Лекция 2

Универсальная односторонняя функция. Перестановки с секретом. Трудный бит

(Конспект: А. Богатов)

Частично использован также конспект А. Куликова 2005 года.

2.1 Кандидаты в односторонние функции

Пусть $\circ$ обозначает конкатенацию строк.

- $f(x \circ y) = x \cdot y$ (сложно раскладывать длинные числа на множители (особенно на 2 простых множителя)),
- $f(x_1 \circ x_2 \circ \ldots \circ x_n \circ I) = (x_1 \circ x_2 \circ \ldots \circ x_n \sum_{i \in I} x_i)$, где $I \subseteq \{1, \ldots, n\}$ (это NP-трудная задача SUBSET SUM; в некотором смысле, задача о рюкзаке — нужно определить, какими элементами из заданного набора набирается заданная сумма).

2.2 Универсальная односторонняя функция

Определение 2.1. Алгоритм A взламывает функцию G с вероятностью $q(n)$, если для бесконечной последовательности длин n_i

$$\Pr_{|x|=n_i} \{G(A(G(x))) = x\} \geq q(n_i).$$

Определение 2.2. $F \Rightarrow G$ (сильный взлом[1] функции F сводится к сильному взлому функции G), если

$$\exists T^{\cdot}\ \forall p'\ \exists p:$$

$$A \text{ взламывает } G \text{ с вероятностью } 1 - \frac{1}{p(n)} \Rightarrow$$

$$T^A \text{ взламывает } F \text{ с вероятностью } 1 - \frac{1}{p'(n)}.$$

Здесь T — полиномиальный вероятностный алгоритм, A используется как *вероятностный* оракул (его случайные биты учитываются при запуске T^A), $p(n)$ и $p'(n)$ —многочлены (положительные при $n \geq 1$).

Определение 2.3. Функция G, вычислимая за полиномиальное время, называется универсальной[2] слабой owf[3], если для любой функции F, вычислимой за полиномиальное время, $F \Rightarrow G$.

Очевидно, что если существует слабая owf, то универсальная слабая owf действительно является слабой owf.

Теорема 2.1. *Пусть $U(M, x) = (M, M(x))$, где M — описание машины, x — вход для этой машины, $M(x)$ — выход машины на входе x, причем моделируем мы в течение времени $|x|^2$, а если не успеваем завершить работу, выдаем x. Тогда взлом любой слабой owf сводится к взлому U.*

Упражнение 2.1. А какое утверждение верно для сильной owf?

Лемма 2.1. *$\forall$ слабой owf f $\exists \tilde{f}$, вычислимая за время $|x|^2$, к взлому которой сводится взлом f.*

Доказательство. Есть функция f — слабая owf; по ней строим $\tilde{f}$, которая заканчивает работу за время, ограниченное квадратом от длины входа:

$$\tilde{f}(x_1 x_2) = f(x_1) x_2.$$

Для того, чтобы $\tilde{f}$ заканчивала работу за квадратичное время, достаточно, чтобы выполнялось неравенство $t_f(x_1) \leq |x_1 x_2|^2 - |x_2|$, где $t_f(x_1)$ – время работы f на входе x_1. Пусть $|x_1| = n$, $|x_2| = m$. Мы можем добиться $t_f(n) \leq (m+n)^2 - m$ выбором подходящего m как многочлена от n, поскольку $t_f(n)$ также ограничено многочленом от n.

[1] Сильный взлом ломает слабые owf.

[2] На самом деле, это обычное понятие полноты.

[3] Вообще-то она может и не являться при этом слабой owf — если таковых нет в природе.

Пусть мы умеем ломать $\tilde{f}$. Тогда функция f ломается следующим образом: берем то, что нам дали (т.е. некоторое значение $f(x_1)$), дописываем случайную строку x_2, ломаем (получая тем самым x_1x_2), убираем с конца x_2. Тем самым, если $\tilde{f}$ мы ломали с вероятностью $1 - \frac{1}{(m+n)^k}$, то и f мы ломаем с вероятностью $1 - \frac{1}{(m+n)^k}$. Поскольку m — фиксированный полином от n, ясно, что мы можем добиться любой необходимой вероятности взлома f, выбирая k. $\square$

