Прикладная Криптография: Симметричные криптосистемы

Макаров Артём

МИФИ 2018

Структура курса

- Лекции: 16 недель

- Сдача разделов: 4 блока
  - Для каждого блока жёсткий дедлайн (без переносов)
  - https://github.com/CryptoCourse/CryptoLabs/wiki/список-лабораторных-работ

- Для сдачи каждого блока:
  - Сдача лабораторных работ для данного блока
  - Сдача лабораторной работы + теория
  - Сдача домашней работы + теория
  - Сдача теории по лекциям

Лабораторные работы

- Образ Linux машины с развёрнутой REST API службой.
- Задача – продемонстрировать атаку на криптосистему систему с уязвимостью.
- Допустимые языки программирования: C++, C#, Python, Java, другие?
- Подробнее на лабораторной работе.

Сдача теории

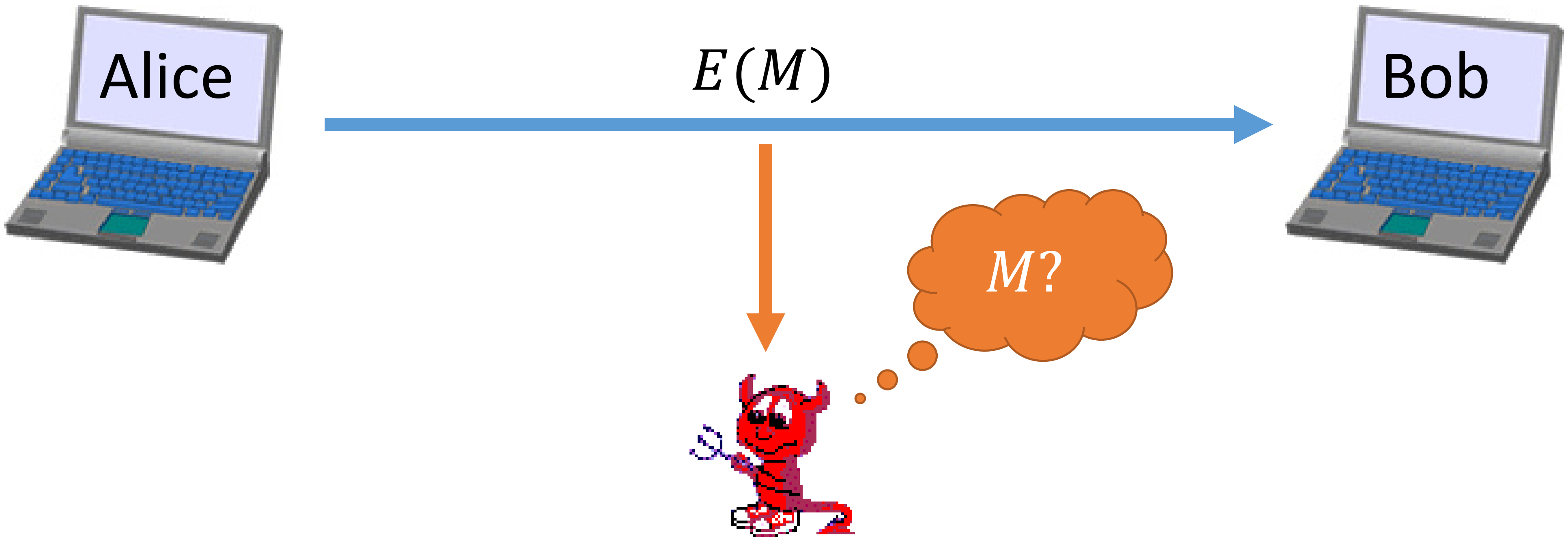
- Сдаётся в формате вопрос – ответ
  - Задаётся набор различных вопросов по пройденному материалу
  - Если на какой то вопрос ответ не получен, или получен не верный ответ – даётся время подумать или поискать ответ
  - Количество попыток – не ограничено внутри блока

- Несправедливости:
  - Разное количество вопросов разным людям
  - Максимальное количество вопросов – не ограничено
  - Возможность не сдать теорию, даже если в гугле были найдены все ответы

Обратная связь и пожелания по курсу

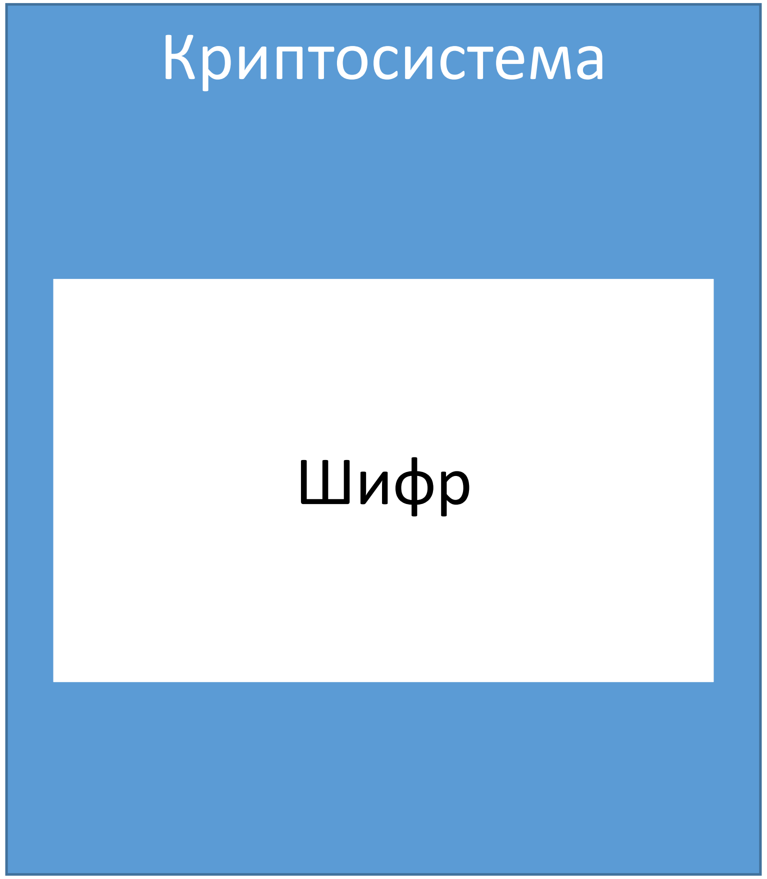
Историческая задача криптографической защиты информации

- Передача зашифрованного сообщения по открытому каналу
- При перехвате зашифрованного сообщения открытый текст должен остаться неизвестным для злоумышленника



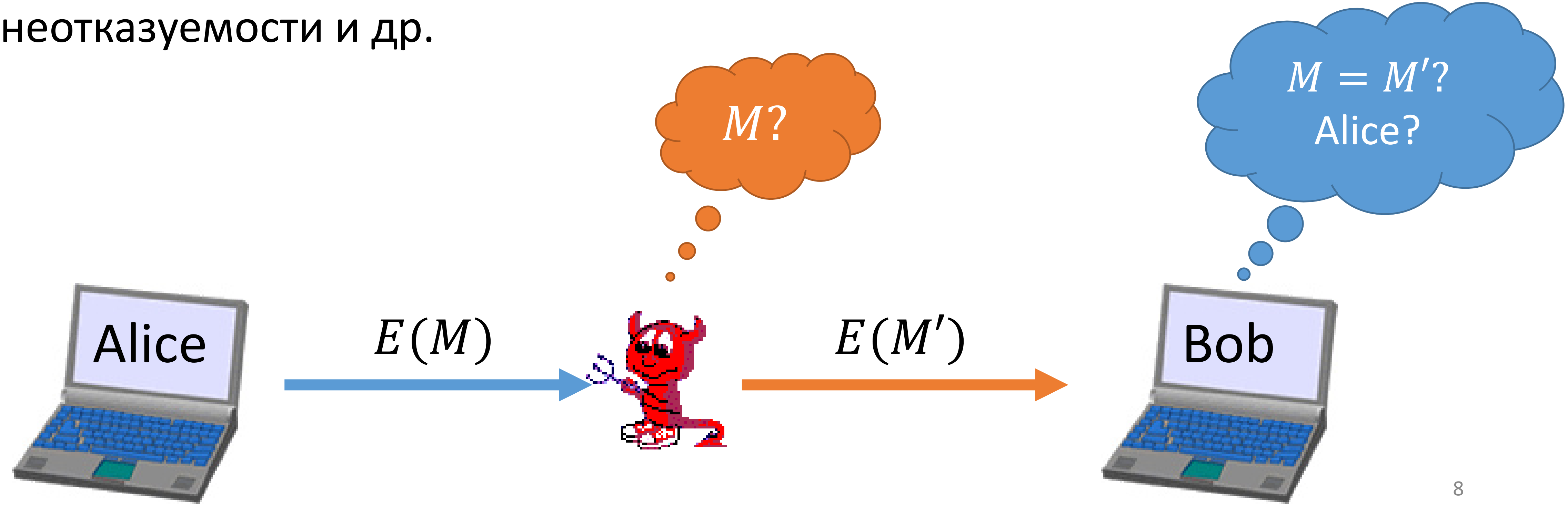
Способы построения и анализа криптосистем

- **Досистемный подход**– построение и анализ криптосистем, которые выглядят «сложными» для создателя;
- Предположении о стойкости исходит «из очевидной сложности взлома» для создателя схемы
- Примеры – шифр Цезаря, шифр простой замены, шифр Вижинера



