Курс лекций по теоретической криптографии

Тема 12. **Интерактивные доказательства с нулевым разглашением**

Шокуров А.В.

2 декабря 2023 г.

Доказательства с нулевым разглашением

Предположим, что Алиса знает доказательство некоторой теоремы и желает убедить Боба, что теорема верна. Самое простое отдать доказательство на проверку Бобу. Но тогда Боб сможет доказывать эту теорему другим участникам самостоятельно. А можно ли убедить Боба так, чтобы не сообщить ему никакой информации, на основании которой он научился бы доказывать эте теорему самостоятельно? Этим двум противоречивым требованиям удовлетворяют протоколы с нулевым разглашением. Это понятие было введено Гольдвассер, Микали и Ракоффом в 1985 году.

Аутентификация

Предположим, что Алиса хочет подтвердить Бобу свою аутентичнность, например, тем, что она знает свой секретный ключ (соответствующий известному открытому ключу данного участника Бобу)и при этом не раскрыть этот ключ Бобу. Он пытается это осуществить с помощью следующего протокола.

Протокол Шнорра интерактивной аутентификации

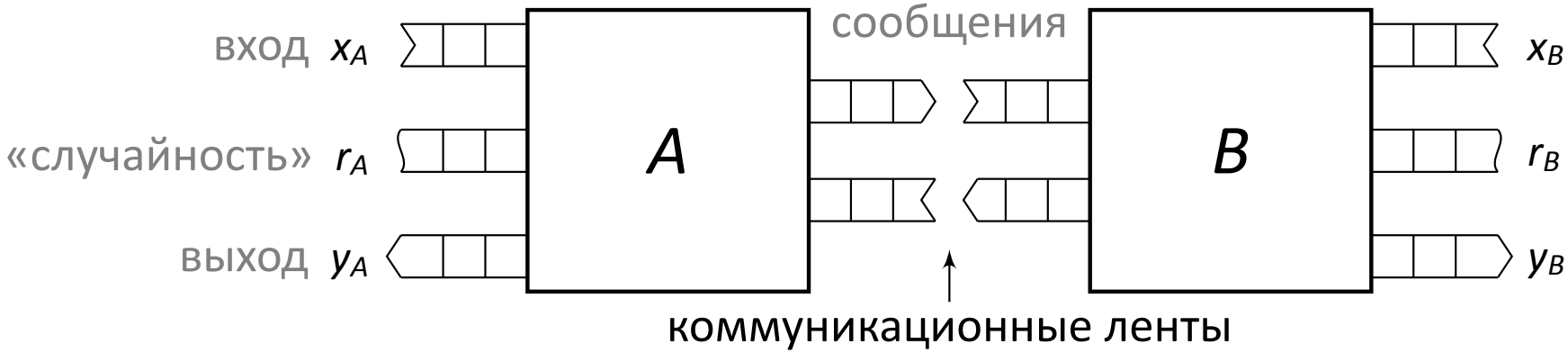
Пусть g — порождающий циклической группы порядка $p-1$,
K — некоторое множество вычетов по модулю p.

A		B	
$\hat{k} \in_{\mathcal{R}} K,\ k = g^{-\hat{k}}$			$(k, \hat{k})$ — открытый и секретный ключи
$\hat{x} \in_{\mathcal{R}} K$			$\hat{x}$ — сеансовый секретный ключ
$x = g^{\hat{x}}$	$\xrightarrow{x}$		x — сеансовый открытый ключ
	$\xleftarrow{h}$	$h \in_{\mathcal{R}} K$	h — challenge
$s = \hat{x} + \hat{k}h$	$\xrightarrow{s}$	$x \stackrel{?}{=} g^s k^h$	$g^{\hat{x}+\hat{k}h} \cdot g^{-\hat{k}h} = g^{\hat{x}} = x$

Это протокол доказательства знания дискретного логарифма $-\log_g k$.

! Здесь нет внешнего противника, но участники протокола, вообще говоря, не доверяют друг другу.
Должна быть обеспечена защита и от нечестного участника, пытающегося доказать (ложную) аутентичность, и от нечестного проверяющего (участвуя в выполнении протокола, он не должен научиться доказывать знание секретного ключа).

