Защита информации

Павел Юдаев

МГТУ им. Баумана, Кафедра ИУ-9

Москва, 2014

Раздел 2 - Поточные шифры и генераторы ПСЧ

Опр.

генератор псевдослучайных чисел (ГПСЧ), pseudo random generator - это функция
$G : \{0,1\}^s \to \{0,1\}^n$, где $n \gg s$, вычислимая за время $poly(n)$ детерминированным алгоритмом.

Опр.

поточный шифр:
$c = E(k, m) = m \oplus G(k)$,
$m = D(k, c) = c \oplus G(k)$

Задача

Может ли поточный шифр быть абсолютно стойким?

Недостатки поточных шифров.

1. Детерминированное шифрование.

Текст 1: “To: Bob. Some text”
Текст 2: “To: Eve. Some text”
Шифротексты этих текстов совпадут, кроме одного фрагмента.

Вывод: при шифровании трафика поточным шифром, необходимо согласовывать новый, независимый ключ для каждой сессии.

Недостатки поточных шифров.

2. Не обеспечивают контроль целостности данных.

$c(m, F(k)) = m \oplus F(k)$
Пусть злоум-к может менять шифротекст.
Пусть $c \to c' : \ c' = c \oplus p$, тогда $m' = m \oplus p$.

Получатель не может это обнаружить, если не использует дополнительные средства контроля.

Например, можно подменить адрес получателя сетевого пакета. В протоколе IPSec подмена адреса сетевого пакета приведет к тому, что сервер, расшифровав пакет, отправит расшифрованный текст другому получателю.

Раздел 2 - Поточные шифры и генераторы ПСЧ

$c = E(k, m) = m \oplus G(k)$

ГПСЧ $G(k) \in \mathrm{PT}$

Стойкость поточного шифра зависит от используемого ГПСЧ.

Опр.

PT - класс детерминированных алгоритмов, имеющих *полиномиальное время работы*. Алгоритм $A \in \mathrm{PT}$, если
$\exists poly\ p(n) : \forall x \in X : len(x) \leq n$
$time(A(x)) \leq p(n)$.

Опр.

Задача разрешимости (decision problem) - задача, имеющая два ответа ДА и НЕТ.

Пример

1) Дано описание функции f. Существует ли $x: \ f(x) = 0$?
2) Дана КНФ f. Существует ли $x: \ f(x) = 1$? (Задача выполнимости, SAT.)

Опр.

Задача поиска (search problem) - задача найти элемент множества, для которого выполнено заданное отношение, или установить, что таких элементов нет.

Пример

Дано описание функции f. Найти $x: \ f(x) = 0$.

Опр.

Функция $\varepsilon(n)$ называется *пренебрежимо малой*, если $\forall$ константы $c > 0$ существует константа $n_0 = n_0(c)$: $\varepsilon(n) < 1/n^c \ \ \forall n > n_0$.

Т.е. при $n \to \infty \ \forall poly(n) \ \varepsilon(n) = o(1/poly(n))$.

На практике, когда число n фиксировано, пользуются фиксированным ε, напр. $\varepsilon = 2^{-80}$.

Опр.

PPT - класс полиномиальных рандомизованных алгоритмов.
Probabilistic polinomial time.
$A \in \mathrm{PPT}$, если
1. Он имеет полиномиальное время работы,
может “подбрасывать монету” и принимать случайные решения; И
2. Если задача имеет два ответа ДА и НЕТ (т.е. это задача разрешимости, decision problem), то
алгоритм дает верный ответ с вероятностью больше 2/3.

Прим.:

можно требовать, чтобы алгоритм A давал правильный ответ с вероятностью более $1/2 + c$, где $c > 0$ - любая константа. Тогда алгоритм, который дает правильный ответ с вероятностью более фикс. $c' < 1$, получается из алгоритма A так. Повторим алгоритм A t раз и примем решения простым большинством по t результатам.

t не зависит от длины входа. $t = t(c, c')$, растет при $c \to 0$ и при $c' \to 1$. Для док-ва исп. неравенство Чебышева.

Опр.