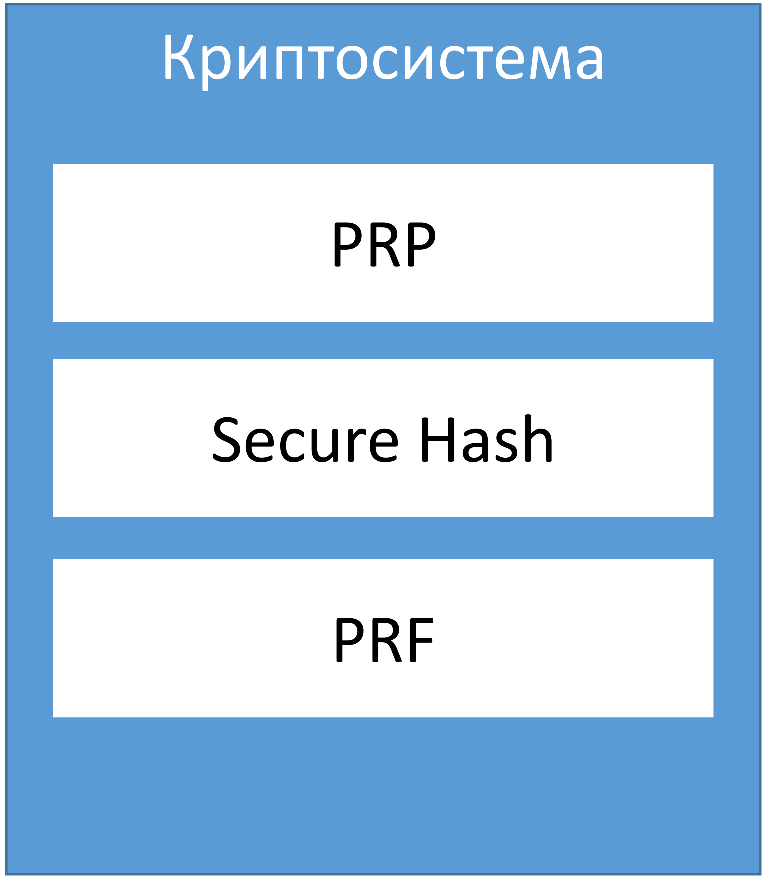
Современная задача криптографической защиты информации

- Передача сообщения по открытому каналу
- Возможен активный злоумышленник
- Обеспечение конфиденциальности, аутентичности, целостности, неотказуемости и др.



Способы построения и анализа криптосистем

- **Системный подход**– построение и анализ криптосистем на основе криптографических примитивов
- Возможно наличие не только средств обеспечения секретности, но и аутентичности, целостности и других
- Предположении о стойкости исходит из анализа системы в целом, через сведение стойкости в сложности вычислительно сложной задачи
- При замене части системы необходимо произвести анализ заново

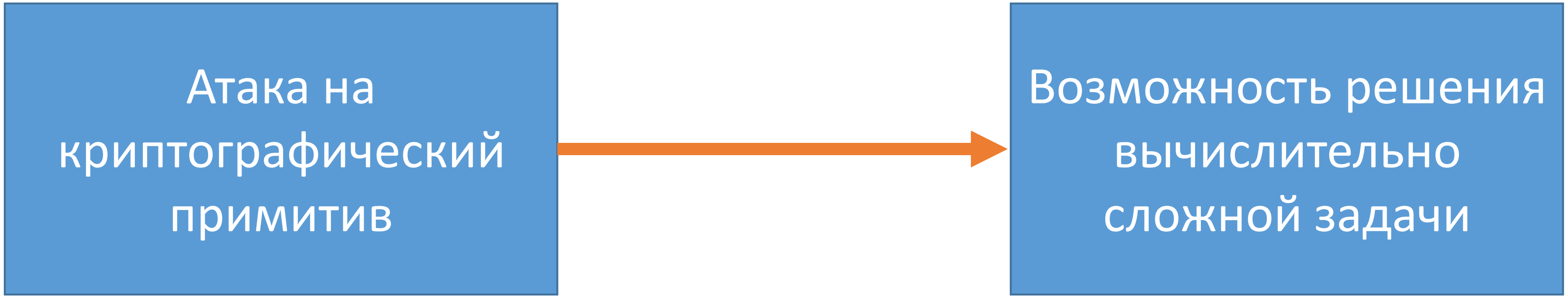


Способы построения и анализа криптосистем

- **Современный подход**– построение и анализ криптосистем на основе абстрактных моделей криптографических примитивов
- Вместо анализа частных свойств примитивов и их взаимодействия производится анализ самой контракции, вне зависимости от используемых примитивов и их стойкости
- Предположении о стойкости исходит из анализа системы в предположении об априорной стойкости примитивов
- При замене части системы нет необходимости проводить повторных анализ

Сведение стойкости (Security Reduction)

- Наиболее распространённый способ доказательства стойкости криптографического примитива является сведение атаки на него в вычислительно сложной задаче. Иными словами показывается, что произвести атаку на примитив так же сложно как вышить вычислительно сложную задачу.



Сведение стойкости криптографический примитивов

- Для симметричных криптосистем стойкость сводится к задаче 3SAT:
    - Пусть дана булевая функция от N переменных
    - Найти вектор решений, при котором значение булевой функции равно 1.
    - NP полная задача

- Для асимметричных криптосистем стойкость может сводится:
    - Задача дискретного логарифмирования в конечных группах
    - Задача факторизации больших целых чисел
    - Задача нахождения кратчайшего вектора решётки
    - Задача декодирования линейных кодов
    - Задача решения многомерных квадратичных многочленов

Шифр Шеннона

Шифр Шеннона - пара функций $\mathrm{E} = (E, D)$, таких что:

- (1) Функция E (**функция зашифрования**) принимает на вход ключ k и сообщение m (называемой открытым текстом, PT) и даёт на выходе шифртекст c (CT), такой что

$$c = E(k, m).$$

Говорят, что c есть **зашифрование** m на ключе k.

- (2) Функция D (**функция расшифрования**) принимает на вход ключ k и шифртекст c и даёт на выходе сообщение m, такое что

$$m = D(k, c)$$

Говорят, что m это **расшифрование** c на ключе k.

Шифр Шеннона

- (3) Функция D обращает функцию E (**свойство корректности**):

$$\forall k, \forall m \; D\big(k, E(k, m)\big) = m.$$

Пусть K – **множество ключей**, M **– множество сообщений**, C **– множество шифртекстов**.

Тогда шифром Шеннона, определённым над (K, M, C) называют пару функций $\mathrm{E} = (E, D)$:

$$E: K \times M \to C,$$
$$D: K \times C \to M,$$

для которых выполняются свойства (1) – (3).

Нотация

$v \in V_n = \{0,1\}^n$ - двоичный вектор длины n

0^n- двоичный вектор $(000 \dots 00) \in V_n$

1^n - двоичный вектор $(111 \dots 11) \in V_n$

$0^k 1^l$- двоичный вектор $(\underbrace{000 \dots 00}_{k}\underbrace{111 \dots 11}_{l}) \in V_{k+l}$

$v' \in \{0,1\}^* = \bigcup_{k=0}^{\infty} \{0,1\}^k$ - двоичный вектор произвольной длины

$v'' \in \{0,1\}^{\leq L} = \bigcup_{k=0}^{L} \{0,1\}^k$ - двоичный вектор, длины не больше L

Нотация

$v \in V_n = \{0,1\}^n$ - двоичный вектор длины n

Пусть $a \in V_n$: $a = (a_0, a_1, \dots, a_{n-1})$, b$\in V_n$: $b = (b_0, b_1, \dots, b_{n-1})$

$ab = (a||b) \in V_{2n}$: $(a||b) = (a_0, a_1, \dots, a_{n-1}, b_0, b_1, \dots, b_{n-1})$ – конкатенация векторов a и b

$v[a]$ - a-я координата вектора v, $a < n$

$v[a, a+1, \dots b] \in V_{b-a+1}$- подвектор, полученный из координат вектора v, $a < b < n$.

Пример: Одноразовый блокнот

Пусть $\mathrm{E} = (E, D)$ – **шифр Шеннона**, для которого $K = M = C = \{0,1\}^L$, где L – фиксированный параметр.

Для ключа $k \in K$ и сообщения $m \in M$ функция **зашифрования** определена как:

$$E(k, m) = k \oplus m.$$

Для ключа $k \in K$ и шифртекста $c \in C$ функция **расшифрования** определена как:

$$D(k, c) = k \oplus c.$$

$\oplus$ - побитное сложение по модулю 2 (XOR).

Корректноть: $D\big(k, E(k, m)\big) = D(k, k \oplus m) = k \oplus (k \oplus m) = (k \oplus k) \oplus m = 0^L \oplus m = m$.

Пример: Одноразовый блокнот переменной длины

Пусть $\mathrm{E} = (E, D)$ – **шифр Шеннона**, для которого $K = \{0,1\}^L, M = C = \{0,1\}^{\leq L}$, где L – фиксированный параметр.

Для ключа $k \in K$ и сообщения $m \in M$: $|m| = l$ функция **зашифрования** определена как:

$$E(k, m) = k[0..l-1] \oplus m.$$

Для ключа $k \in K$ и шифртекста $c \in C$: $|c| = l$ функция **расшифрования** определена как:

$$D(k, c) = k[0..l-1] \oplus c.$$

$\oplus$ - побитное сложение по модулю 2 (XOR).

