

Прикладная Криптография: Симметричные криптосистемы Абсолютная и Семантическая стойкость

Макаров Артём

МИФИ 2020

Структура курса



- Лекции: 16 недель
- Сдача разделов: 3 блока
 - Для каждого блока жёсткий дедлайн (без переносов)
 - <https://github.com/CryptoCourse/CryptoLectures/wiki/Список-домашних-работ-и-лекций>
 - <https://github.com/CryptoCourse/CryptoLabs/wiki/список-лабораторных-работ>
 - **Штраф за пропуск дедлайна: -5/100 к итоговой оценке за семестр за каждый дедлайн в неделю**
- Для сдачи каждого блока:
 - Сдача лабораторных работ для данного блока
 - Сдача домашних работ
 - Сдача теории по лабораторным и домашним

СВЯЗЬ



<https://vk.com/zmacr>

vk.com
(Вопросы)



<https://discord.gg/Vb38A6H>

Discord
(сдача лаб и дз)



<https://t.me/f1589>

t.me
(Вопросы)

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

Материалы прошлого года

- Курс обновляется в момент чтения. Материалы прошлого года доступны, но еженедельно обновляются.
- Доверять и использовать нужно только текущие материалы, т.е. материалы всех прошедших в семестре лекций и лабораторных заданий текущего блока.
- Не рекомендуется выполнять задания «наперёд», так как материал может измениться

Материалы прошлого года

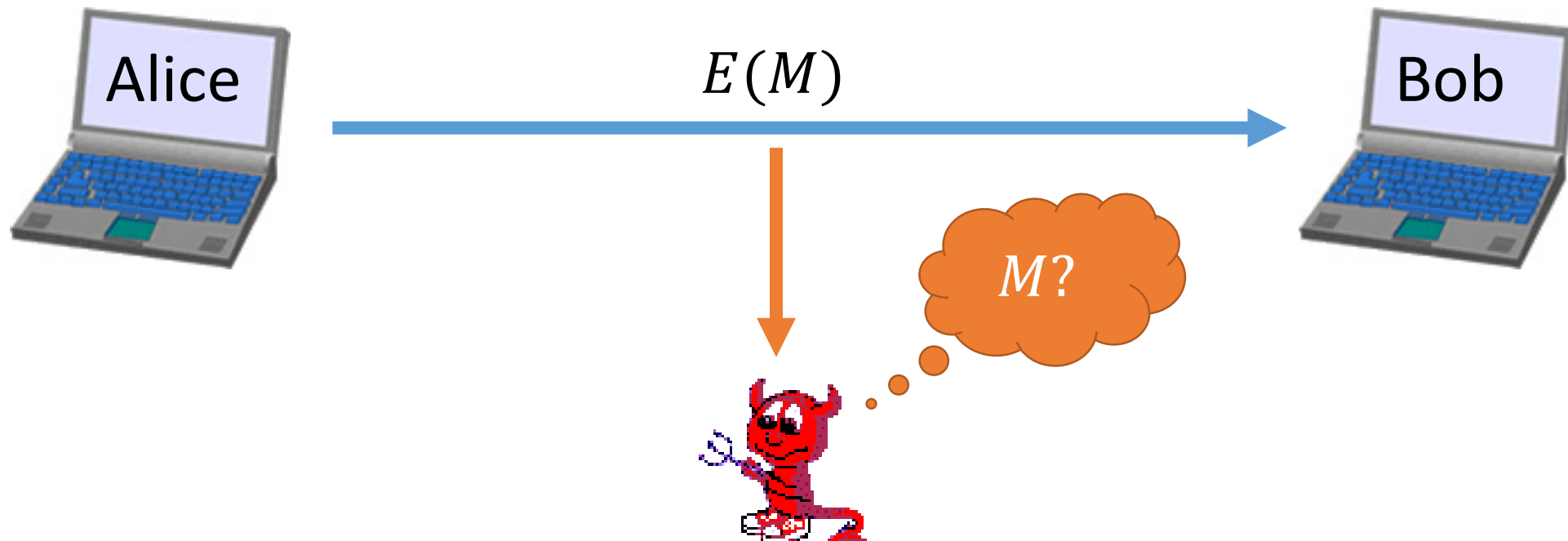
Название	Описание	Блок	Сроки сдачи
Атака при многократном использовании одноразового блокнота	link	1	07.09.19 - 21.09.19(06:00)
Атака на аутентичность при использовании поточных шифров	link	1	07.09.19 - 21.09.19(06:00)
X Атака на аутентичность блочного шифра в режиме CBC	meh	2	20.09.18 - 01.11.18(06:00)

