(n)}$, — псевдослучайный генератор тогда и только тогда, когда

1. g — полиномиально вычислима;
2. $m(n) > n$ для всех $n \in \mathbb{N}$;
3. $\{g(\upsilon_n)\}_n$ удовлетворяет условию непредсказуемости следующего бита.

Доказательство теоремы Яо ($\Rightarrow$)

Теорема (Yao)

$\{\xi_n\}_n$ — псевдослучайное семейство $\iff$ $\{\xi_n\}_n$ удовлетворяет условию НСБ.

Доказательство. ($\Rightarrow$) Пусть существуют п. в. м. Т. P («предсказатель») и полином p:

$$\Pr_i\left[P(1^n, \xi_n^{[1\,\ldots i-1]}) = \xi_n^{[i]}\right] > \frac{1}{2} + \frac{1}{p(n)}$$ для бесконечно многих n.

Построим «различитель» — п. в. м. Т. D, работающую на входах $(1^n, x)$, $x \in \mathbb{B}^{m(n)}$, так:

1. $i \in_{\mathcal{U}} \{1, \ldots, m(n)\}$
2. $b = P(1^n, x^{[1\,\ldots i-1]})$
3. если $b = x^{[i]}$, то $D(1^n, x) = 1$, иначе $D(1^n, x) = 0$

- $\Pr\left[D(1^n, \xi_n) = 1\right] = \Pr_i\left[P(1^n, \xi_n^{[1\,\ldots i-1]}) = \xi_n^{[i]}\right] > \frac{1}{2} + \frac{1}{p(n)}$
- $\Pr\left[D(1^n, \upsilon_{m(n)}) = 1\right] = \Pr_i\left[P(1^n, \upsilon_{m(n)}^{[1\,\ldots i-1]}) = \upsilon_{m(n)}^{[i]}\right] =$

$$= \sum_{k=1}^{m(n)} \Pr_i\left[P(1^n, \upsilon_{m(n)}^{[1\,\ldots k-1]}) = \upsilon_{m(n)}^{[k]},\ i = k\right] =$$

$$= \sum_{k=1}^{m(n)} \Pr\left[P(1^n, \upsilon_{m(n)}^{[1\,\ldots k-1]}) = \upsilon_{m(n)}^{[k]}\right] \cdot \Pr_i\left[i = k\right] = m(n) \cdot \frac{1}{2} \cdot \frac{1}{m(n)} = \frac{1}{2}$$

Разность этих вероятностей $> \frac{1}{p(n)}$ для бесконечно многих n — противоречие.

Доказательство теоремы Яо ($\Leftarrow$)

($\Leftarrow$) *От противного.* Предположим, $\{\xi_n\}$ и $\{\upsilon_{m(n)}\}$ не вычислительно неразличимы: существуют такие п. в. м. Т. D («различитель») и полином p, что для бесконечно многих n

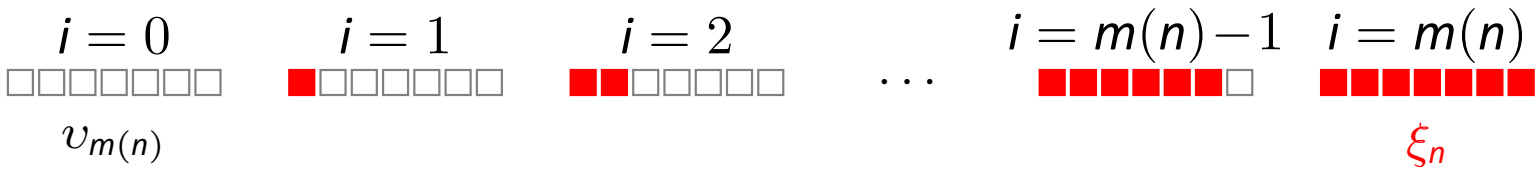
$$\Pr\bigl[D(1^n, \xi_n) = 1\bigr] - \Pr\bigl[D(1^n, \upsilon_{m(n)}) = 1\bigr] > \frac{1}{p(n)}$$

(модуль можно снять, так как либо верно это, либо то же самое с минусом — тогда возьмём вместо D машину D': $D'(1^n, x) = 1 \leftrightarrow D(1^n, x) \neq 1$).