Корректноть: $D\big(k, E(k, m)\big) = D(k, k \oplus m) = k \oplus (k \oplus m) = (k \oplus k) \oplus m = 0^L \oplus m = m$.

Пример: Шифр подстановки

Пусть Σ – конечный алфавит. Пусть $\mathrm{E} = (E, D)$ – **шифр Шеннона**. для которого $M = C = \Sigma^L$, где L – фиксированный параметр. K – множество $S(\Sigma)$ всех подстановок над Σ.

Для ключа $k \in K$ и сообщения $m \in M$: $|m| = l$ функция **зашифрования** определена как:

$$E(k, m) = (k(m[0]), k(m[1]), \dots, k(m[L-1])).$$

Для ключа $k \in K$ и шифртекста $c \in C$: $|c| = l$ функция **расшифрования** определена как:

$$D(k, c) = (k^{-1}(c[0]), k^{-1}(c[1]), \dots, k^{-1}(c[L-1])).$$

$\bigoplus$ - побитное сложение по модулю 2 (XOR).

Корректноть: $D\big(k, E(k, m)\big) = D(k, k \bigoplus m) = (k^{-1}(k(m[0])), \dots, k^{-1}(k(m[L-1])) = (m[0], \dots, m[L-1]) = m$

Пример: Аддитивный одноразовый блокнот

Пусть $\mathrm{E} = (E, D)$ – **шифр Шеннона**, для которого $K = M = C = \{0, \dots, n-1\}$, где n – фиксированный параметр.

Для ключа $k \in K$ и сообщения $m \in M$ функция **зашифрования** определена как:

$$E(k, m) = m + k \bmod n$$

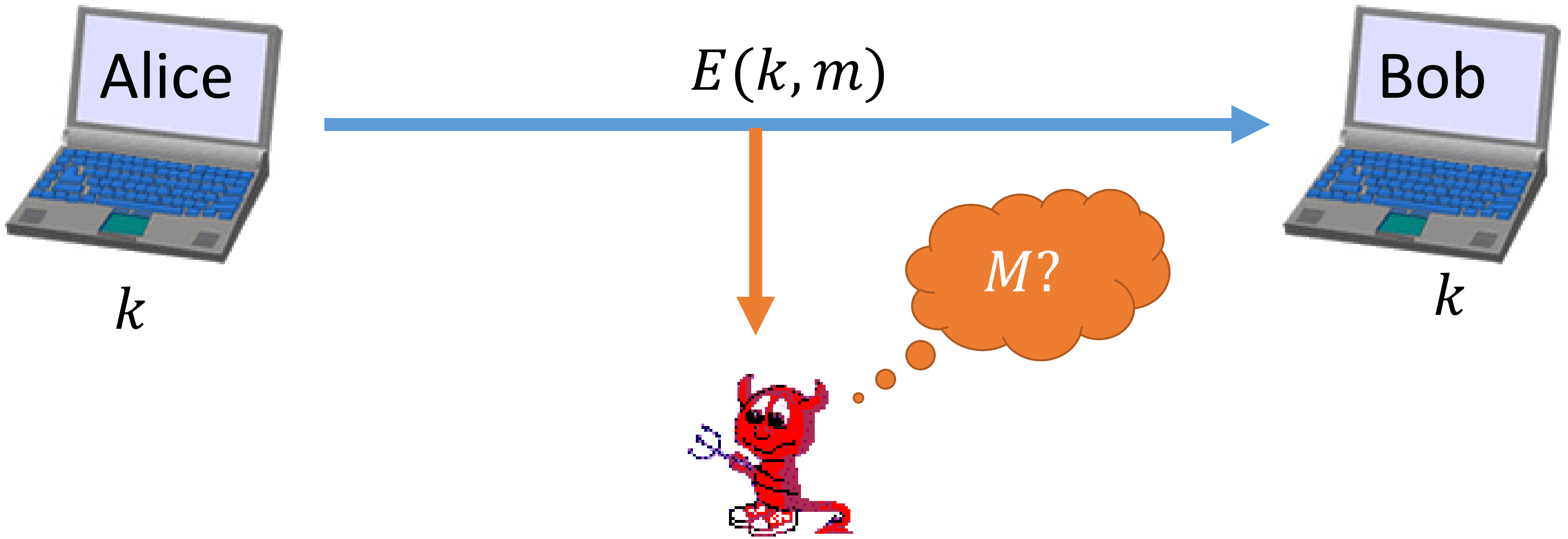
Для ключа $k \in K$ и шифртекста $c \in C$ функция **расшифрования** определена как:

$$D(k, c) = c - k \bmod n$$

Корректноть: $D\big(k, E(k, m)\big) = D(k, m + k) = (m + k) - k = m$.

Цель шифра Шеннона

- Цель шифра Шеннона – обеспечение **секретности** передаваемых сообщений по открытому каналу
- Для обеспечения секретности необходим общий секретный ключ $k \in K$, неизвестный для злоумышленника



Понятие стойкости

Очевидный вопрос – что понимать под стойкостью шифра?

Стойкость – метрика качества шифра.

- Попытка 1: размер ключа
  - Чем больше ключ, тем сложнее перебрать все возможные варианты. Длина ключа как параметр стойкости.
  - Но возможны и другие атаки, кроме перебора, например частотный анализ
  - Пример – шифр подстановки, $|\Sigma| = 27, K = S(\Sigma): |K| \sim 10^{28}$, но возможна полиномиальная частотная атака

Понятие стойкости

- Попытка 2: малая вероятность расшифрования
  - Чем меньше вероятность расшифрования для злоумышленника, тем более стойкий шифр. Вероятность расшифрования как параметр стойкости.
  - Но тогда шифр определённый на коротких сообщениях, например 1 бит, менее стойкий чем шифр, определённый на длинных сообщениях, так как велика возможность «угадать» сообщение.
  - Иными словами, невозможно обеспечить стойкость при шфировании однобитного сообщения

Понятие стойкости

- Попытка 3: **равная** вероятность расшифрования
  - При данном шифртексте вероятность расшифрованы его в любой открытый текст **одинакова**
  - Пример нестойкого шифра: $M = \{0,1\}^n$, $\mathrm{E} = (E, D)$ – шифр Шеннона над (K, M, C):

$$\mathrm{K}_0 \subset K: E(k_0, \mathrm{m}_0) = \mathrm{c}\,,$$
$$\mathrm{K}_1 \subset K: E(k_1, \mathrm{m}_1) = c,$$
$$|K_0| > |K_1|$$
$$m_0, m_1 \in M: m_0 \neq m_1$$

Вероятность угадывания при выборе m_0 ($|K_0| = 800, |K_1| = 600$):

$$\frac{|K_0|}{|K_0| + |K_1|} \approx 57\% > 50\%$$

Абсолютная стойкость

Определение 1.1. Пусть $\mathrm{E} = (E, D)$ – шифр шеннона над (K, M, C). Рассмотрим вероятностный эксперимент, в котором случайная величина $\boldsymbol{k}$ равномерна распределена на K $(\boldsymbol{k} \in_R K)$.

Если $\forall m_0, m_2 \in M$ и $c \in C$ имеем:

$$\Pr[E(\boldsymbol{k}, m_0) = c] = \Pr[E(\boldsymbol{k}, m_1) = c]$$

То шифр E называется **абсолютно стойким шифром Шеннона**.

Абсолютная стойкость защищает против **любых** (не только эффективных) противников.

Эквивалентные определения абсолютной стойкости

Теорема 1.1. Пусть $\mathrm{E} = (E, D)$ **-** шифр Шеннона над (K, M, C). Тогда следующие определения эквивалентны:

- (1) E – абсолютно стойкий
- (2) $\forall c \in C\ \exists N_c(c): \forall m \in M\ |\{k \in K: E(k, m) = c\}| = N_c$
- (3) Если $\boldsymbol{k} \in_R K$ тогда все случайные величины $E(\boldsymbol{k}, m)$ имеют одинаковое распределение

$\rhd$ $(2) \Rightarrow (3)$ Переформулируем (2): для каждого $c \in C$ существует число $P_c(c)$, такое что $\forall m \in M\ \Pr[E(\boldsymbol{k}, m) = c] = P_c$, $\boldsymbol{k} \in_R K$. $P_c = \frac{N_c}{|K|}$. $\lhd$

Эквивалентные определения абсолютной стойкости

Теорема 1.1. Пусть $\mathrm{E} = (E, D)$ **-** шифр Шеннона над (K, M, C). Тогда следующие определения эквивалентны:

- (1) E – абсолютно стойкий
- (2) $\forall c \in C\ \exists N_c(c){:}\ \forall m \in M\ |\{k \in K{:}\ E(k, m) = c\}| = N_c$
- (3) Если $\boldsymbol{k} \in_R K$ тогда все случайные величины $E(\boldsymbol{k}, m)$ имеют одинаковое распределение

$\rhd$ $(1) \Rightarrow (2)$ Пусть $c \in C$ фиксированный шифртекст. Выберем произвольное сообщение $m_0 \in M$. Пусть $P_c = \Pr[E(\boldsymbol{k}, m_0) = c]$. $(1) \Rightarrow$ $\forall m \in M \Pr[E(\boldsymbol{k}, m) = c] = \Pr[E(\boldsymbol{k}, m_0) = c] = Pc$. $\lhd$

Эквивалентные определения абсолютной стойкости

Теорема 1.1. Пусть $\mathrm{E} = (E, D)$ **-** шифр Шеннона над (K, M, C). Тогда следующие определения эквивалентны:

- (1) E – абсолютно стойкий
- (2) $\forall c \in C\ \exists N_c(c) \colon \forall m \in M\ |\{k \in K \colon E(k, m) = c\}| = N_c$
- (3) Если $\boldsymbol{k} \in_R K$ тогда все случайные величины $E(\boldsymbol{k}, m)$ имеют одинаковое распределение

$\rhd$ $(2) \Rightarrow (1)$. Фиксируем $m_0, m_1 \in M, c \in C$ (2) $\Rightarrow$ $\Pr[E(\boldsymbol{k}, m_0) = c] = P_c = \Pr[E(\boldsymbol{k}, m_1) = c]$. $\lhd$

Одноразовый блокнот – абсолютно стойкий шифр

Теорема 1.2. Пусть $\mathrm{E} = (E, D)$ - одноразовый блокнот при $K = M = C = \{0,1\}^L$ для параметра L. Тогда E – абсолютно стойкий шифр.

