Курс лекций по теоретической криптографии

Тема 13. **Криптосистемы с открытым ключом**

Шокуров А.В.

Определение

$n \in \mathbb{N}$ — параметр стойкости,

$M_n \subseteq \mathbb{B}^*$ — пространство сообщений (открытых текстов).

Определение

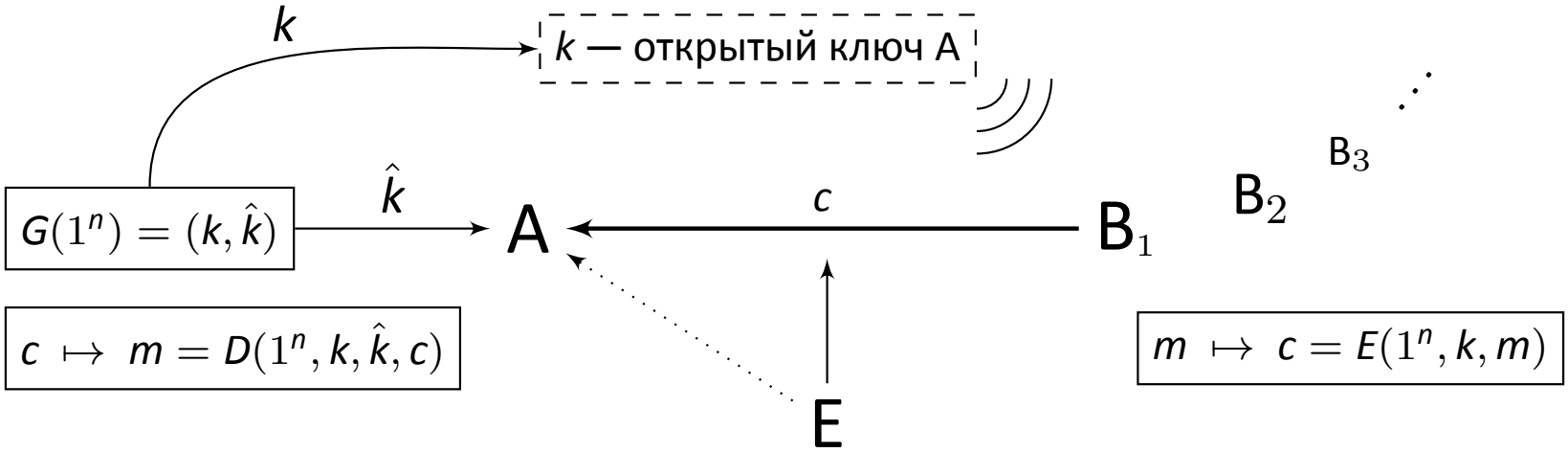
Система (вероятностного) шифрования[a] *с открытым ключом* (public-key (probabilistic) encryption system) — это тройка алгоритмов (G, E, D):

- генератор ключей G — п. в. м. Т., $G(1^n) = (k, \hat{k})$ — *открытый и секретный ключи* (public key and private key, *ключ шифрования и ключ дешифрования*), (можно считать, что пара $(k, \hat{k})$ выбирается из $K_n = \operatorname{supp} G(1^n)$ согласно вероятностному распределению, задаваемому случайной величиной $G(1^n)$);
- алгоритм шифрования E — п. (в.) м. Т.,
  для $m \in M_n$, $(k, \hat{k}) \in K_n$ $\quad E(1^n, k, m) = c$ — *криптограмма (шифртекст)* открытого текста m на ключе k;
- алгоритм дешифрования D — п. д. м. Т.,
  $\forall n \ \forall m \in M_n \ \forall (k, \hat{k}) \in K_n \quad D\big(1^n, k, \hat{k}, E(1^n, k, m)\big) = m$ (с вероятностью 1).

---

[a] Синонимы: *криптосистема*, *шифр* (cryptosystem, cipher).