Лекция	Описание	Блок	Сроки сдачи домашней работы
1	Абсолютная и семантическая стойкость (лекция , задание)	1	14.09.19
2	Поточные шифры (лекция , задание)	1	XXXX
3	Практические аспекты (лекция)	1	null

Обратная связь и пожелания по курсу

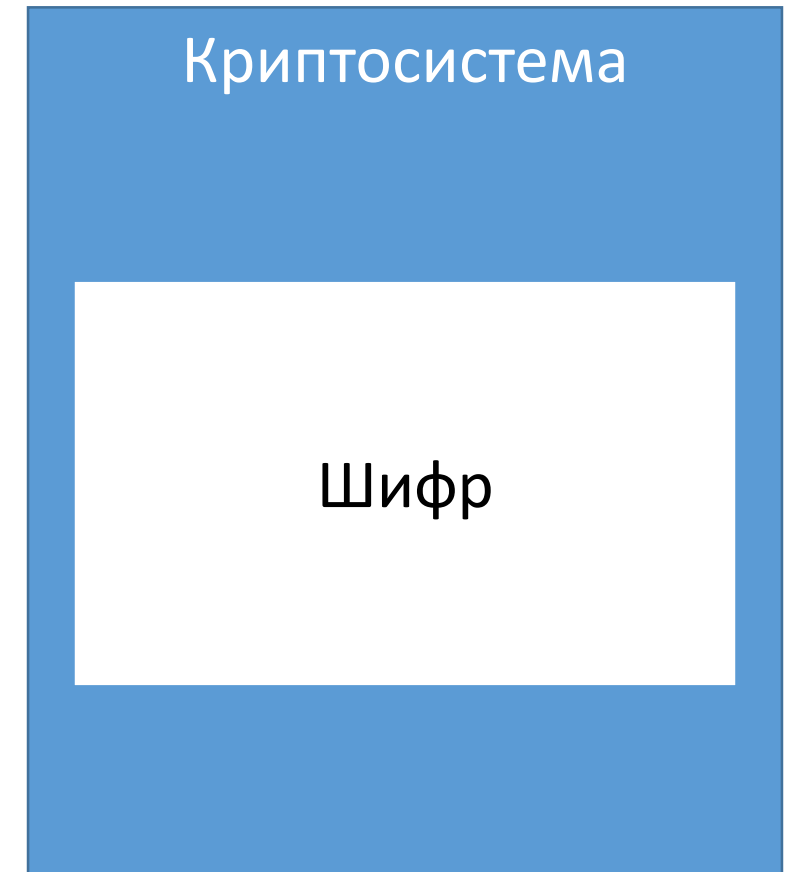
Историческая задача криптографической защите информации