$\rhd$ Для фиксированного сообщения $m \in M$, шифртекста $c \in C$ и ключа $k \in K$, уникального для сообщения m : $k = m \oplus c$ имеем определение (2) из **Теоремы 1.1** $\lhd$

Одноразовый блокнот переменной длины – не абсолютно стойкий шифр

Теорема 1.3. Пусть $\mathrm{E} = (E, D)$ - одноразовый блокнот переменной длины при $K = \{0,1\}^L, M = C = \{0,1\}^{\leq L}$ для параметра L. Тогда E – **не** абсолютно стойкий шифр.

⊳ Пусть $m_0 \in M\colon |m_0| = 1, m_1 \in M\colon |m_1| > 1, c \in C\colon |c| = 1$

$$a = \Pr[E(k, m_0) = c] = 0.5$$
$$\mathrm{b} = \Pr[E(k, m_1) = c] = 0$$
$$a \neq b.$$

Иными словами не выполняется **Определение 1.1**. (Абсолютная стойкость). ⊲

Эквивалентные определения абсолютной стойкости

Теорема 1.4. Пусть $\mathrm{E} = (E, D)$ - шифр Шеннона на (K, M, C). Рассмотрим вероятностный эксперимент для равномерно распределённой $\boldsymbol{k} \in_R K$.

Тогда E – абсолютно стойкий тогда и только тогда, когда для произвольного предиката $\phi: C \to \{0,1\}$ и $\forall m_0, m_1 \in M$

$$\Pr[\phi(E(\boldsymbol{k}, m_0)) = 1] = \Pr\left[\phi\left(E(\boldsymbol{k}, m_1)\right) = 1\right]$$

⊳Пусть $S = \{c \in C: \phi(c) = 1\}$. Так как E – абсолютно стойкий имеем

$$\Pr[\phi(E(\boldsymbol{k}, m_0)) = 1] = \sum_{c \in S} \Pr[E(\boldsymbol{k}, m_0) = c] =$$

$$\sum_{c \in S} \Pr[E(\boldsymbol{k}, m_1) = c] = \Pr[\phi(E(\boldsymbol{k}, m_1)) = 1]$$

Эквивалентные определения абсолютной стойкости

Теорема 1.4. Пусть $\mathrm{E} = (E, D)$ - шифр Шеннона на (K, M, C). Рассмотрим вероятностный эксперимент для равномерно распределённой $\boldsymbol{k} \in_R K$.

Тогда E – абсолютно стойкий тогда и только тогда, когда для произвольного предиката $\phi: C \to \{0,1\}$ и $\forall m_0, m_1 \in M$

$$\Pr[\phi(E(\boldsymbol{k}, m_0) = 1] = \Pr\big[\phi\big(E(\boldsymbol{k}, m_1)\big) = 1\big]$$

Пусть E – **не** абсолютно стойкий. То есть

$$\Pr[E(\boldsymbol{k}, m_0) = c] \neq \Pr[E(\boldsymbol{k}, m_1) = c].$$

Фиксируем $c \in C$. Пусть $\phi: \phi(c) = 1, \phi(c') = 0, c' \neq c$.

$$\Pr[\phi(E(\boldsymbol{k}, m_0) = 1] = \Pr[E(\boldsymbol{k}, m_0) = c]$$
$$\neq \Pr[E(\boldsymbol{k}, m_1) = c] = \Pr[\phi(E(\boldsymbol{k}, m_1) = 1] \lhd$$

Эквивалентные определения абсолютной стойкости

Теорема 1.4. Пусть $\mathrm{E} = (E, D)$ - шифр Шеннона на (K, M, C). Рассмотрим вероятностный эксперимент для равномерно распределённой $\boldsymbol{k} \in_R K$.

Тогда E – абсолютно стойкий тогда и только тогда, когда для произвольного предиката $\phi: C \to \{0,1\}$ и $\forall m_0, m_1 \in M$

$$\Pr[\phi(E(\boldsymbol{k}, m_0) = 1] = \Pr\left[\phi\left(E(\boldsymbol{k}, m_1)\right) = 1\right]$$

Иными словами: при использовании произвольного предиката на шифртекстах абсолютно стойкого шифра злоумышленник не получает информации об открытом тексте.

Эквивалентные определения абсолютной стойкости

Теорема 1.5. Пусть $\mathrm{E} = (E, D)$ - шифр Шеннона на (K, M, C). Рассмотрим вероятностный эксперимент для $\boldsymbol{k} \in_R K,\ \boldsymbol{m} \in_R M.\ \boldsymbol{m}$ и $\boldsymbol{k}$ – независимы. Введём случайную величину $\boldsymbol{c} = E(\boldsymbol{k}, \boldsymbol{m})$ Тогда:

- Если E – абсолютно стойкий, тогда $\boldsymbol{c}$ и $\boldsymbol{m}$ независимы:
- Если $\boldsymbol{c}$ и $\boldsymbol{m}$ независимы, и каждое сообщение из M выберется с вероятностью, отличной от 0, то E – абсолютно стойкий.

Иными словами, для абсолютно стойкого шифра верно равенство:

$$\Pr[\boldsymbol{m} = m | \boldsymbol{c} = c] = \Pr[\boldsymbol{m} = m]$$

То есть наличие шифртекста не даёт злоумышленнику никаких преимуществ.

Энтропия

Мера неопределённости в поведении сигнала, количество информации передаваемое сигналом, величина измерения – бит.

$H(\boldsymbol{x}) = -\Pr[\boldsymbol{x}] \log_2 \Pr[\boldsymbol{x}]$ - энтропия случайной величины $\boldsymbol{x}$.

Пусть $\boldsymbol{x} \in_R \{0,1\}^n$, тогда $H(\boldsymbol{x}) \leq n$. $H(\boldsymbol{x}) = n$ если $\boldsymbol{x}$ – равномерно распеределенная

$H(\boldsymbol{x}|\boldsymbol{y}) = \sum_{a \in X} \Pr[\boldsymbol{x} = a] \, H(\boldsymbol{x}|\boldsymbol{y} = a)$ - условная энтропия случайной величины $\boldsymbol{x}$. $H(\boldsymbol{x}|\boldsymbol{y}) \leq H(\boldsymbol{x})$, $H(\boldsymbol{x}|\boldsymbol{y}) = H(\boldsymbol{x})$, если $\boldsymbol{x}$ и $\boldsymbol{y}$ независимы.

Эквивалентные определения

Теорема 1.6. Пусть $\mathrm{E} = (E, D)$ - шифр Шеннона на (K, M, C). Пусть $\boldsymbol{m} \in_R M, \boldsymbol{c} \in_R C$. Тогда шифр E – абсолютно стойкий, если $H(m) = H(m|c)$

Иными словами шифртекст не даёт никакой информации об открытом тексте.

Принцип действия абсолютно стойкого шифра – «применить» энтропию (неопределённость) равномерно распределённого ключа к сообщению для получния равномерно распределённого шифртекста.

Плохие новости

Теорема 1.7 (Шеннона). Пусть $\mathrm{E} = (E, D)$ шифр Шеннона на (K, M, C). Если E – абсолютно стойкий, то

- $|K| \geq |M|$
- $H(\boldsymbol{k}) \geq H(\boldsymbol{m}), \boldsymbol{k} \in_R K, \boldsymbol{m} \in_R M$

Простое объяснение – невозможно получить равномерно распределённую случайную величину длины m, используя детерминированный алгоритм над равномерно распределённой случайной величиной длины $n < m$.

Иными словами для шифрования 1 Gb данных **любым** абсолютно стойким шифром потребуется ключ размера как минимум 1 Gb.

Вычислимый шифр

Вычислимый шифр на (K, M, C) – пара **эффективных** алгоритмов $\mathrm{E} = (E, D)$, где $E: K \times M \to C$ – вероятностная функция защифрования, $D: K \times C \to M$ – функция расшифрования.

- Обозначим процедуры зашифрования как $c \overset{R}{\leftarrow} E(k, m)$.
- Обозначим выбор равномерно распределённого ключа как $k \overset{R}{\leftarrow} K$.

При этом $\forall k \in K, m \in M, c \overset{R}{\leftarrow} E(\mathrm{k}, \mathrm{m}), m' \leftarrow D(k, c) \Pr[m = m'] = 1$ (**свойство корректности**).