Построим «предсказатель следующего бита» — п. в. м. Т. P, работающую на входах $(1^n, x)$, где $x \in \mathbb{B}^{\leqslant m(n)-1}$, следующим образом:

1. $y \in_{\mathcal{U}} \mathbb{B}^{m(n)-|x|}$;
2. $z = D(1^n, xy)$;
3. выход $P(1^n, x) = \begin{cases} y^{[1]}, & \text{если } z = 1, \\ \overline{y^{[1]}}, & \text{если } z \neq 1. \end{cases}$

Обозначим $\delta_i(n) = \Pr\bigl[D(1^n, \xi_n^{[1\,\ldots i]} \upsilon_{m(n)}^{[i+1\,\ldots m(n)]}) = 1\bigr], \quad 0 \leqslant i \leqslant m(n).$



Доказательство теоремы Яо ($\Leftarrow$)

Зафиксируем *i* и рассмотрим вероятность $\Pr\big[P(1^n, \xi_n^{[1\,\ldots i-1]}) = \xi_n^{[i]}\big] =$

[далее ради краткости будем опускать 1^n и писать m вместо $m(n)$]

$$= \Pr\big[P(\xi_n^{[1\,\ldots i-1]}) = \xi_n^{[i]},\ \upsilon_m^{[i]} = \xi_n^{[i]}\big] \qquad + \Pr\big[P(\xi_n^{[1\,\ldots i-1]}) = \xi_n^{[i]},\ \upsilon_m^{[i]} = \overline{\xi_n^{[i]}}\big] =$$

$$= \Pr\big[P(\xi_n^{[1\,\ldots i-1]}) = \upsilon_m^{[i]},\ \upsilon_m^{[i]} = \xi_n^{[i]}\big] \qquad + \Pr\big[P(\xi_n^{[1\,\ldots i-1]}) = \overline{\upsilon_m^{[i]}},\ \upsilon_m^{[i]} = \overline{\xi_n^{[i]}}\big] =$$

$$= \Pr\big[D(\xi_n^{[1\,\ldots i-1]}\upsilon_m^{[i\,\ldots m]}) = 1,\ \upsilon_m^{[i]} = \xi_n^{[i]}\big] + \Pr\big[D(\xi_n^{[1\,\ldots i-1]}\upsilon_m^{[i\,\ldots m]}) \neq 1,\ \upsilon_m^{[i]} = \overline{\xi_n^{[i]}}\big] =$$

$$= \sum_{b\in\mathbb{B}} \Pr\big[D(\xi_n^{[1\,\ldots i-1]} b \upsilon_m^{[i+1\,\ldots m]}) = 1,\ \upsilon_m^{[i]} = b,\ \xi_n^{[i]} = b\big] +$$

$$+ \sum_{b\in\mathbb{B}} \Pr\big[D(\xi_n^{[1\,\ldots i-1]} b \upsilon_m^{[i+1\,\ldots m]}) \neq 1,\ \upsilon_m^{[i]} = b,\ \overline{\xi_n^{[i]}} = b\big] =$$

$$= \tfrac{1}{2}\Pr\big[D(\xi_n^{[1\,\ldots i]}\upsilon_m^{[i+1\,\ldots m]}) = 1\big] \qquad + \tfrac{1}{2}\Pr\big[D(\xi_n^{[1\,\ldots i-1]}\overline{\xi_n^{[i]}}\upsilon_m^{[i+1\,\ldots m]}) \neq 1\big] =$$

$$= \tfrac{1}{2}\delta_i(n) + \tfrac{1}{2}(1 - 2\delta_{i-1}(n) + \delta_i(n)) = \tfrac{1}{2} + \delta_i(n) - \delta_{i-1}(n).$$

Здесь использовано следующее соображение:

$$\delta_{i-1}(n) = \Pr\big[D(\xi_n^{[1\,\ldots i-1]}\upsilon_m^{[i\,\ldots m]}) = 1\big] =$$

$$\Pr\big[D(\xi_n^{[1\,\ldots i-1]}\xi_n^{[i]}\upsilon_m^{[i+1\,\ldots m]}) = 1,\ \upsilon_m^{[i]} = \xi_n^{[i]}\big] + \Pr\big[D(\xi_n^{[1\,\ldots i-1]}\overline{\xi_n^{[i]}}\upsilon_m^{[i+1\,\ldots m]}) = 1,\ \upsilon_m^{[i]} = \overline{\xi_n^{[i]}}\big] =$$

$$= \tfrac{1}{2}\Pr\big[D(\xi_n^{[1\,\ldots i]}\upsilon_m^{[i+1\,\ldots m]}) = 1\big] + \tfrac{1}{2}\Pr\big[D(\xi_n^{[1\,\ldots i-1]}\overline{\xi_n^{[i]}}\upsilon_m^{[i+1\,\ldots m]}) = 1\big] \ \Rightarrow$$

$$\Rightarrow\ \Pr\big[D(\xi_n^{[1\,\ldots i-1]}\overline{\xi_n^{[i]}}\upsilon_m^{[i+1\,\ldots m]}) = 1\big] = 2\delta_{i-1}(n) - \delta_i(n).$$