Интерактивная пара машин Тьюринга



- Вход может быть общим у обеих машин, но они могут иметь и свои частные входы.
- Раунд $\sim$ период активности одной машины с записью слова (сообщения) на коммуникационную ленту. Если раунд только один, то соответствующий протокол называется *неинтерактивным* (*unidirectional*).
- Выход пары (результат выполнения протокола) (A, B) — это (y_A, y_B) (или только y_B, если машина B останавливается последней).
- «Интерактивная» машина *полиномиальна*, если время её работы (число тактов до остановки) при совместном вычислении с любой другой машиной Тьюринга на общем входе x ограничено величиной $\mathrm{poly}(|x|)$. В этом случае число раундов полиномиаль- но, но размер входящих сообщений — вообще говоря, необязательно.

Интерактивное доказательство

P — доказывающий (prover),
V — проверяющий (verifier), выдающий 1 или 0 (доказательство принимается или нет);
$L \subseteq \mathbb{B}^*$ — язык, соответствующий некоторой распознавательной задаче.

Определение

Интерактивное доказательство (*протокол интерактивного доказательства*) для языка L — это интерактивная пара машин Тьюринга (P, V) с полиномиальной V, для которых выполнены условия:

- (полнота / completeness) для любого $x \in L$ $\quad \Pr\big[(P, V)(x) = 1\big] \geqslant \frac{2}{3}$,
- (корректность / soundness) для любой интерактивной в. м. Т. P' и для любого $x \notin L$ $\quad \Pr\big[(P', V)(x) = 1\big] \leqslant \frac{1}{3}$.

Определение (с дополнительными частными входами)

Интерактивное доказательство для языка L — это интерактивная пара машин Тьюринга (P, V), где V — полиномиальная машина, для которых выполнены условия:

- (полнота) $\forall x \in L \ \exists x_P \ \forall x_V$ $\quad \Pr\big[\big(P(x_P), V(x_V)\big)(x) = 1\big] \geqslant \frac{2}{3}$,
- (корректность) для любой интерактивной в. м. Т. P' и $\forall x \notin L \ \forall x_P \ \forall x_V$ $\quad \Pr\big[\big(P'(x_P), V(x_V)\big)(x) = 1\big] \leqslant \frac{1}{3}$.

Интерактивое доказательство

NB

1. При замене границ полноты и корректности соответственно на $1 - \frac{1}{2^{q(|x|)}}$ и $\frac{1}{2^{q(|x|)}}$ для произвольного полинома $q(\cdot)$ получаются эквивалентные определения.
2. P можно считать детерминированной машиной: отсутствие ограничения на время работы позволяет ему перебрать все вероятные возможности.

Полнота означает, что если входное слово принадлежит языку L и оба участника следуют протоколу, то доказательство всегда будет принято.
Второе требование — требование корректности, защищает Боба от нечестной Алисы, которая пытается его обмануть, ”доказывая”ложное утверждение. При этом Алиса может отклоняться от протокола, в том числе используя вместо машины P другую машину Тьюринга P^*. Требуется, чтобы вероятность обмана была минимальной.

Сложностной класс IP

Определение

IP — класс языков, имеющих интерактивное доказательство.

? Покажите, что $\mathrm{NP} \subseteq \mathrm{IP}$ и $\mathrm{BPP} \subseteq \mathrm{IP}$.

Теорема (Shamir)

$$\mathrm{IP} = \mathrm{PSPACE}$$

PSPACE — класс всех языков, распознаваемых машинами Тьюринга, которые задействуют не более чем полиномиальное (от длины входа) число ячеек на своей рабочей ленте.

Свойство нулевого разглашения

В определении интерактивного доказательства заложена «защита» проверяющего от нечестного доказывающего P'. Как быть, если проверяющий нечестный и захочет впоследствии восстановить доказательство, чтобы демонстрировать его другим?

Свойство нулевого разглашения:
нечестный проверяющий V' может получить из общения с P только то, что мог бы вычислить и сам V', если бы знал, что доказываемое утверждение ($x \in L$) верно.

Это свойство определяется в терминах неотличимости реальной транскрипции взаимодействия P и V' от транскрипции, сгенерированной некоторым полиномиальным симулятором, не взаимодействовавшим с P.