Доказательство теоремы 2.1. Рассмотрим произвольную owf M^*, которая заканчивает работу за время $|x|^2$, и покажем, что $M^* \Rightarrow U$. Для достаточно длинных входов машины U она запускает машину M^* на доле входов $\mu = \frac{1}{2^{\lceil M^* \cdot \mathrm{const} \rceil}} = \mathrm{const}$. Если мы не взламываем лишь долю $\frac{1}{n^k}$ от всех входов U, то должны взламывать значительную долю входов из сектора, соответствующего машине M^*; именно, мы взламываем долю $\mu - \frac{1}{n^k}$, что составляет

$$1 - \frac{1}{\mu n^k} \tag{2.1}$$

по отношению ко всем входам машины M^* длины $n - |M^*|$. Ясно, что для любой требуемой вероятности взлома M^* мы можем подобрать достаточно большие k и n, для которых (2.1) будет больше искомой. $\square$

Упражнение 2.2. Что произойдет в случае семейств односторонних функций (сильных либо слабых)?

Упражнение 2.3. Что произойдет, если соперник — детерминированный? Если он задан схемами?

Упражнение 2.4. Доказать, что если существует owf, то существует и *не*универсальная owf.

2.3 Функции с секретом (trapdoor functions)

Понятие «функция с секретом» почти бессмысленно. Поэтому будем рассматривать *семейства* таких функций. Ограничимся инъективными функциями (перестановками).

Определение 2.4. Односторонняя функция с секретом (trapdoor permutation family, tdpf) — это полиномиальный по времени алгоритм

$$G:\ (1^n, r_g) \mapsto (e, d, s),$$

где n — параметр надежности (он же у нас будет длиной входа), r_g — строка случайных битов генератора, e, d, s — булевы схемы (d — [секретный] decryptor, e — [публичный] encryptor, s — [публичный] sampler),

- $e\colon \{0,1\}^n \to \{0,1\}^{\varepsilon(n)}$,
- $d\colon \{0,1\}^{\varepsilon(n)} \to \{0,1\}^n$,
- $s(r_s) \in \{0,1\}^n$,

и

$$\forall A\ \forall p\ \exists n\ \forall n > N\ \Pr\{A(1^n, e(x), s, e) \in e^{-1}(e(x))\} < \frac{1}{p(n)}$$

(здесь $x = s(r_s)$; $(e,d,s) = G(1^n, r_g)$; вероятность берется по r_g, r_s, случайным битам A),

$$\forall x \in \operatorname{Im} s \quad d(e(x)) = x.$$

Если из этого определения убрать d, то получим семейство односторонних функций (по умолчанию сильных).

На основе tdpf строятся криптосистемы с открытым ключом.

Упражнение 2.5. Изменится ли что-то существенное, если s станет функцией от e?

Упражнение 2.6. А если никакого s не будет (т.е. $s = \mathrm{Id}$)?

Упражнение 2.7. Выполнить для tdpf те упражнения, что были для owff.

Упражнение 2.8. Что, если d либо e использует случайные биты и иногда ошибается?

Пример 2.1 (RSA).

$$s(x) = x,$$

$$e(x) = x^{\varepsilon} \bmod n, \quad d(x) = x^{\delta} \bmod n,$$

где

$$\varepsilon \cdot \delta \equiv 1 \pmod{(p-1)(q-1)},$$

$$n = pq, \quad p, q \in \mathbb{P}.$$

Является ли такое семейство tdpf, неизвестно, но на нём основаны реально использующиеся протоколы.

2.4 Трудный бит

Пусть $y = f(x)$; если противник не сможет вычислить x, но может, например, узнать все нечетные биты y, это также нехорошо.

Определение 2.5. $B\colon \{0,1\}^n \to \{0,1\}$ называется *трудным битом* (*hardcore predicate*) для функции f, если

$$\forall k\ \forall A\ \exists N\ \forall n>N \quad \Pr\{A(f(x)) = B(x)\} < \frac{1}{2} + \frac{1}{n^k}, \tag{2.2}$$

где A — вероятностный полиномиальный по времени противник, а вероятность в определении берется по его случайным числам и по $x \in \{0,1\}^n$.

Оказывается, из любой инъективной owf можно сделать такую (инъективную) owf, у которой есть трудный бит. (В частности, это же можно проделать и для семейства перестановок с секретом.)