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



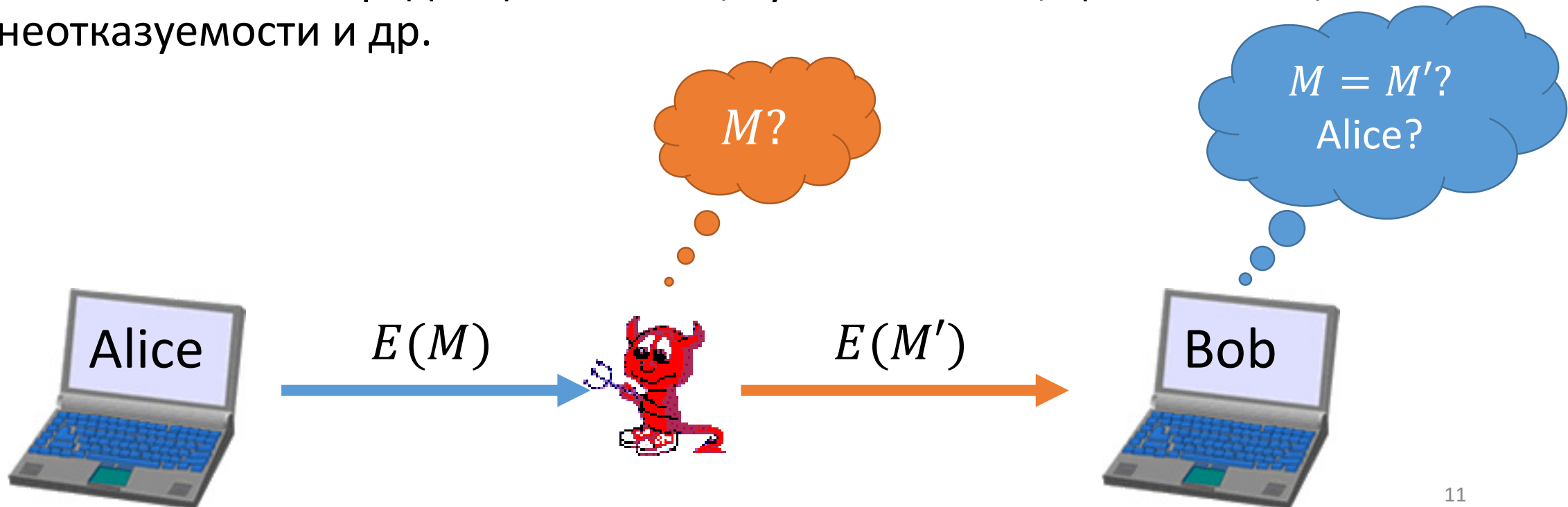
Способы построения и анализа криптосистем

- **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.



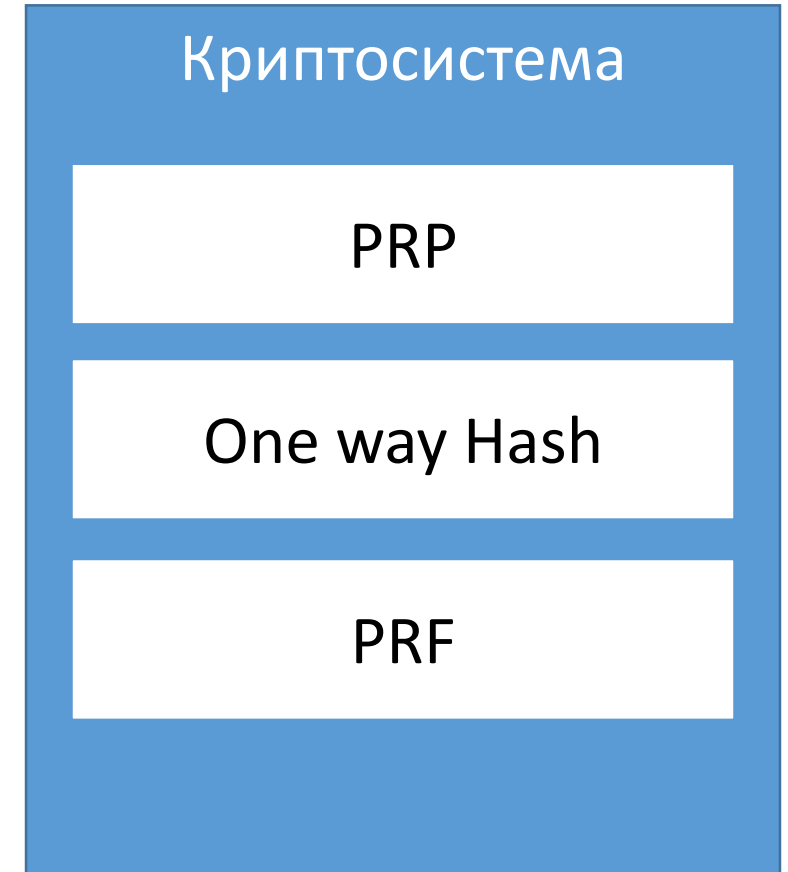
Способы построения и анализа криптосистем

- **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 конструкции, 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 практической стойкости криптографического примитива является сведение атаки на него к вычислительно сложной задаче. Иными словами показывается, что произвести атаку на примитив так же сложно как решить вычислительно сложную задачу.



Сведение стойкости (Security Reduction)

Доказательство стойкости криптосистемы показывается сведением её к стойкости криптографических примитив. При современном подходе описание системы использует только абстрактные модели примитивов (PRF, PRP, и другие).