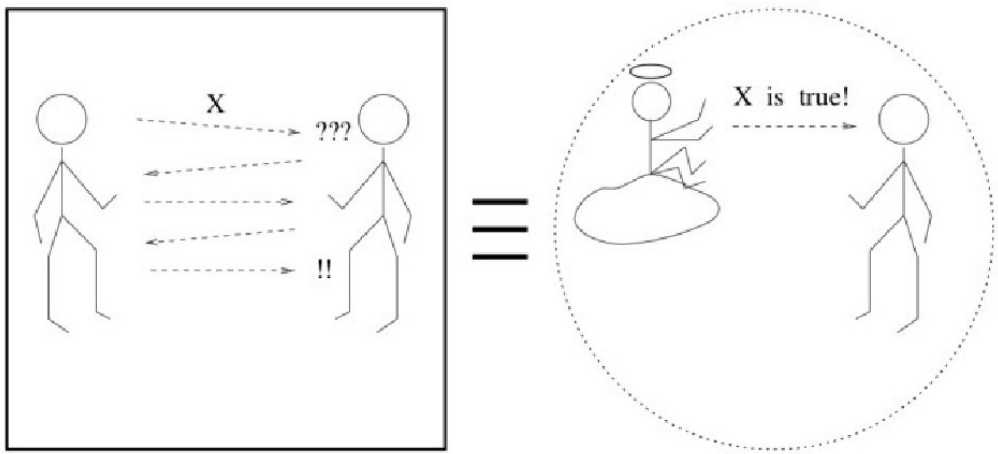


Иллюстрация из книги: *O. Goldreich.* Foundations of cryptography. Volume 1 (Basic tools). — Cambridge University Press, 2001. — P. 184.

Понятие неотличимости можно формализовать разными способами.

Понятия неотличимости

ξ, ζ — случайные величины со значениями в конечном или счётном множестве X.

Статистическое расстояние между ξ и ζ:

$$\Delta(\xi, \zeta) = \frac{1}{2} \sum_{z \in X} \left| \Pr\left[\xi = z\right] - \Pr\left[\zeta = z\right] \right|,$$

$$\Delta(\xi, \zeta) = \max_{Z \subseteq X} \left| \Pr\left[\xi \in Z\right] - \Pr\left[\zeta \in Z\right] \right|.$$

Пусть $\{\xi_x\}_{x \in L}, \{\zeta_x\}_{x \in L}$ — семейства случайных величин, индексированных строками из бесконечного множества $L \subseteq \mathbb{B}^*$.

Определение

Семейства $\{\xi_x\}, \{\zeta_x\}$ называются *статистически неразличимыми* (*statistically close / indistinguishable*), если $\Delta(\xi_x, \zeta_x) = \mathrm{negl}(|x|)$.

Определение

Семейства $\{\xi_x\}, \{\zeta_x\}$ называются *вычислительно неразличимыми* (*computationally indistinguishable*), если для любой п. в. м. Т. D

$$\left| \Pr\left[D(x, \xi_x) = 1\right] - \Pr\left[D(x, \zeta_x) = 1\right] \right| = \mathrm{negl}(|x|).$$

Интерактивные доказательства с нулевым разглашением

Рассмотрим выполнение интерактивной пары машин Тьюринга (P, V) и обозначим через $view_V^P(x)$ транскрипцию этого выполнения, состоящую из использованного префикса случайной строки машины V, последовательности пересылаемых сообщений в хронологическом порядке, выхода (P, V).
Свойство нулевого разглашения для интерактивного доказательства (P, V) определяется как неотличимость семейств случайных величин $\{view_{V'}^P(x)\}_x$ для произвольного нечестного проверяющего V' и $\{S(x)\}_x$ для некоторого полиномиального симулятора S, не имеющего доступа к P.

Интерактивные доказательства с нулевым разглашением

Определение

Интерактивное доказательство (P, V) для бесконечного языка L называется *доказательством с абсолютно нулевым разглашением* (*perfect zero-knowledge proof*), если для любой интерактивной п. в. м. Т. V' существует п. в. м. Т. S, такая, что семейства $\{view^P_{V'}(x)\}_{x\in L}$ и $\{S(x)\}_{x\in L}$ распределены одинаково,
а точнее, для любого $x \in L$

- $\Pr\left[S(x) \neq \perp\right] \geqslant \frac{1}{\mathrm{poly}(|x|)}$,
- при условии $S(x) \neq \perp$ $\quad \Pr\left[S(x) = z\right] = \Pr\left[view^P_{V'}(x) = z\right]$.

NB Эквивалентно, можно говорить о некотором *полиномиальном в среднем* симуляторе:
« ...существует п. в ср. в. м. Т. S: для любого $x \in L$ $\quad \Pr\left[S(x) = z\right] = \Pr\left[view^P_{V'}(x) = z\right]$ ».