Схема применения и пример



Криптосистема Эль-Гамаля

g — порождающий элемент (циклической) группы $\langle g \rangle_q$ — открытая информация

- G: $\hat{k} \in_{\mathcal{R}} \{1, \ldots, q-1\}$ — секретный ключ, x
  $k = g^{\hat{k}}$ — соответствующий открытый ключ; g^x
- E: $y \in_{\mathcal{R}} \{1, \ldots, q-1\}$, y
  $s = k^y$, g^{xy}
  $\left.\begin{array}{l} c_1 = g^y \\ c_2 = m \cdot s \end{array}\right\}$ $c = (c_1, c_2)$ — шифртекст для m; $\begin{array}{l} g^y \\ m \cdot g^{xy} \end{array}$
- D: $s = c_1^{\hat{k}}$, $(g^y)^x$
  $m = c_2 \cdot s^{-1}$. $m \cdot g^{xy} \cdot g^{-xy}$

задача CDH:

$(g, g^x, g^y) \overset{?}{\rightsquigarrow} g^{xy}$

Модель противника: атаки

Стойкость криптосистемы определяется относительно конкретного противника.

Ему известна криптосистема, то есть алгоритмы G, E, D, и параметр стойкости n.

Параметр стойкости n и ключ k во время работы противника не меняются.

Основные типы атак:

1. *атака с известным открытым ключом* (key-only attack, **KOA**):
   противнику известен k и вся остальная открытая информация — стандартное предположение

   сюда входят
   - атака с известными шифртекстами (KCA),
     $c_1, c_2, \ldots, c_l$, где $c_i = E(1^n, k, m_i)$ для некоторых $m_i \in M_n$;
   - атака с известными открытыми текстами (KPA),
     $(m_1, c_1), (m_2, c_2), \ldots, (m_l, c_l)$, где $c_i = E(1^n, k, m_i)$;
   - атака с выбором открытых текстов (CPA):
     $m_1, \ldots, m_l \rightsquigarrow (m_1, c_1), \ldots, (m_l, c_l)$, где $c_i = E(1^n, k, m_i)$;

2. *атака с выбором шифртекстов* (**CCA**):
   $c_1, \ldots, c_l \rightsquigarrow (m_1, c_1), \ldots, (m_l, c_l)$, где $m_i = D(1^n, k, \hat{k}, c_i)$.

$1 \leqslant i \leqslant l, \quad l \leqslant \mathrm{poly}(n)$

Атака 2 может быть адаптивной или неадаптивной.

Модель противника: угрозы

Основные типы угроз:

1. полное раскрытие (total breaking):
   найти некоторый $\hat{k}'$, соответствующий k, то есть такой, что $(k, \hat{k}') \in K_n$;
2. извлечение открытого текста:
   по известной информации и случайному значению $E(1^n, k, m)$ найти m;
3. извлечение частичной информации об открытом тексте:
   для некоторой функции $f : \mathbb{B}^* \to \mathbb{B}^*$ по известной информации и случайному значению $E(1^n, k, m)$ найти $f(m)$;
4. различение двух шифртекстов:
   выбрав $m^0, m^1 \in M_n$, $|m^0| = |m^1|$, и получив $E(1^n, k, m^b)$ для $b \in_u \{0, 1\}$, определить b, то есть какое из двух сообщений было зашифровано.

Криптосистема *IND-стойкая*, если она удовлетворяет условию неразличимости двух шифртекстов (indistinguishability of encryptions), то есть никакой противник не может осуществить угрозу 4 с вероятностью, существенно превосходящей вероятность случайного угадывания.

! Если E — детерминированный алгоритм, то (G, E, D) не может быть IND-стойкой.

? Объясните, почему это так.

Определение стойкости

Определение

Криптосистема (G, E, D) называется *стойкой против угрозы извлечения открытого текста на основе атаки с известным открытым ключом*, если
для любой п. в. м. Т. A

$$\Pr_{m \in_{\mathcal{U}} M_n}\left[A\left(1^n, k, E(1^n, k, m)\right) = m\right] = \mathrm{negl}(n).$$

NB Вероятность здесь определяется, помимо случайного выбора m, также вероятностными машинами G (генерирующей k), A и E.

Как формализация требований, которые предъявляются к преобразованиям, выполняемым в стойких (в смысле этого определения) криптосистемах, возникло понятие *семейства функций с секретом*.