Семантическая стойкость

Пусть $\mathrm{E} = (E, D)$ - вычислимый шифр на (K, M, C).

Теорема 1.3 $\Rightarrow E$ – абсолютно стойкий, если $\forall \phi: C \rightarrow \{0,1\}, \mathrm{k} \in_R K$ – равномерно распределённый и выполняется

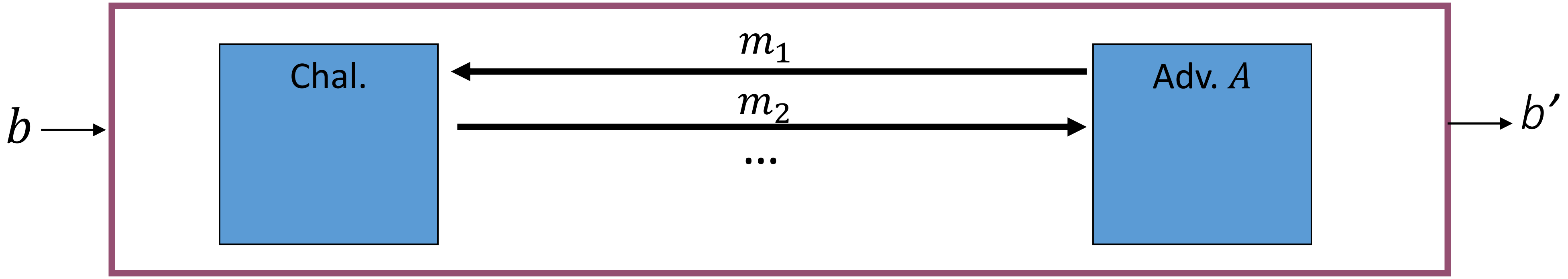
$$\Pr[\phi(E(\boldsymbol{k}, m_0) = 1] = \Pr\big[\phi\big(E(\boldsymbol{k}, m_1)\big) = 1\big]$$

Ослабим свойство абсолютной стойкости: вместо требования равенства вероятностей потребуем чтобы их разность не превосходила величину ϵ:

$$\Pr[\phi(E(\boldsymbol{k}, m_0) = 1] - \Pr\big[\phi\big(E(\boldsymbol{k}, m_1)\big) = 1\big] \leq \epsilon$$

Понятие игры

- Игра состоит из двух сторон – **противника** A (**Adversary**) и **претендента** (**Challenger**), моделируемые **эффективными** алгоритмами. При этом алгоритм A – вероятностный
- **Входом** игры называется некоторая величина b
- **Ход игры** – атакующий и претендент обмениваются сообщениями согласно некоторому фиксированному протоколу
- **Результатом** игры называется некоторая величина b'

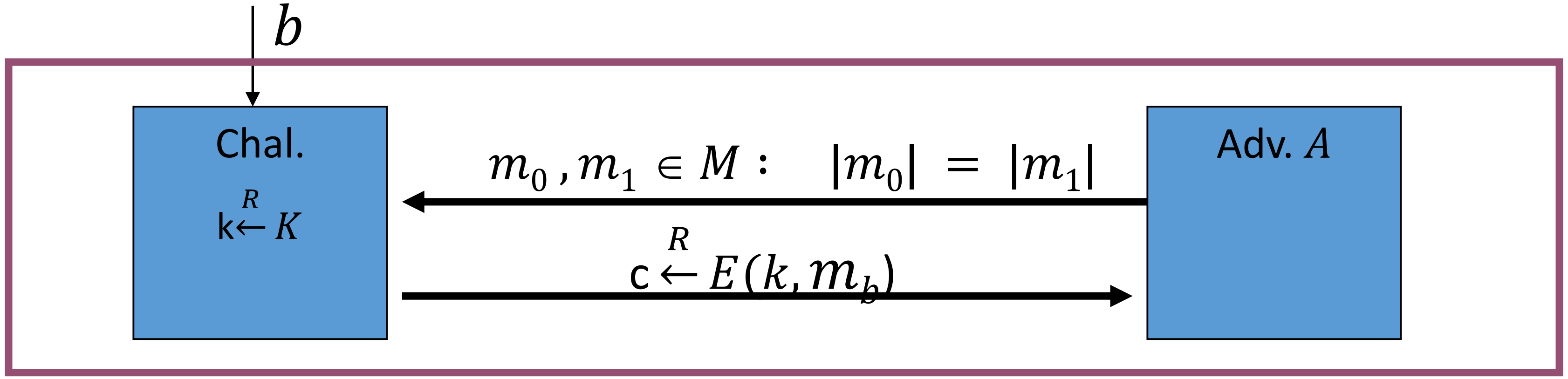


Понятие игры на различимость, определения

- **Входом** игры называется случайное число $b \in \{0,1\}$, неизвестное для атакующего, определяющего эксперимент
- **Экспериментом** $(Exp\ b)$ называется «режим» претендента при его общении с атакующим
- **Ход игры** – атакующий и претендент обмениваются сообщениями согласно некоторому фиксированному протоколу
- **Цель игры** – атакующий пытается угадать число b (угадать эксперимент)
- **Результатом** игры называется число $b' \in \{0,1\}$ – выход алгоритма A

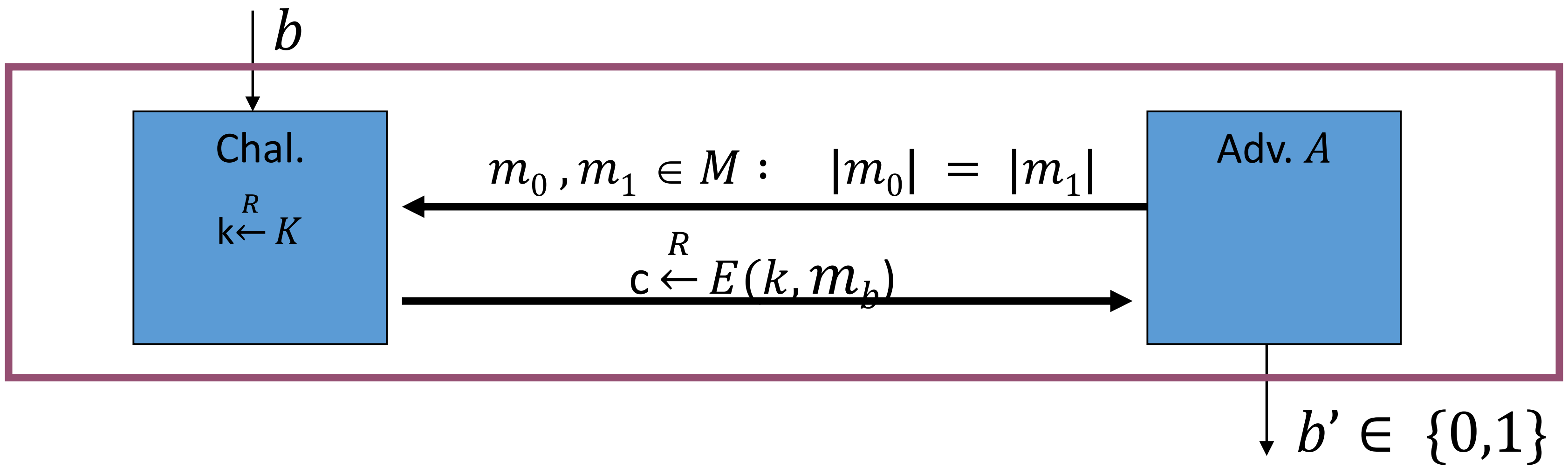
Игра: семантическая стойкость (одноразовое использование ключа)

Для $\mathrm{E} = (E, D)$ - вычислимого шифра на (K, M, C) и противника A определим два эксперимента, Experiment 1 и Experiment 2 следующим образом:



Игра: семантическая стойкость (одноразовое использование ключа)

- Противник выбирает сообщения $m_0, m_1 \in M$ **одинаковой длины**
- Претендент выбирает $k \xleftarrow{R} K$, c$\xleftarrow{R} E(k, m_b)$ и отправляет атакующему
- Противник выставляет бит $b' \in \{0,1\}$ как результат игры

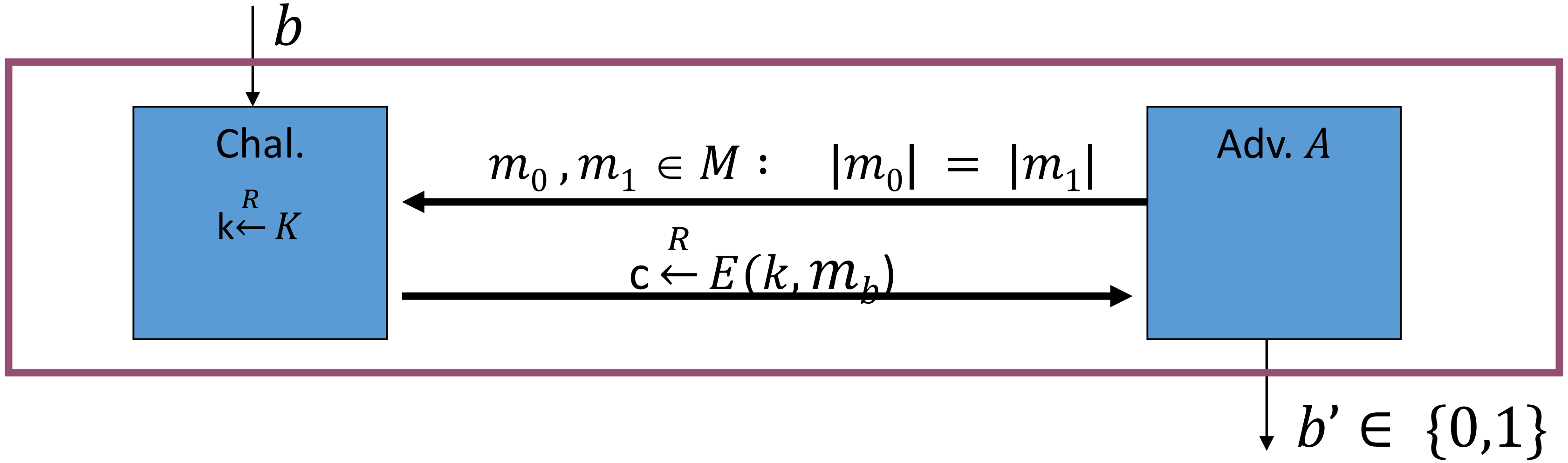


Игра: семантическая стойкость (одноразовое использование ключа)

Пусть W_b - событие того, что $b' = 1$ в эксперименте b.