P - класс задач, разрешимых детерминированными алгоритмами за полиномиальное время.

Опр.

PP - класс задач, разрешимых рандомизированными алгоритмами за полиномиальное время. Если задача имеет два ответа, то вероятность правильного ответа алгоритма - более $1/2$.

Опр.

$\mathrm{BPP} \subseteq \mathrm{PP}$. (Bounded probabilistic polynomial)
Это класс задач, которые разрешимы рандомизированными алгоритмами за полиномиальное время. При этом, если задача имеет два ответа ДА и НЕТ, то
алгоритм дает верный ответ с вероятностью больше константы 2/3.

Что дает BPP по сравнению с PP?
Пусть вероятность правильного ответа $c = 2/3$ нам не достаточна. Хотим достичь $c' < 1$.
Для этого алгоритм из BPP повторим t раз и
примем решение простым большинством по t результатам.
Здесь t не зависит от длины входа. $t = t(c, c')$, растет при $c \to 1/2$ и при $c' \to 1$.)

Пусть можно точно предсказать следующие значения ГПСЧ.
Т.е.
$\exists i,\ \exists A \in \mathrm{PT} : G(k)|_{1,..i} \xrightarrow{A} G(k)|_{i+1,...,n}$

Если злоумышленник знает $m|_{1,..i}$ и $c|_{1,..i}$,
тогда $G(k)|_{1,..i} = m|_{1,..i} \oplus c|_{1,..i}$ и находит $G(k)|_{i+1,...,n}$.

Если $\exists i, A : G(k)|_{1,..i} \xrightarrow{A} G(k)|_{i+1}$,
то дешифрование бит за битом.

Обозн.: $k \xleftarrow{R} K$ - случайный равновероятный выбор элемента k из множества K

Опр.

ГПСЧ $G: K \to \{0,1\}^n$ называется *предсказуемым*, если существует алгоритм $A \in \mathrm{PPT}$,
не пренебрежимо малая функция $\gamma(n)$ и
$\exists i(n),\ 0 \le i \le n-1$:
при $k \xleftarrow{R} K$ $P[A(G(k)|_{1,\ldots,i}) = G(k)|_{i+1}] > 1/2 + \gamma(n)$.

$P[\cdot]$ вычисляется по случайному выбору ключа k и алгоритму A.

Задача

Дать определение непредсказуемого ГПСЧ.

Задача

Пусть $G: K \to \{0,1\}^n : \forall k\, XOR(G(k)) = 0$. Является ли он предсказуемым?

Статистические тесты:

- тест на случайность: количество серий из 0 и серий из 1 длины k для разных k
- тест на случайность: количество фикс. шаблонов длины k
- тест на равномерность,
  проверка гипотезы о законе распределения с. в.:
  сравнить количество всевозможных пар битов; можно для троек и т.д. (тест по критерию Пирсона χ^2)
- тест на автокорреляцию, для любого сдвига она мала:
  пусть $s_i \in \{0, 1\}$ - символы последовательности.

$$C(t) = \frac{1}{N}\sum_{i=0}^{N-1}(2s_i - 1)(2s_{i+t} - 1) = \begin{cases} 1 & \text{if } t = 0 \\ < \beta/\sqrt{N} & \text{if } t \neq 0 \end{cases}$$

- тест Maurer'a: степень возможного сжатия посл-ти.

Обзор: Википедия, Тестирование псевдослучайных последовательностей.

Однако из стат. тестов не следует, что этот PRG непредсказуемый!

Пример

$X_n = frac((n + X_0)\sqrt{2})$ - равномерный на [0,1], т.к. $\sqrt{2}$ - иррациональное число, но предсказуемый.

Пример

LCG - linear congruent generator, и LFBR - linear feedback shift register. Статистически хорошие, но все их параметры злоум-к однозначно определяет по довольно короткой псевдослучайной посл-ти.

Пример

LCG в простейшем виде: $X_{n+1} = (aX_n + b)\ mod\ m$
параметры: a, b, m

Пусть m известно, a, b - нет. Если знаем X_1, X_2, X_3, решаем систему 2 линейных уравнений относительно a, b.