Сведение стойкости (Security Reduction)

Пусть A – стойкая система. Показать что система B стойкая. ($A \rightarrow B$).
(Показать сведение стойкости системы B к стойкости системы A .)

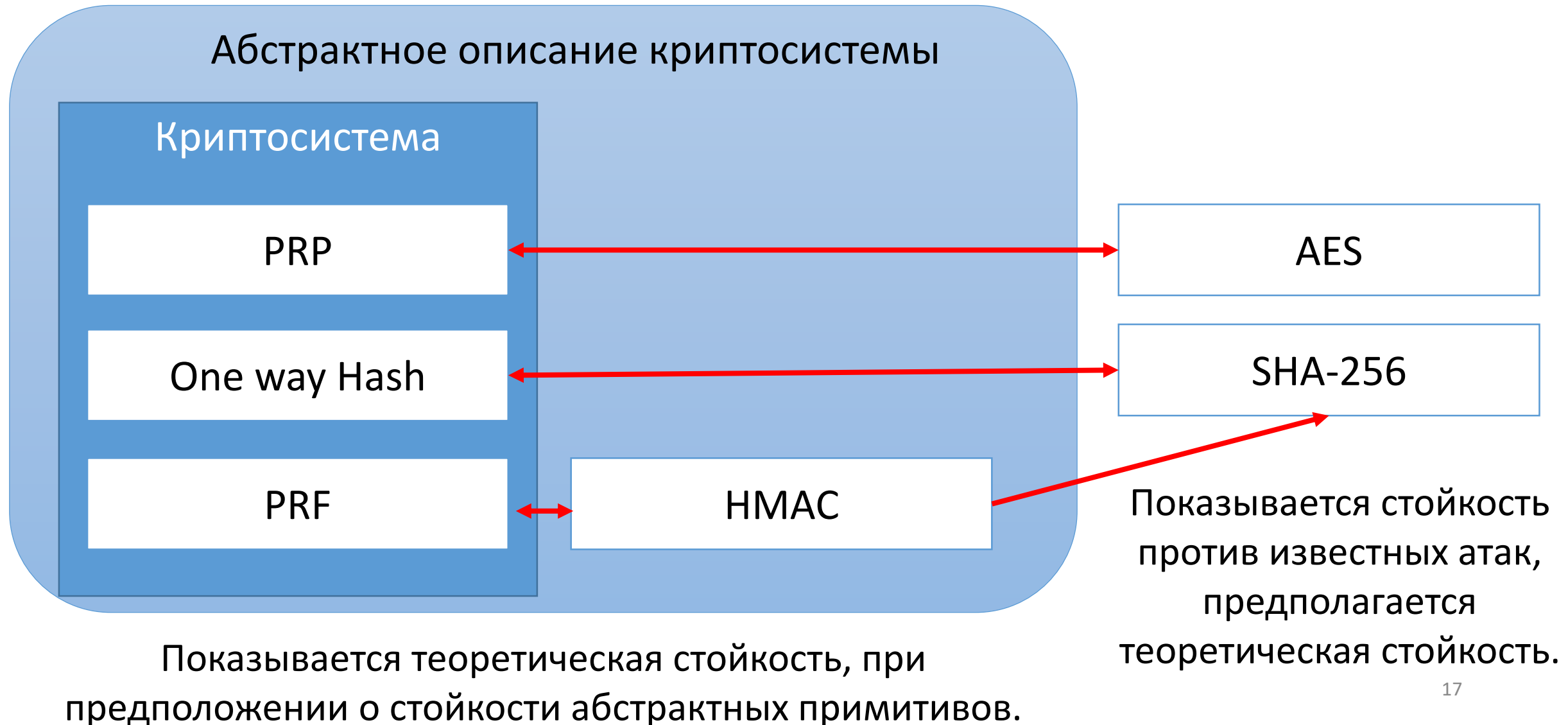
▷ От противного. Пусть существует атака на систему B . Попробуем использовать эту атаку для построения атаки на систему A .

(Строим атаку на систему A).

Следовательно, из предположения нестойкости системы B
(предположения о наличии атаки) мы построили атаку на систему A ,
 $\bar{B} \rightarrow \bar{A}$.