Преимуществом (**Advantage**) противника A против алгоритма E в семантической игре есть величина:

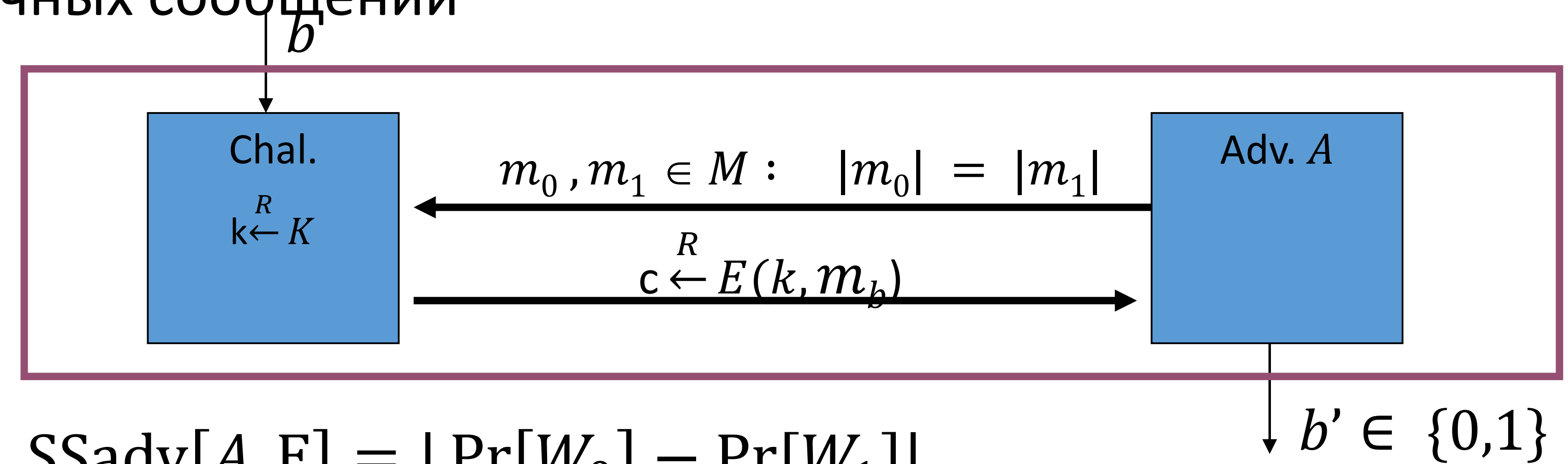
$$\text{SSadv}[A, \text{E}] = |\Pr[W_0] - \Pr[W_1]|$$



Семантическая стойкость (одноразовое использование ключа)

Шифр E - (одноразово) **семантически стойкий**, если для всех эффективных противников A величина $\mathrm{SSadv}[A, \mathrm{E}]$ – пренебрежимо малая величина

Иными словами – вычислительно невозможно отличить шифрректсты различных сообщений



$$\mathrm{SSadv}[A, \mathrm{E}] = |\Pr[W_0] - \Pr[W_1]|$$

Семантическая стойкость

- «Ослабленная» версия абсолютной стойкости: только эффективные противники и разность вероятностей расшифрования в заданные сообщения не превосходит ϵ.
- Позволяет использовать короткие ключи

Примеры:

- Одноразовый блокнот – семантически стойкий шифр
- Одноразовый блокнот переменной длины – семантически стойкий шифр
- Шифр подстановки – не семантически стойкий шифр

Построение атаки на семантическую стойкость

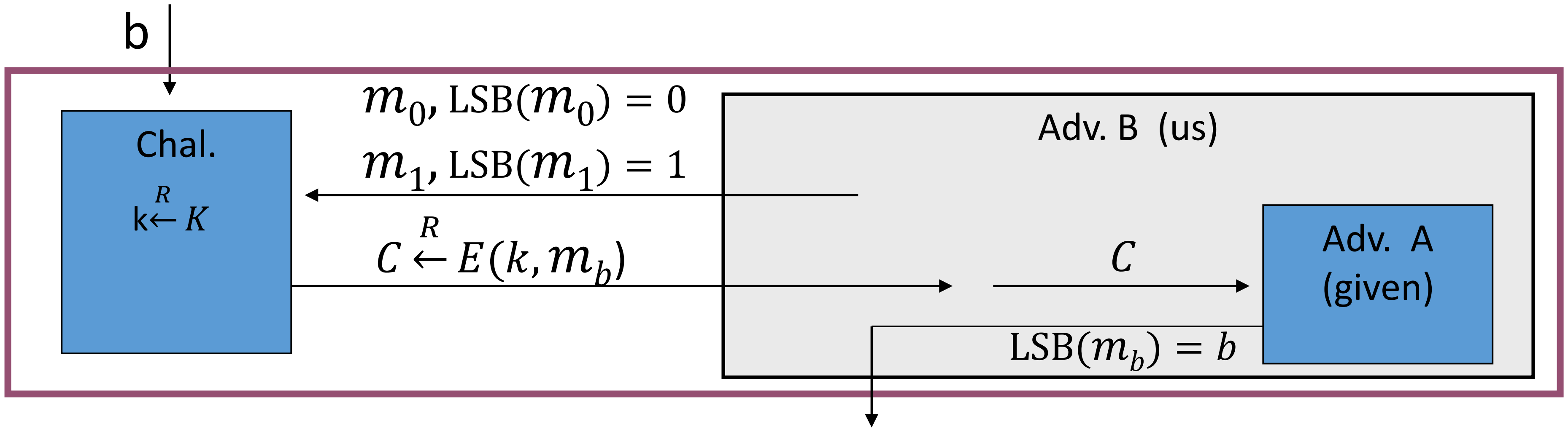
Пусть A – алгоритм позволяющий получить наименее значимый бит (LSB) открытого текста через шифртекст с $\stackrel{R}{\leftarrow} E(k, m)$. Тогда $\mathrm{E} = (E, D)$ – не семантически стойкий шифр.

⊳Построим эффективный алгоритм B, позволяющий выиграть игру на семантическую стойкость.

- Генерация двух сообщений m_0, m_1
- Получение шифртекста c
- Передача шифртекста на вход алгоритма A
- Передача полученного наименее значимого бита как результата игры. ⊲

Построение атаки на семантическую стойкость

Пусть A – алгоритм позволяющий получить наименее значимый бит (LSB) открытого текста через шифртекст с $\xleftarrow{R} E(k, m)$. Тогда $\mathrm{E} = (E, D)$ – не семантически стойкий шифр.



$$\mathrm{SSadv}[A, \mathrm{E}] = |\Pr[W_0] - \Pr[W_1]| = |1 - 0| = 1$$

Доказательства сведением (Reduction proof)

Пусть $\mathrm{E} = (E, D)$ - вычислимый семантически стойкий шифр на (K, M, C).
Тогда $\mathrm{E}' = (E', D')$: $\begin{cases} (c_0, c_1) = \ E'(k, m) = E(k, m)||E(k, m) \\ D'(k, (c_0, c_1)) = D(k, c_0) \end{cases}$ –
семантически стойкий шифр.

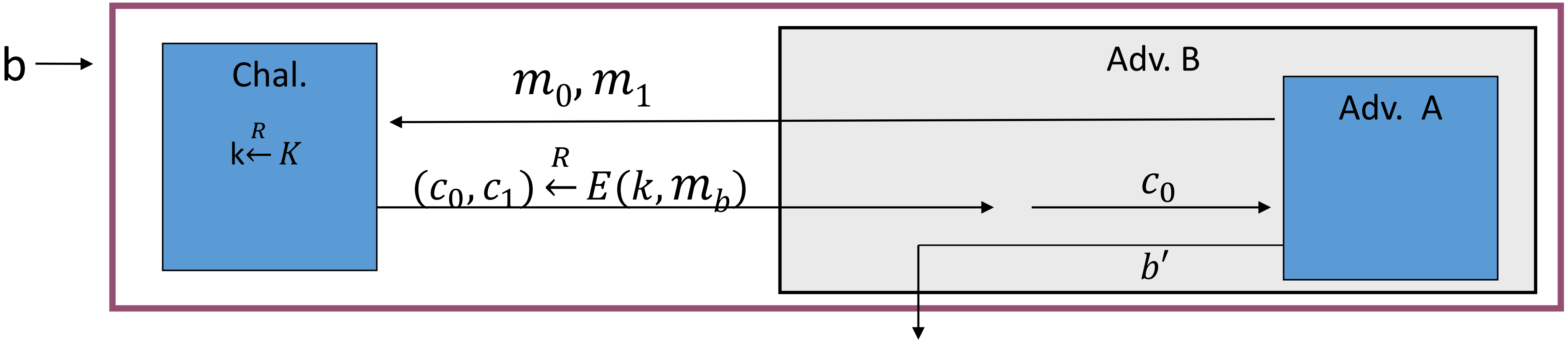
⊳ От противного. Пусть E' - не семантически стойкий шифр. Тогда $\exists$ противник A: $\mathrm{SSadv}[A, \mathrm{E}'] \geq e$, где e – не пренебрежимо малая величина.