Это одностороннее семейство функций, которое тем не менее допускает инвертирование своих функций эффективным алгоритмом, но только если он «знает» некий секрет (о функции).

Семейства функций с секретом

Для произвольных полиномов $l(\cdot)$, $m(\cdot)$ рассмотрим семейство функций вида

$$F = \bigcup_{n \in \mathbb{N}} F_n, \quad \text{где } F_n = \{f_{n,d} : \mathbb{B}^{l(n)} \to \mathbb{B}^{m(n)}\}_{d \in I_n}, \quad I_n \subseteq \mathbb{B}^*.$$

Определение

F называется *семейством функций с секретом* (trapdoor function family), если для него существует тройка[a] алгоритмов (Γ, C, R), удовлетворяющих следующим условиям:

1. Γ — п. в. м. т.: $\Gamma(1^n) = \big(\Gamma_1(1^n), \Gamma_2(1^n)\big) = (d, s)$,
   d — описание (индекс) функции из семейства, s — «секрет» для этой функции;
2. C — п. д. м. т.: $\big(\Gamma_1, C\big)$ — генератор для семейства F: $C(1^n, d, x) = f_{n,d}(x)$;
3. R — п. д. м. т.:
   $\forall\,(d,s) \in \operatorname{supp}\Gamma(1^n)\ \ \forall\, x \in \mathbb{B}^{l(n)} \quad R\big(1^n, d, s, f_{n,d}(x)\big) = x' \in f_{n,d}^{-1}\big(f_{n,d}(x)\big)$;
4. для любой п. в. м. т. A $\Pr\limits_{\substack{x \in_{\mathcal{U}} \mathbb{B}^{l(n)} \\ d = \Gamma_1(1^n)}} \big[A(1^n, d, f_{n,d}(x)) \in f_{n,d}^{-1}\big(f_{n,d}(x)\big)\big] = \operatorname{negl}(n)$.

---

[a] Называемая *генератором функций с секретом*.

1 + 2 $\Rightarrow$ семейство F полиномиально вычислимо.

1 + 2 + 4 $\Rightarrow$ F одностороннее.

Криптосистемы с открытым ключом из перестановок с секретом

Утверждение

Если существует семейство перестановок с секретом, то существует система (детерминированного) шифрования с открытым ключом, стойкая против угрозы извлечения открытого текста на основе атаки с известным открытым ключом.

Пусть $F = \{f_{n,d} : \mathbb{B}^n \to \mathbb{B}^n\}_{n,d}$ — семейство перестановок с секретом.

n — параметр стойкости, $M_n = \mathbb{B}^n$, $m \in M_n$.

- $G \equiv \Gamma$: $G(1^n) = (d, s)$,
  $k = d$ — открытый ключ, $\hat{k} = s$ — секретный ключ;
- $E \equiv C$: $E(1^n, d, m) = C(1^n, d, m) = f_{n,d}(m) = c$ — криптограмма;
- $D \equiv R$: $D(1^n, d, s, c) = R\big(1^n, d, s, c\big) = f_{n,d}^{-1}(c) = m$.

(G, E, D) — криптосистема с открытым ключом, и она стойкая по определению семейства перестановок с секретом.

NB Необходимым условием существования стойких (в указанном смысле) криптосистем с открытым ключом является существование односторонних функций: функция f, такая, что $f(r) = k \Leftrightarrow G(r; 1^n) = (k, \hat{k})$, должна быть односторонней, иначе противник с существенной вероятностью по k сможет находить r и вычислять $\hat{k}$.

Пример семейства функций с секретом

Семейство функций Рабúна

$\mathbb{P}_n$ — множество простых чисел p длины n в двоичной записи, таких, что $p \equiv 3 \mod 4$.

$N = pq$, где $p \neq q$, $p, q \in \mathbb{P}_n$, называется *числом Блюма* в таком случае.

- Γ: $\Gamma(1^n) = \big(N, (p, q)\big)$ для различных $p, q \in_{\mathcal{R}} \mathbb{P}_n$.
- C: $f_{n,N}(x) = x^2 \mod N$ по всем $x \in \mathbb{Z}_N^*$.