Но система A – стойкая, следовательно предположение не верно и B – стойкая. ◁

Сведение стойкости (Security Reduction)



Сведение стойкости криптографических примитивов

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

Шифр Шеннона

Шифр Шеннона - пара функций $E = (E, D)$, таких что:

- (1) Функция E (**функция зашифрования**) принимает на вход ключ k и сообщение m (называемой открытым текстом, РТ) и даёт на выходе шифртекст c (СТ), 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**):
$$\forall k, \forall m \ D(k, E(k, m)) = m.$$

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

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

$$E: K \times M \rightarrow C,$$

$$D: K \times C \rightarrow M,$$

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

Нотация

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

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 ($|v| = n$)

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

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

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

Нотация

$x \in_R X$ - $x \in X$, выбранный случайно равновероятно

$x \leftarrow_R X$ – выбор случайного равновероятного $x \in X$

О.Т. (Р.Т.) – Открытый текст (Plain Text)

Ш.Т (С.Т.) – Шифртекст (Cipher Text)

Пример: Одноразовый блокнот

Пусть $E = (E, D)$ – **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).

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

Пример: Одноразовый блокнот переменной длины

Пусть $E = (E, D)$ – **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).

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

Пример: Шифр подстановки

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

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

Корректность: $D(k, E(k, m)) =$

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

Пример: Аддитивный одноразовый блокнот

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

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

$$E(k, m) = (m + k) \bmod n, \text{ по координатам}$$

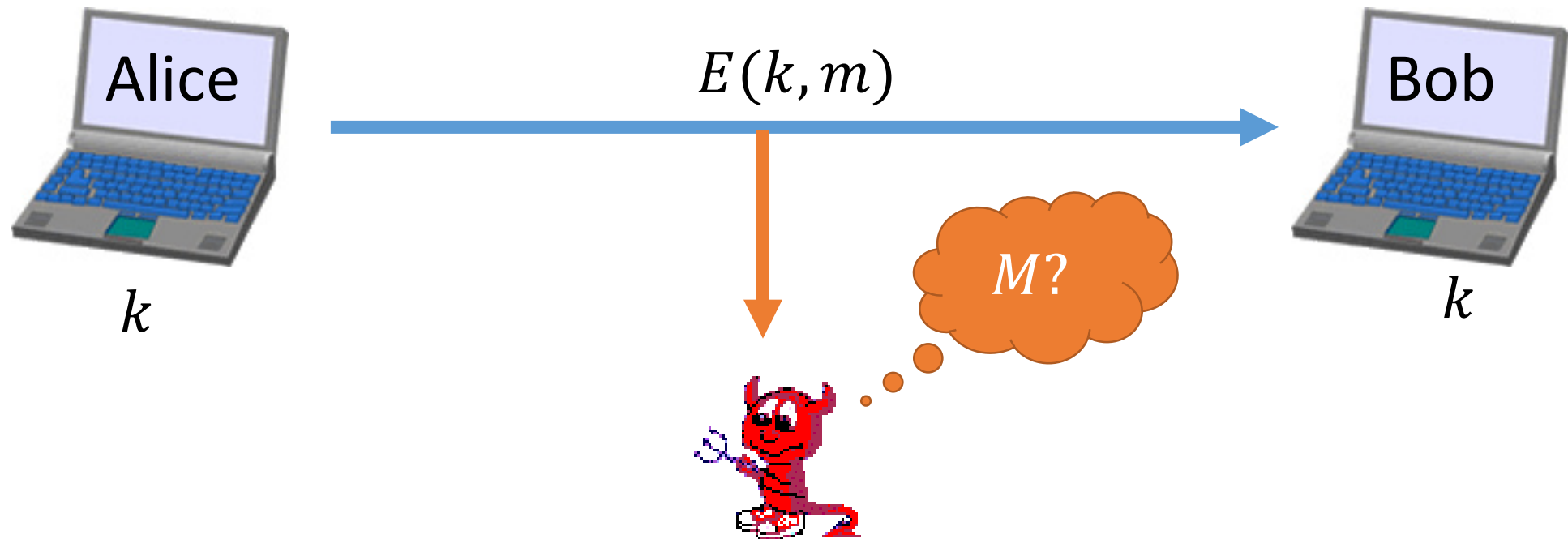
Для ключа $k \in K$ и шифртекста $c \in C$ функция **расшифрования** определена как:

$$D(k, c) = (c - k) \bmod n, \text{ по координатам}$$

Корректность: $D(k, E(k, m)) = D(k, m + k) = (m + k) - k = m.$

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 шифровании однобитного сообщения