Упражнение 2.9. Использует ли нижеприведённое доказательство тот факт, что $|f(x)| = |f(x')|$, если $|x| = |x'|$?

Теорема 2.2 (Голдрейха-Левина). *Если f является инъективной owf, то $\tilde{f}(x,r) = (f(x), r)$ тоже является односторонней и имеет трудный бит $B(x,r) = \langle x, r\rangle$, где $\langle x,r\rangle = x_1 r_1 \oplus x_2 r_2 \oplus \ldots$.*

Доказательство. Пусть мы умеем угадывать трудный бит. Построим противника, ломающего f. Казалось бы,

$$x_i = \langle f^{-1}(y), r\rangle \oplus \langle f^{-1}(y), r \oplus e_i\rangle = B(x,r) \oplus B(x, r\oplus e_i)$$

($r \oplus e_i$ означает, что мы поменяли i-й бит в r), так что мы можем угадать любой бит x_i из x. Однако правильное вычисление противником $B(x,r)$ и $B(x, r\oplus e_i)$ – зависимые события. Поэтому $B(x,r)$ мы не будем у него выяснять — это один и тот же бит для всех i, и мы можем перебрать два его возможных значения. А вот $B(x, r\oplus e_i)$ — свой для каждого i.

На этом можно было бы уже остановиться, если бы нам не предстояло уменьшать вероятность ошибки противника, повторяя его для разных r. Это бы привело к очень большому перебору; поэтому мы будем проделывать не совсем независимые эксперименты, выбрав лишь логарифмическое число случайных строк r^i; благодаря приведённой ниже конструкции мы сможем породить из них много попарно независимых строк, трудные биты $B(x,r)$ для которых будут просто вычисляться через перебираемые нами трудные биты для исходных r^i.

Что же касается x, мы можем безбоязненно повторять вычисления для одного и того же x (вернее, $f(x)$) несмотря на то, что вероятность в определении трудного бита берётся по всем x. Дело в том, что тех x, для которых вероятность успеха противника достаточно велика, много, как доказывается в следующей лемме.

Лемма 2.2. *Пусть соперник ломает наш трудный бит с вероятностью $1/2+\varepsilon$ (т.е., в терминах (2.2) $\varepsilon = \varepsilon(n) = 1/n^k$). Пусть*

$$S_n = \{x \mid \Pr\{A(f(x), r) = B(x, r)\} \geq \frac{1}{2} + \frac{\varepsilon}{2}\}.$$

Тогда $|S_n| \geq \frac{\varepsilon}{2} \cdot 2^n$.

Доказательство леммы.

$$S(x) := \Pr\{A(f(x), r) = B(x, r)\}$$

$$|\overline{S_n}| = 2^n \Pr_x \left\{S(x) < \frac{1}{2} + \frac{\varepsilon}{2}\right\} = 2^n \Pr_x \left\{1 - S(x) \geq \frac{1}{2} - \frac{\varepsilon}{2}\right\}$$

$$E(1 - S(x)) = 1 - \left(\frac{1}{2} + \varepsilon\right) = \frac{1}{2} - \varepsilon$$

Неравенство Маркова:

$$\Pr\{\alpha > \alpha'\} \leq \frac{E\alpha}{\alpha'},$$

где α – неотрицательная случайная величина.
У нас $E\alpha = \frac{1}{2} - \varepsilon, \quad \alpha' = \frac{1}{2} - \frac{\varepsilon}{2}$.

$$2^n \frac{\frac{1}{2} - \varepsilon}{\frac{1}{2} - \frac{\varepsilon}{2}} = 2^n \frac{1 - 2\varepsilon}{1 - \varepsilon} \leq 2^n \left(1 - \frac{\varepsilon}{2}\right)$$