Теорема

Если для любой п. в. м. Т. A при $p, q \in_{\mathcal{R}} \mathbb{P}_n$ $\Pr\big[A(1^n, pq) \in \{p, q\}\big] = \mathrm{negl}(n)$,[a]
то $\{f_{n,N}\}$ — семейство функций с секретом.

[a] То есть если задача факторизации чисел Блюма трудна в среднем.

- R: $R\big(1^n, N, (p, q), z\big) = y$ при $z = x^2 \mod N$ для некоторого $x \in \mathbb{Z}_N^*$, где

$$\begin{cases} y \equiv \pm z^{\frac{p+1}{4}} \mod p \\ y \equiv \pm z^{\frac{q+1}{4}} \mod q \end{cases}$$ — четыре варианта, каждый из которых даёт единственное значение y (по китайской теореме об остатках),

$$\left.\begin{aligned} y^2 \equiv z^{\frac{p+1}{2}} \equiv x^{p+1} \equiv x^{p-1}x^2 \equiv x^2 \equiv z \mod p \\ y^2 \equiv z^{\frac{q+1}{2}} \equiv x^{q+1} \equiv x^{q-1}x^2 \equiv x^2 \equiv z \mod q \end{aligned}\right\} \Rightarrow y^2 \equiv z \mod N.$$

NB Если особым образом ограничить области определения функций $f_{n,N}$ (брать только элементы нечётного порядка из $\mathbb{Z}_N^*$), то получится семейство *перестановок* с секретом.

Доказательство теоремы

Достаточно проверить выполнение условия 4 определения семейства функций с секретом. Предположим, что A — полиномиальная вероятностная машина Тьюринга и пусть

$$\varepsilon(N) = Pr\{f_{n,N}(A(N, f_{n,N}(z))) = f_{n,N}(z)\}$$

ее вероятность успеха инвертирования $f_{n,N}$. Здесь $z \in_R \mathbb{Z}_N^*$. Построим машину Тьюринга S^A с оракулом A для факторизации N. На входе N машина S выполняет следующие шаги:

1. выбирает $z \in_R \mathbb{Z}_N^*$ и вычисляет $f_{n,N}$:
2. вычисляет $a = \text{Н.О.Д.}(N, f_{n,N}(z))$. Если $a \neq 1$, то выдает $(N/a, a)$ и останавливается:
3. обращается к машине A как к оракулу, передавая ей в качестве запроса пару $(N, f_{n,N}(z))$. Пусть $z' = A(N, f_{n,N}(z))$:
4. вычисляет $b = \text{Н.О.Д.}(N, z + z')$. Выдает $(N/b, b)$ и останавливается.

Доказательство теоремы

Если на шаге 2 $a \neq 1$, то получаем разложение $N = pq$. При условии $a = 1$ случайная величина z равномерно распределена в $\mathbb{Z}_N^*$. В этом случае вероятность успеха инвертирования $f_{n,N}(z)$ по определению машины A равна $\varepsilon(N)$. Заметим, что существует ровно четыре возможных допустимых значений для z'. Без ограничения общности можно считать, что z' принадлежит множеству из двух корней $\{z_1, N - z_1\}$. Поскольку значение z распределено равномерно на множестве всех четырех прообразов $f_{n,N}(z)$, то с вероятностью $1/2$ элемент z принадлежит другой паре $(z_2, N - z_2)$. Тогда $(z + z')(z - z') = z^2 - (z')^2 = 0 \,(\mathrm{mod} N)$ и $z + z' \neq 0 \,(\mathrm{mod} N)$. Тогда $\text{Н.О.Д.}(z + z', N) \in \{p, q\}$.
Следовательно, машина S^A находит разложение числа N на множители p и q с вероятностью не менее $\varepsilon(N)/2$.

IND-стойкость систем вероятностного шифрования с открытым ключом

До сих пор рассматривалось детерминированное шифрование.

Пусть теперь алгоритм E — вероятностный.