Понятие стойкости

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

$$K_0 \subset K: E(k_0, m_0) = c,$$

$$K_1 \subset K: E(k_1, m_1) = c,$$

$$|K_0| > |K_1|$$

$$m_0, m_1 \in M: m_0 \neq m_1; (k_0, k_1) \in (K_0 \times K_1)$$

Вероятность расшифровать C как m_0 ($|K_0| = 800, |K_1| = 600$):

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

Абсолютная стойкость

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

Если $\forall m_0, m_1 \in M$ и $c \in C$ имеем:

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

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 $E = (E, D)$ - 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 $\mathbf{k} \in_R K$ тогда все случайные величины $E(\mathbf{k}, m)$ имеют одинаковое распределение

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

Эквивалентные определения абсолютной стойкости

Теорема 1.1. Пусть $E = (E, D)$ - 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 $\mathbf{k} \in_R K$ тогда все случайные величины $E(\mathbf{k}, m)$ имеют одинаковое распределение

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

Эквивалентные определения абсолютной стойкости

Теорема 1.1. Пусть $E = (E, D)$ - 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 $\mathbf{k} \in_R K$ тогда все случайные величины $E(\mathbf{k}, m)$ имеют одинаковое распределение

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

Одноразовый блокнот – абсолютно стойкий шифр

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

▷ 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** ◁

Одноразовый блокнот переменной длины – не абсолютно стойкий шифр

Теорема 1.3. Пусть $E = (E, D)$ - 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 $m_0 \in M: |m_0| = 1$, $m_1 \in M: |m_1| > 1$, $c \in C: |c| = 1$

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

(Шифртекст длинны 1 не может иметь открытый текст длинны > 1)

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

Предикат

Пусть имеется некоторый элемент $s \in S$.

Пусть мы хотим получить некоторую информацию обладая s . Пусть функция $F(s)$ – есть функция «получения» некоторой информации из s .

Предикатом на множестве S назовём булеву функцию $\phi: S \rightarrow \{0,1\}$.

Тогда вычисление предиката $F(s) = \phi(s)$ есть минимальная функция «получения» информации из s (функция получения информации, с выходом 1 бит).

Альтернативная трактовка предиката – бинарная различимость элементов множества.

Эквивалентные определения абсолютной стойкости

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

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

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

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

$$\begin{aligned} \Pr[\phi(E(\mathbf{k}, m_0)) = 1] &= \sum_{c \in S} \Pr[E(\mathbf{k}, m_0) = c] = \\ &= \sum_{c \in S} \Pr[E(\mathbf{k}, m_1) = c] = \Pr[\phi(E(\mathbf{k}, m_1)) = 1] \end{aligned}$$

Эквивалентные определения абсолютной стойкости

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

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

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

Пусть E – **не** абсолютно стойкий. То есть $\exists c_0 \in C$:

$$\Pr[E(\mathbf{k}, m_0) = c_0] \neq \Pr[E(\mathbf{k}, m_1) = c_0].$$

Пусть $\phi: \phi(c_0) = 1, \phi(c') = 0, \forall c' \neq c_0$

$$\begin{aligned} \Pr[\phi(E(\mathbf{k}, m_0)) = 1] &= \Pr[E(\mathbf{k}, m_0) = c_0] \neq \\ \Pr[E(\mathbf{k}, m_1) = c_0] &= \Pr[\phi(E(\mathbf{k}, m_1)) = 1] \end{aligned}$$

◁

Эквивалентные определения абсолютной стойкости

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

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

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

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

- Если E – абсолютно стойкий, тогда \mathbf{c} и \mathbf{m} независимы:
- Если \mathbf{c} и \mathbf{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:

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

То есть наличие шифртекста не даёт злоумышленнику никаких преимуществ.

Энтропия

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

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

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

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

Энтропия

Эквивалентные определения

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

Иными словами шифртекст не даёт никакой информации об открытом тексте.

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

Плохие новости

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

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

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.

Семантическая стойкость

Продолжение следует...