Интерактивные доказательства с нулевым разглашением

Определение

Интерактивное доказательство (P, V) для бесконечного языка L называется *доказательством со статистически нулевым разглашением*

(*statistical / almost-perfect zero-knowledge proof*),

если для любой интерактивной п. в. м. Т. V' существует п. в. м. Т. S, такая, что семейства $\{view^P_{V'}(x)\}_{x\in L}$ и $\{S(x)\}_{x\in L}$ статистически неразличимы.

Определение

Интерактивное доказательство (P, V) для бесконечного языка L называется *доказательством с (вычислительно) нулевым разглашением*

(*computational zero-knowledge proof*),

если для любой интерактивной п. в. м. Т. V' существует п. в. м. Т. S, такая, что семейства $\{view^P_{V'}(x)\}_{x\in L}$ и $\{S(x)\}_{x\in L}$ вычислительно неразличимы.

Соответствующие классы языков: PZK, $\mathrm{SZK} \equiv \mathrm{APZK}$, $\mathrm{ZK} \equiv \mathrm{CZK}$.

? Докажите, что справедлива цепочка вложений: $\mathrm{BPP} \subseteq \mathrm{PZK} \subseteq \mathrm{SZK} \subseteq \mathrm{ZK} \subseteq \mathrm{IP}$.

NB $\mathrm{ZK} = \mathrm{IP}$, если существует односторонняя функция [Ben-Or *et al.*]: «всё, что доказуемо, доказуемо с нулевым разглашением».

Интерактивное доказательство изоморфизма графов

Будем рассматривать графы с множеством вершин $U = \{1, \ldots, n\}$, $n \in \mathbb{N}$, закодированные естественным способом (матрицами смежности, списком рёбер и т. п.).
$L_{\mathrm{GI}} = \{(G_0, G_1) \mid G_0 \cong G_1\}$ — язык ИЗОМОРФИЗМ ГРАФОВ.

[$G_0 \cong G_1 \quad \Leftrightarrow \quad$ существует такая перестановка φ на U, что $\varphi(G_0) = G_1$.]
Пусть φ — изоморфизм между G_0 и G_1. Следующие четыре шага выполняются в цикле m раз, где m — число ребер графа, каждый раз с независимыми случайными величинами.

1. Алиса выбирает случайную перестановку π на множестве U и вычисляет $H = \pi G_1$ и посылает этот граф Бобу.
2. Боб выбирает случайный бит α и посылает его Алисе.
3. Если $\alpha = 1$, то Алиса посылает Бобу перестановку π, в противном случае — перестановку $\pi \circ \varphi$.
4. Если перестановка, полученная Бобом, не является изоморфизмом между G и H, то Боб останавливается и отвергает доказательство. В противном случае выполнение протокола продолжается.

Интерактивное доказательство изоморфизма графов

NB $L_{\mathrm{GI}} \in \mathrm{NP}$, но неизвестно, является ли он NP-полным, принадлежит ли к классу BPP.

Протокол Π_{GI} интерактивного доказательства для L_{GI}

Общий вход: (G_0, G_1), где $G_0 = (U, E_0)$, $G_1 = (U, E_1)$, $|U| = n$, $|E_0| = |E_1|$.

$|E_1|$ раз

P		V	
$\pi \in_{\mathcal{U}} S_n, \quad c \in_{\mathcal{U}} \mathbb{B}$			
$H = \pi(G_c)$	$\xrightarrow{H}$		H — новый граф, изоморфный
$\psi = \begin{cases} \pi, & \text{если } b = c \\ \pi \circ \varphi, & \text{если } b \neq c \end{cases}$	$\xleftarrow{b}$	$b \in_{\mathcal{U}} \mathbb{B}$	b — номер случайного из двух
(где φ — изоморфизм $G_{\bar{c}} \mapsto G_c$)	$\xrightarrow{\psi}$		ψ — изоморфизм $G_1 \mapsto H$ либо изоморфизм $G_0 \mapsto H$
		$\psi(G_b) \stackrel{?}{=} H$	

$(P, V)(G_0, G_1) = 1$, т. е. V принимает доказательство того, что $(G_0, G_1) \in L_{\mathrm{GI}}$, если и только если во всех $|E_1|$ циклах проверки $\psi(G_b) \stackrel{?}{=} H$ прошли успешно.

Нулевое разглашение при доказательстве изоморфизма графов

Утверждение

Π_{GI} — абсолютно полный [a] протокол интерактивного доказательства с абсолютно нулевым разглашением для языка L_{GI}.

---

[a] То есть с границей полноты 1.

Доказательство.