Оракул $\mathcal{O}$, определяемый для криптосистемы (G, E, D),

- в начале работы выполняет алгоритм G и получает ключи $(k, \hat{k}) = G(1^n)$;
- после этого принимает запрос (m^0, m^1), где $m^0, m^1 \in M_n$, получив который, выбирает случайный бит $b \in_u \{0, 1\}$ и возвращает $E(1^n, k, m^b)$, то есть в зависимости от значения b выдаёт либо $E(1^n, k, m^0)$, либо $E(1^n, k, m^1)$;
- ответив на запрос, завершает свою работу.

Определение

Система (G, E, D) вероятностного шифрования с открытым ключом называется *полиномиально стойкой* (IND-KOA-стойкой), если для любой п. в. м. Т. A с вышеописанным оракулом $\mathcal{O}$

$$\Pr\left[A^{\mathcal{O}}(1^n, k) = b\right] \leqslant \frac{1}{2} + \mathrm{negl}(n).$$

NB Вероятность определяется выбором b, а также машинами G, A, E.

Гипотетический пример полиномиально стойкой криптосистемы

$y \in \mathbb{Z}_M^*$ называется *квадратичным вычетом* (quadratic residue) по модулю M, если $y \equiv a^2 \mod M$ для некоторого a.

$p, q \in \mathbb{P}$
$N = pq$

Символ Лежандра для $y \in \mathbb{Z}_p^*$ $\quad \left(\frac{y}{p}\right) = \begin{cases} 1, & \text{если } y \text{ — квадратичный вычет } \mod p \\ -1, & \text{если } y \text{ — квадратичный невычет } \mod p \end{cases}$

Символ Якоби для $y \in \mathbb{Z}_N^*$ $\quad \left(\frac{y}{N}\right) = \left(\frac{y}{p}\right) \cdot \left(\frac{y}{q}\right)$

$Y^1 = \left\{y \in \mathbb{Z}_N^* : \left(\frac{y}{N}\right) = 1\right\}$

В этом множестве половина — квадратичные вычеты и по p, и по q,
половина — квадратичные невычеты по p и по q одновременно.

NB Принадлежность $y \overset{?}{\in} Y^1$ эффективно распознаваема: существует полиномиальный алгоритм вычисления символа Якоби (даже при неизвестном разложении N на простые множители).

Определим предикат на Y^1: $\quad QR_N(y) = \begin{cases} 1, & \text{если } y \text{ — квадратичный вычет } \mod N, \\ 0, & \text{если } y \text{ — квадратичный невычет } \mod N. \end{cases}$

Гипотетический пример полиномиально стойкой криптосистемы

Задача вычисления $QR_N(y)$ для Y^1 при неизвестных p и q считается трудной:

Предположение QR

Для любой п. в. м. Т. A $\Pr\limits_{\substack{p,q \\ y \in_{\mathcal{R}} Y^1}} \left[A(N, y) = QR_N(y)\right] \leqslant \frac{1}{2} + \mathrm{negl}(n)$.

Криптосистема Гольдвассер—Микали

- $G(1^n) = \big((N, z), (p, q)\big)$ для различных $p, q \in_{\mathcal{R}} \mathbb{P}$, $N = pq$ и произвольного квадратичного невычета z из Y^1.
- E применяется отдельно к каждому биту открытого текста $\Rightarrow$ считаем $m \in \mathbb{B}$,
  $c = E\big((N, z), m\big) = z^m x^2 \mod N$, где $x \in_{\mathcal{R}} \mathbb{Z}_N^*$.
  Значит, $E\big((N, z), 0\big)$ — случайный квадратичный вычет по модулю N,
  $E\big((N, z), 1\big)$ — случайный квадратичный невычет по модулю N.
- $D\big((p, q), c\big) = 1 - QR_N(c)$ — эффективно вычислимо при известных p и q
  (достаточно проверить $c^{\frac{p-1}{2}} \stackrel{?}{\equiv} 1 \mod p$ и $c^{\frac{q-1}{2}} \stackrel{?}{\equiv} 1 \mod q$).

Теорема

Если предположение QR справедливо,
то криптосистема Гольдвассер—Микали полиномиально стойкая.