Лемма доказана. □

Итак, опишем конструкцию, обращающую f при помощи взломщика для B, формально. Положим $l = (2k+2)\lceil \log_2 n \rceil$ (если вероятность успеха противника составляет $1/2 + 1/n^k$) и выберем l случайных строчек в соответствии с равномерным распределением: $r^1, \ldots, r^l$. Эти строки — кандидаты на роль r. Выберем также l битов (обозначим их $\rho^1, \ldots, \rho^l$), после чего проделаем следующее: для всех непустых подмножеств J множества $\{1, \ldots, l\}$ вычислим

$$r^J = \bigoplus_{j \in J} r^j,$$

$$\rho^J = \bigoplus_{j \in J} \rho^j$$

(заметим, что если ρ^j — правильные биты для r^j, то ρ^J — правильные биты для r^J), и далее для всех i вычислим

$$x_i^J = \rho^J \oplus \bar{B}(y, r^J \oplus \bar{e}_i),$$
$$x'_i = \operatorname*{maj}_J x_i^J.$$

Лемма 2.3. *Величины r^J из доказательства теоремы равномерно распределены и попарно независимы.*

Доказательство. То, что они равномерно распределены, очевидно. Если $K \subseteq J$, то

$$\begin{aligned}
&\mathbf{P}\{r^J = t,\ r^K = t'\} = \\
&\qquad \mathbf{P}\{r^{J\setminus K} = t \oplus t',\ r^K = t'\} \overset{(J\setminus K)\cap K=\emptyset}{=} \\
&\qquad \mathbf{P}\{r^{J\setminus K} = t \oplus t'\} \cdot \mathbf{P}\{r^K = t'\} \overset{\substack{\text{равномерно}\\ \text{распределены!}}}{=} \\
&\qquad \mathbf{P}\{r^J = t\} \cdot \mathbf{P}\{r^K = t'\}.
\end{aligned}$$

Значит, можно считать, что $J \setminus K \neq \emptyset$ и $K \setminus J \neq \emptyset$. Тогда

$$\begin{aligned}
&\mathbf{P}\{r^J = t,\ r^K = t'\} = \\
&\qquad \sum_{t''} \mathbf{P}\{r^J = t,\ r^K = t',\ r^{J\cap K} = t''\} = \\
&\qquad \sum_{t''} \mathbf{P}\{r^{J\setminus K} = t,\ r^{K\setminus J} = t',\ r^{J\cap K} = t''\} = \\
&\qquad \mathbf{P}\{r^{J\setminus K} = t\} \cdot \mathbf{P}\{r^{K\setminus J} = t'\} \cdot \underbrace{\sum_{t''} \mathbf{P}\{r^{J\cap K} = t''\}}_{1} \overset{\substack{\text{равномерно}\\ \text{распределены!}}}{=} \\
&\qquad \mathbf{P}\{r^J = t\} \cdot \mathbf{P}\{r^K = t'\}.
\end{aligned}$$

□

Для фиксированного i оценим вероятность того, что среди x_i^J было больше половины правильных (т.е. что $x'_i = x_i$). Обозначим через

$$\zeta_i^J = \{x_i = x_i^J\}$$

вероятность успеха в одном испытании (однократном вычислении $B(x, r')$). Обозначим $m = 2^l - 1$.

Лемма 2.4. *Для достаточно больших* n

$$\Pr\left\{\sum_J \zeta_i^J \le \frac{m}{2}\right\} < \frac{1}{2n}.$$

Доказательство. Вероятность успеха в одном испытании равна $\frac{1}{2} + \frac{\varepsilon}{2}$, если $x \in S_n$ (а мы знаем, что S_n достаточно велико и можно им ограничиться). Испытания попарно независимы, поэтому

$$E \sum \zeta_i^J = m\left(\frac{1}{2} + \frac{\varepsilon}{2}\right) \Rightarrow \frac{m}{2} = E - \frac{m\varepsilon}{2}$$

Применим неравенство Чебышёва ($\Pr\{\alpha < E\alpha - \delta\} < \frac{D\alpha}{\delta^2}$):

$$\Pr\left\{\sum_J \zeta_i^J < E - \frac{m\varepsilon}{2}\right\} < \frac{4D\sum \zeta_i^J}{m^2\varepsilon^2} < \frac{4}{m\varepsilon^2} \le \frac{4}{n^2}$$

для достаточно больших n (здесь использовано, что благодаря попарной независимости $D\sum \zeta_i^J = mD\zeta_i^J < m$). □

И утверждение теоремы можно считать доказанным.

□

Упражнение 2.10. Убедиться, что утверждение теоремы выполнено и для неинъективной owf.

Упражнение 2.11. Конструкция использует известную ей вероятность успеха противника; как от этого избавиться?