- Полнота: $G_0 \cong G_1$, P, V — честные $\Rightarrow H \cong G_0$ и $H \cong G_1$,

$$\Pr\big[(P, V)(G_0, G_1) = 1\big] = 1.$$

- Корректность: $G_0 \ncong G_1$, P' — нечестный $\Rightarrow \forall H$ либо $H \ncong G_0$, либо $H \ncong G_1$,

$$\Pr\big[(P', V)(G_0, G_1) = 1\big] \leqslant \tfrac{1}{2} \text{ в каждом цикле и } \leqslant \tfrac{1}{2^{|\varepsilon_1|}}.$$

- Нулевое разглашение: $G_0 \cong G_1$, V' — нечестный — произвольная инт. п. в. м. Т.

Симулятор S будет выполнять роль доказывающего для V' (не зная φ и не имея возможности его вычислить), и он должен выдать V' столько же информации, сколько и P, но ему не нужно убеждать V' в знании изоморфизма.

Определим S, как машину, цикл работы которой на входе (G_0, G_1) выглядит следующим образом:

Нулевое разглашение при доказательстве изоморфизма графов

1. выбрать $c' \in_{\mathcal{U}} \mathbb{B}$, $\pi' \in_{\mathcal{U}} S_n$,
   вычислить $H' = \pi'(G_{c'})$;
2. запустить $V'(r; G_0, G_1)$, где $r \in_{\mathcal{U}} \mathbb{B}^{l(n)}$, используя H' в первом раунде;
3. получив b' от V', проверить его: если $b' \notin \mathbb{B}$, положить $b' = 1$;
4. если $b' = c'$, послать V' перестановку π',
   после завершения V' возвратить $S(G_0, G_1) = (r, H', b', \pi', V'(r; G_0, G_1))$,
   если $b' \neq c'$, возвратить $\perp$.

Случайные величины $\pi'(G_0)$ и $\pi'(G_1)$ распределены равномерно на множестве графов, изоморфных данным $\Rightarrow$ H' распределён так же, как H в $(P, V')(G_0, G_1)$, и случайные величины c' и $H' = \pi'(G_{c'})$ независимы $\Rightarrow$ b' и c' независимы $\Rightarrow$

$$\Pr[b' = c'] = \Pr[b' = 0, c' = 0] + \Pr[b' = 1, c' = 1] = \tfrac{1}{2}\big(\Pr[b' = 0] + \Pr[b' = 1]\big) = \tfrac{1}{2}$$

$$\Rightarrow \quad \Pr[S(G_0, G_1) \neq \perp] = \tfrac{1}{2}.$$

Когда $S(G_0, G_1) \neq \perp$, случайная величина $S(G_0, G_1) = (r, H', b', \pi', V'(r; G_0, G_1))$ распределена так же, как $\mathit{view}^P_{V'}(G_0, G_1) = (r, H, b', \pi, V'(r; G_0, G_1))$. □

Доказательства с нулевым разглашением для языков из класса NP

Для криптографии особый интерес представляют те протоколы, в которых P полиномиален, если на его (частный) вход подаётся некая дополнительная информация.

Бинарное отношение $R \subseteq \mathbb{B}^* \times \mathbb{B}^*$ — NP-*отношение*, если R распознаётся некоторой п. д. м. Т. (то есть $R \in \mathrm{P}$) и $\forall x \, \forall y \;\; \big((x, y) \in R \Rightarrow |y| \leqslant \mathrm{poly}(|x|)\big)$.
Такой y называется тогда NP-*доказательством* для x относительно R.

Пусть $L_R = \{x \in \mathbb{B}^* \mid \exists\, y \in \mathbb{B}^* \;\; (x, y) \in R\}$, тогда $\mathrm{NP} = \{L_R \mid R - \text{NP-отношение}\}$.[1]

Теорема (Goldreich, Micali, Wigderson)

Пусть существует односторонняя функция.
Пусть R — NP-отношение, такое, что язык L_R бесконечен.
Тогда существует протокол интерактивного доказательства (P, V)
с вычислительно нулевым разглашением для языка L_R, в котором P работает за полиномиальное время на любом общем входе $x \in L_R$, если ему на вход подаётся произвольное NP-доказательство для x относительно R.

[1] Ср. с определением:
$\mathrm{NP} = \{L \subseteq \mathbb{B}^* \mid \exists R \in \mathrm{P} \;\; \exists_{\text{полином}}\, p \;\; \forall x \in \mathbb{B}^* \quad x \in L \Leftrightarrow (\exists\, y \in \mathbb{B}^{p(|x|)} \;\; (x, y) \in R)\}$.