Доказательство теоремы Яо ($\Leftarrow$)

Пусть теперь ***i* случайно**, $i \in_{\mathcal{U}} \{1, \ldots, m(n)\}$:

$$\Pr_i\left[P(1^n, \xi_n^{[1\ldots i-1]}) = \xi_n^{[i]}\right] = \frac{1}{m(n)} \sum_{k=1}^{m(n)} \Pr\left[P(1^n, \xi_n^{[1\ldots k-1]}) = \xi_n^{[k]}\right] =$$

$$= \frac{1}{2} + \frac{1}{m(n)} \sum_{k=1}^{m(n)} (\delta_k(n) - \delta_{k-1}(n)) = \frac{1}{2} + \frac{1}{m(n)} (\delta_{m(n)}(n) - \delta_0(n)) =$$

$$= \frac{1}{2} + \frac{1}{m(n)} (\Pr\left[D(1^n, \xi_n) = 1\right] - \Pr\left[D(1^n, \upsilon_{m(n)}) = 1\right]) > \frac{1}{2} + \frac{1}{m(n) \cdot p(n)}$$

для бесконечно многих n — противоречие с условием непредсказуемости следующего бита. $\square$

NB

Во второй части доказательства использован т. н. *гибридный аргумент (метод)*:
для двух семейств случайных величин построена последовательность из полиномиального числа гибридов, «плавно» переводящих одно семейство в другое;
из попарной неразличимости соседних гибридов следует неразличимость крайних гибридов, по распределению совпадающих с исходными семействами.

Достаточное условие существования псевдослучайных генераторов

Утверждение (Yao)

Если существует односторонняя перестановка, то существует псевдослучайный генератор.

Схема доказательства. Пусть f — односторонняя перестановка.

- Продолжим её на всё $\mathbb{B}^*$ и построим f', как в теореме Гольдрайха—Левина.
- $f' : \bigcup_n \mathbb{B}^{2n} \to \mathbb{B}^*$ — односторонняя перестановка с трудным предикатом $b(\cdot)$.
- Определим $g : x \mapsto f'(x)b(x), \quad x \in \mathbb{B}^*$.
- Далее доказывается, что g — псевдослучайный генератор. □

NB То, что f перестановка, нужно не только для обеспечения правильных длин значений, но и для того, чтобы $f(x)$ было равномерно распределено на $\mathbb{B}^n$ при $x \in_{\mathcal{U}} \mathbb{B}^n$.

NB Другой вариант построения g из функции f и трудного предиката b для f

$$g(x) = b(x)b(f(x))b(f(f(x))) \ldots b(f^{m(|x|)-1}(x))$$

даёт псевдослучайный генератор со значениями полиномиальной от $|x|$ длины.

Даже конструкции, увеличивающей длину строки на один бит, оказывается достаточно, чтобы говорить о наличии псевдослучайных генераторов с произвольным $m(n)$.

Теорема «о растягивании выхода» псевдослучайного генератора

Теорема

Если существует некоторый псевдослучайный генератор g', то для любого полинома $m(\cdot)$ с условием $m(n) > n$ по всем $n \in \mathbb{N}$ существует такой псевдослучайный генератор g'', что $g(\mathbb{B}^n) \subset \mathbb{B}^{m(n)}$ для всех $n \in \mathbb{N}$.

Схема доказательства.

(1) Определим g на всех $x \in \mathbb{B}^n$ для любого $n \in \mathbb{N}$:

$$g(x) = g'(x)^{[1\,\ldots n+1]}.$$

Тогда g — псевдослучайный генератор.