Построим эффективный алгоритм B для семантической игры против шифра E с использованием алгоритма A, показав тем самым что E – не семантический стойкий $\Rightarrow$ противоречие $\Rightarrow$ E' – семантический стойкий.
⊲

Доказательства сведением (Reduction proof)

Пусть $\mathrm{E} = (E, D)$ - вычислимый семантически стойкий шифр на (K, M, C).
Тогда $\mathrm{E}' = (E', D')$: $\begin{cases} (c_0, c_1) = E'(k,m) = E(k,m)||E(k,m) \\ D'(k,(c_0,c_1)) = D(k,c_0) \end{cases}$ –
семантически стойкий шифр.

$\mathrm{SSadv}[A, \mathrm{E}'] \geq e$, где e – не пренебрежимо малая величина.



$\mathrm{SSadv}[B, \mathrm{E}] = \mathrm{SSadv}[A, \mathrm{E}'] \geq e$

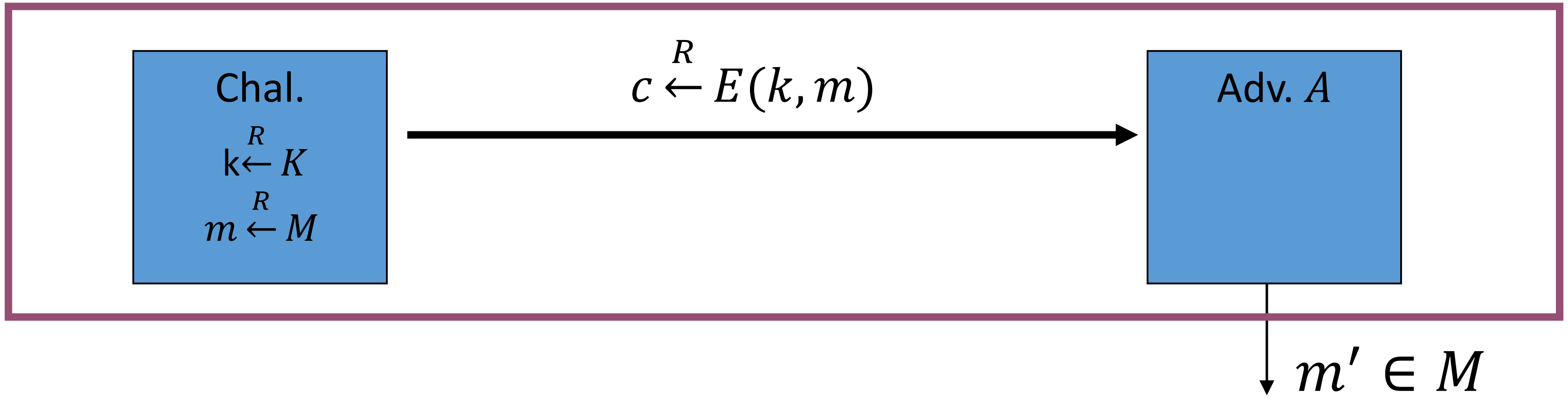
Восстановление сообщений

Атака на восстановление сообщений: имея зашифрованное сообщение $c \leftarrow E(k, m), m \in M$, восстановить сообщение m, с вероятностью больше $1/|M|$.

Опишем игру на восстановление сообщений.

- Претендент вычисляет $m \overset{R}{\leftarrow} M, k \overset{R}{\leftarrow} K, c \overset{R}{\leftarrow} E(k, m)$ и отправляет c противнику.
- Противник возвращает m' как результат игры.

Восстановление сообщений

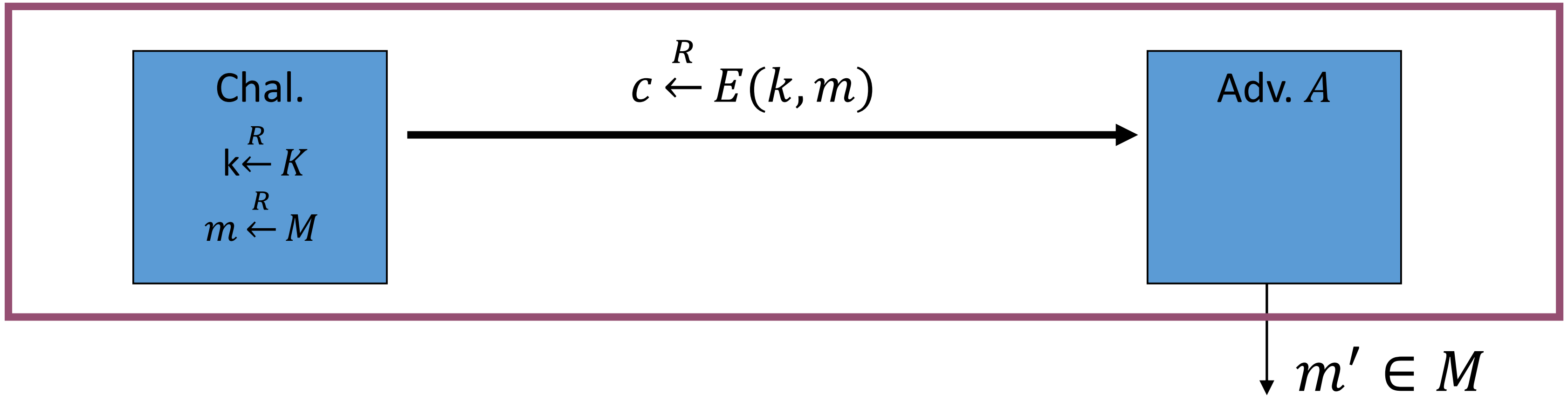


Пусть W – событие, при котором $m' = m$.

Преимуществом алгоритма A против шифра E при атаке на восстановление сообщений является величина

$$\mathrm{MRadv}[A, \mathrm{E}] = \left| \Pr[W] - \frac{1}{|M|} \right|$$

Восстановление сообщений



$$\mathrm{MRadv}[A, \mathrm{E}] = \left| \Pr[W] - \frac{1}{|M|} \right|$$

Шифр E называется **стойким к атаке на восстановление сообщений**, если $\forall A$ величина $\mathrm{MRadv}[A, \mathrm{E}] < \epsilon$, где ϵ - пренебрежимо малая величина.

Восстановление сообщений

Теорема 1.8. Атака на восстановление сообщений более слабая, чем на семантическую стойкость (Если шифр $\mathrm{E} = (E, D)$ семантически стойкий на (K, M, C), то он стойкий к атаке на восстановление сообщений)

⊳ Пусть A – эффективный алгоритм. Обозначим p – вероятность выиграть игру на восстановление сообщений для алгоритма A:

$$\mathrm{MRadv}[A, \mathrm{E}] = \left| p - \frac{1}{|M|} \right|.$$

Построим эффективный алгоритм B для игры на семантическую стойкость простив алгоритма E, для которого

$$\mathrm{MRadv}[A, \mathrm{E}] \leq \mathrm{SSadv}[B, \mathrm{E}].$$

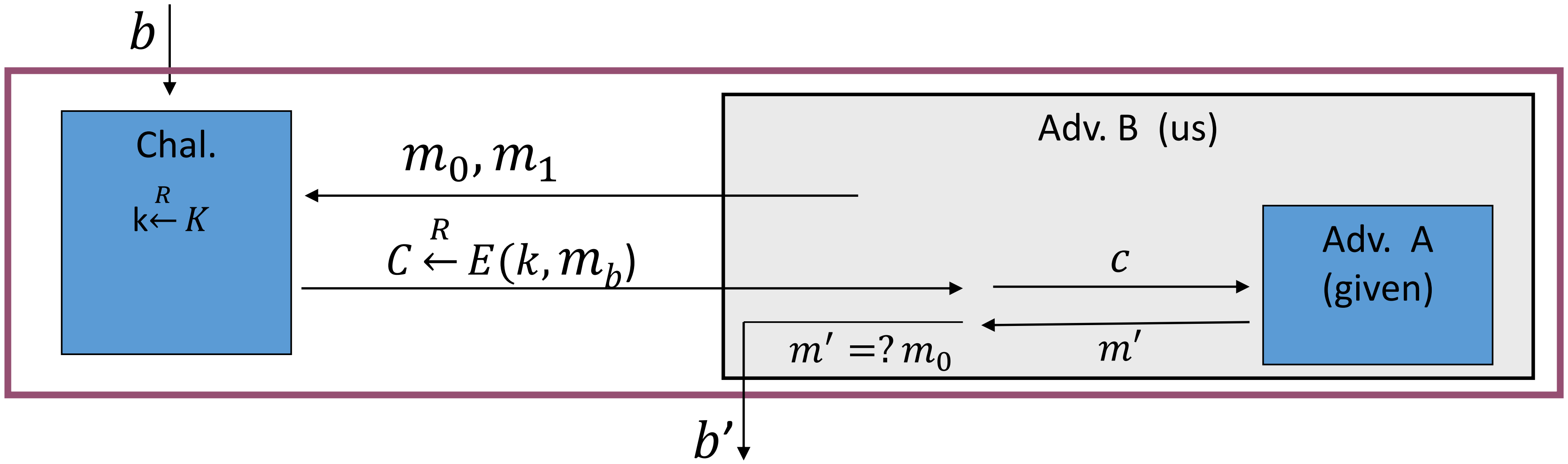
Восстановление сообщений

Теорема 1.8. Атака на восстановление сообщений более слабая, чем на семантическую стойкость (Если шифр $\mathrm{E} = (E, D)$ семантически стойкий на (K, M, C), то он стойкий к атаке на восстановление сообщений)