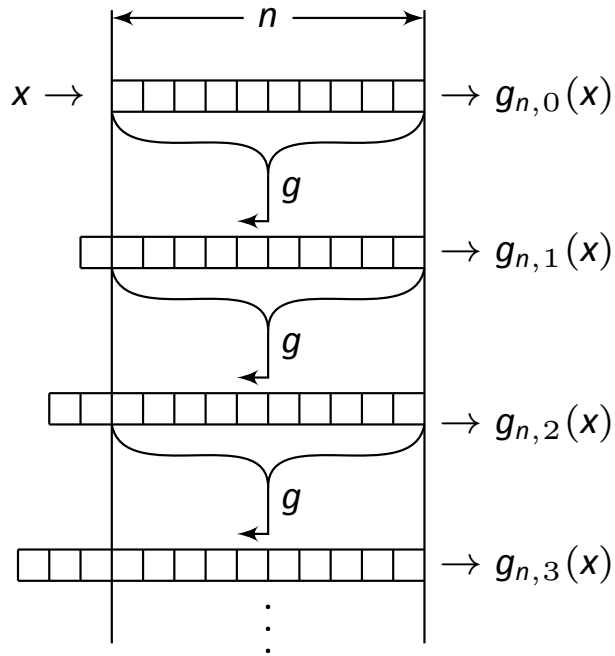
(2) Определим g'' на всех $x \in \mathbb{B}^n$ для любого $n \in \mathbb{N}$:

$$g_{n,0}(x) = x$$
$$g_{n,1}(x) = g(x) = g(x)^{[1]} g_{n,0}(g(x)^{[2\,\ldots n+1]})$$
$$\vdots$$
$$g_{n,k+1}(x) = g(x)^{[1]} g_{n,k}(g(x)^{[2\,\ldots n+1]}),\ k \geqslant 0$$

($g_{n,k}$ увеличивает длину на k)

$$g''(x) = g_{n,m(n)-n}(x)$$

g'' — псевдослучайный генератор. □



Необходимое условие существования псевдослучайного генератора

Утверждение

Пусть g — псевдослучайный генератор, $\forall n \in \mathbb{N}\ \ g(\mathbb{B}^n) \subset \mathbb{B}^{2n}$.
Тогда f, определённая равенством $f(xy) = g(x)$ для всех $x, y \in \mathbb{B}^n,\ n \in \mathbb{N}$, — односторонняя функция.

Доказательство. Очевидно, что f полиномиально вычислима. Докажем её односторонность от противного:
пусть для п. в. м. Т. A, некоторого полинома p и бесконечно многих n

$$\Pr\left[A\left(f(\upsilon_{2n})\right) \in f^{-1}\left(f(\upsilon_{2n})\right)\right] \geqslant \frac{1}{p(n)}.$$

Построим «различитель» D: на входе $w \in \mathbb{B}^*$

1. $z = A(w)$,
2. $D(w) = \begin{cases} 1, & \text{если } f(z) = w, \\ 0 & \text{иначе.} \end{cases}$

По определению f, $f(\upsilon_{2n}) = g(\upsilon_n)$ как случайные величины $\Longrightarrow$

$$\Pr\left[D(g(\upsilon_n)) = 1\right] = \Pr\left[f(A(g(\upsilon_n))) = g(\upsilon_n)\right] = \Pr\left[A\left(f(\upsilon_{2n})\right) \in f^{-1}\left(f(\upsilon_{2n})\right)\right] \geqslant \frac{1}{p(n)}.$$

Теорема Хостада—Импальяццо—Левина—Луби

Итак, $\Pr\bigl[D(g(\upsilon_n)) = 1\bigr] \geqslant \frac{1}{p(n)}$ для бесконечно многих n. С другой стороны, по определению f, $|f(\mathbb{B}^{2n})| = |g(\mathbb{B}^n)| \leqslant 2^n \quad \Longrightarrow$

$$\begin{aligned}\Pr\bigl[D(\upsilon_{2n}) = 1\bigr] &= \Pr\bigl[f(A(\upsilon_{2n})) = \upsilon_{2n}\bigr] \leqslant \Pr\bigl[\upsilon_{2n} \in f(\mathbb{B}^{2n})\bigr] \\ &= \frac{|f(\mathbb{B}^{2n})|}{|\mathbb{B}^{2n}|} \leqslant \frac{2^n}{2^{2n}} = \frac{1}{2^n}.\end{aligned}$$

Значит, для бесконечно многих n

$\Pr\bigl[D(g(\upsilon_n)) = 1\bigr] - \Pr\bigl[D(\upsilon_{2n}) = 1\bigr] \geqslant \frac{1}{p(n)} - \frac{1}{2^n} > \frac{1}{2p(n)}$ — противоречие с псевдослучайностью $g(\upsilon_n)$. □

NB С помощью теоремы «о растягивании выхода» можно снять ограничения на g в условии утверждения.

Теорема (Håstad, Impagliazzo, Levin, Luby)

Псевдослучайные генераторы существуют тогда и только тогда, когда существуют односторонние функции.