Построим алгоритм B. Алгоритм B генерирует два сообщения m_0 и m_1и оправляет их претенденту в игре на семантическую стойкость. Претендент отвечает шифртекстом c одного из сообщений, которых алгоритм B пересылает алгоритму A, получая восстановленное сообщение m'. Если $m' = m_0$ то выводит $b' = 0$, иначе $b' = 1$.

Восстановление сообщений

Теорема 1.8. Атака на восстановление сообщений более слабая, чем на семантическую стойкость (Если шифр $\mathrm{E} = (E, D)$ семантически стойкий на (K, M, C), то он стойкий к атаке на восстановление сообщений)



Восстановление сообщений

Теорема 1.8. Атака на восстановление сообщений более слабая, чем на семантическую стойкость (Если шифр $\mathrm{E} = (E, D)$ семантически стойкий на (K, M, C), то он стойкий к атаке на восстановление сообщений)

Для $b = 0{,}1$ пусть p_b - вероятность того, что алгоритм B выдаст значение $b' = 1$, при шифровании сообщения m_b. Тогда $\mathrm{SSadv}[B, \mathrm{E}] = |p_0 - p_1|$. С другой стороны, если c есть зашифрование m_0 то вероятность $p_0 = p$ (Вероятность выиграть игру на восстановление для A). Если же c есть зашифрование m_1, то $p_1 = \mathrm{Pr}[m_1 = m'] = 1/|M|$. Следовательно

$$SSadv[B, \mathrm{E}] = |p_1 - p_0| = \left|\frac{1}{|M|} - p\right| = MRadv[A, \mathrm{E}]$$

⇒ атака на восстановление даёт атаку на семантическую стойкость. ⊲

Восстановление битов сообщения

Пусть $\mathrm{E} = (E, D)$ шифр на (K, M, C). $M = \{0,1\}^L$. Пусть $par(m)$ – функция вычисления чётности сообщения $m \in M$.

Определим игру на восстановление битов.

- Претендент вычисляет $m \xleftarrow{R} M, k \xleftarrow{R} K, c \xleftarrow{R} E(k, m)$ и отправляет c противнику.
- Противник возвращает $\mathrm{b}' \in \{0,1\}$ как результат игры.

Пусть W – событие, при котором $b' = par(m)$.

Преимуществом алгоритма A против шифра E при атаке на восстановление битов является величина

$$\mathrm{PARadv}[A, \mathrm{E}] = |\Pr[W] - 1/2|$$

Восстановление битов сообщения

Пусть $\mathrm{E} = (E, D)$ шифр на (K, M, C). $M = \{0,1\}^L$. Пусть $par(m)$ – функция вычисления чётности сообщения $m \in M$.

Шифр E называется **стойким к восстановлению битов**, если величина $\mathrm{PARadv}[A, \mathrm{E}] < \epsilon$, где ϵ – пренебрежимо малая величина.

Вычисление индивидуальных битов сообщений

Теорема 1.9. Атака на восстановление битов сообщения более слабая, чем на семантическую стойкость (Если шифр $\mathrm{E} = (E, D)$ семантически стойкий на (K, M, C), то он стойкий к атаке на восстановление битов сообщения)

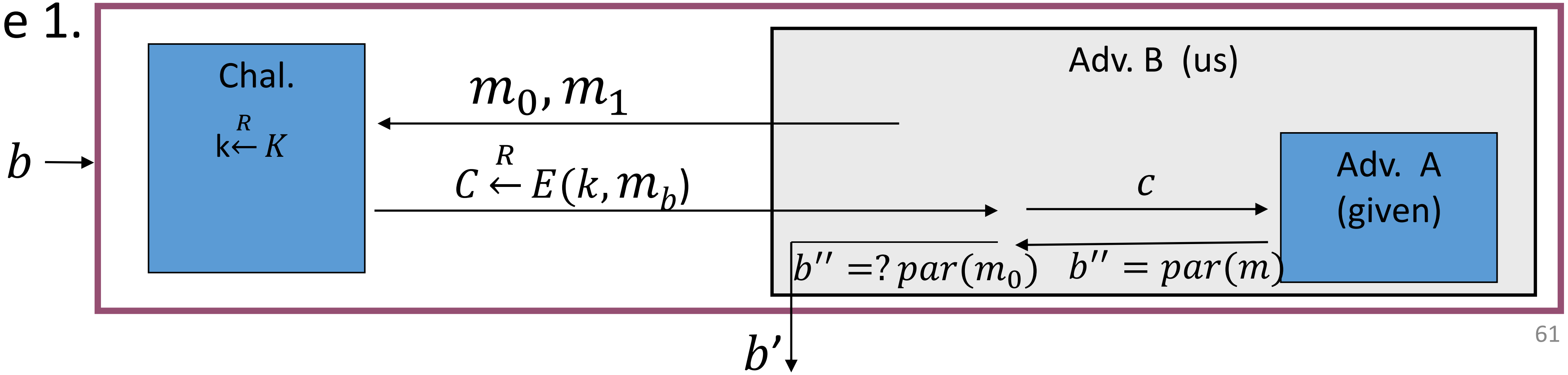
⊳ Построим эффективный алгоритм B для игры на семантическую стойкость простив алгоритма E, для которого

$$\mathrm{PARadv}[A, \mathrm{E}] = \frac{1}{2}\mathrm{SSadv}[B, \mathrm{E}].$$

Вычисление индивидуальных битов сообщений

Теорема 1.9. Атака на восстановление битов сообщения более слабая, чем на семантическую стойкость (Если шифр $\mathrm{E} = (E, D)$ семантически стойкий на (K, M, C), то он стойкий к атаке на восстановление битов сообщения)

Противник B генерирует сообщения $m_0, m_1 \leftarrow m_0 \oplus (0^{L-1}1)$ и отправляет претенденту, получая шифртекст c, который он передаёт алгоритму A. После получения значения b'' если $b'' = par(m_0)$ то $b' = 0$, иначе 1.



Вычисление индивидуальных битов сообщений

Теорема 1.9. Атака на восстановление битов сообщения более слабая, чем на семантическую стойкость (Если шифр $\mathrm{E} = (E, D)$ семантически стойкий на (K, M, C), то он стойкий к атаке на восстановление битов сообщения)

Пусть A: $\mathrm{PARadv}[A, \mathrm{E}] = \epsilon$, т.е. вероятность угадать чётность есть $\frac{1}{2} + \epsilon$.

Для $b = 0{,}1$ пусть p_b - вероятность того, что алгоритм B выдаст значение $b' = 1$. Тогда $SSadv[B, \mathrm{E}] = |p_1 - p_0| = 2\epsilon = \mathrm{PARadv}[A, \mathrm{E}]$.

$$p_0 = \frac{1}{2} + \epsilon \text{ (верная чётность } m_0\text{)},$$

$$p_1 = 1 - p_0 = \frac{1}{2} - \epsilon \text{(неверная чётность } m_1\text{)}. \triangleleft$$

Семантическая стойкость (альтернативная формулировка)

Теорема 1.10. (обобщение 1.9) Пусть задана игра на семантическую стойкость для алгоритма A против шифра $\mathrm{E} = (E, D)$ на (K, M, C)**.** Определим $\mathrm{SSadv}^*[A, \mathrm{E}] = \left|\Pr[W] - \frac{1}{2}\right|$, где W – событие, при котором $\mathrm{b}' = \mathrm{b}$. Тогда $\mathrm{SSadv}\,[A, \mathrm{E}] = 2 * \mathrm{SSadv}^*[A, \mathrm{E}]$

⊳доказательство аналогично **Теореме 1.9.** ⊲

Выводы

- Модель абсолютно стойкого шифра делает его сложно применимым в практическом смысле
  - Требуется размер ключа равный размеру сообщения
  - Невозможно добиться стойкости при переменной длине сообщений
- Семантически стойкий шифр – ослабленная модель абсолютно стойкого шифра, пригодная для практического применения
  - Стойкость к восстановлению сообщений
  - Стойкость к восстановлению битов сообщений
- Игровая модель – модель, позволяющая вводить определения стойкости для криптографический примитивов
  - Доказательства стойкости методом сведения (reduction)
  - Построение атак через моделирование игры