Найти m тоже легко:

http://security.stackexchange.com/questions/4268/cracking-a-linear-congruential-generator

Раздел 2 - Поточные шифры и генераторы ПСЧ

Опр.

алгоритм A назовем *оракулом* для ГПСЧ G, если он принимает на вход значение $G(k)$ и
выдает значение 0 или 1 со следующим смыслом:
$A(G(k)) = 0$, если A считает $G(k)$ не случайной последовательностью, $A(G(k)) = 1$ иначе.
(Или наоборот.)

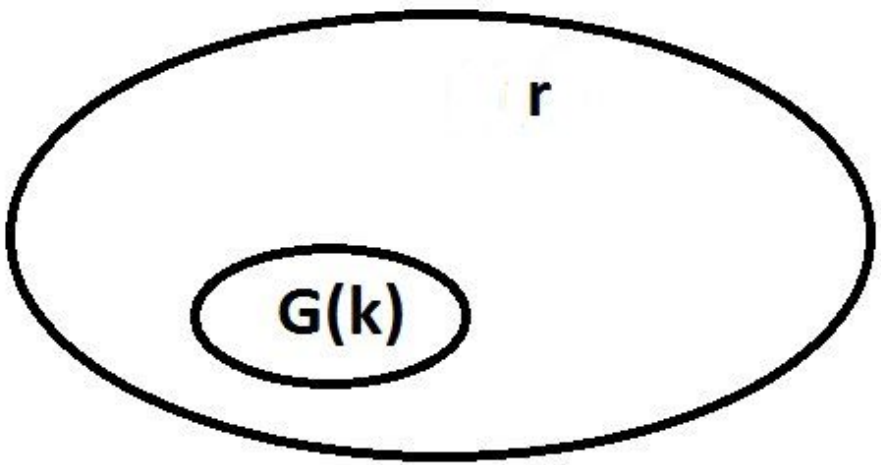
Опр.

ГПСЧ $G : K \to \{0,1\}^n$ называется *криптографически стойким*, если $\forall A \in \mathrm{PPT}$ при $k \xleftarrow{R} K$, $r \xleftarrow{R} \{0,1\}^n$ величина

$Adv[A, G] := |P[A(G(k)) = 1] - P[A(r) = 1]| < \varepsilon(n)$, где $\varepsilon(n)$ - пренебрежимо малая функция.

$P[\cdot]$ вычисляется по случайному выбору k и r и алгоритму A.

Т.е., применяя алгоритмы из PPT,
результат вычисления $G(k)$ при $k \xleftarrow{R} K$
не удается (не известен алгоритм)
отличить от результата выбора r при $r \xleftarrow{R} \{0,1\}^n$.



Задача

пусть $G : K \to \{0,1\}^n$ такой, что $(G(k))[1] = 0$ для $2/3$ ключей из K. Пусть оракул A выдает 1, если $x[1] = 0$, иначе 0. Найдите, чему равно
$Adv[A, G] = |P[A(G(k)) = 1] - P[A(r) = 1]|$.

Утверждение

Если генератор псевдослучайных чисел предсказуемый, то он не криптостойкий.

Док-во

G - предсказуемый $\Rightarrow \exists$ алгоритм , предсказывающий следующий бит с вероятностью более $1/2 + \gamma$, где γ не пренебрежимо малая величина.
Определим оракул B: если $A(X|_{1,\dots,i}) = X|_{i+1}$, выдать 1, иначе 0. Тогда
$P(B(r) = 1) = 1/2$
$P(B(G(k)) = 1) > 1/2 + \gamma$, значит
$Adv[B, G] > \gamma$, ч.т.д.

Следствие: криптостойкий ГПСЧ непредсказуемый.

Теорема 1 (Теорема Yao, 1982)

Пусть $G : K \to \{0,1\}^n$ - ГПСЧ. Пусть $\forall i \in \{0,..,n-1\}$ G не предсказуемый в позиции $i+1$. Тогда G - криптостойкий ГПСЧ.

Без доказательства.

Как построить генератор длинных посл-тей ПСЧ?

Псевдослучайные функции (ПСФ)

Пусть $K = \{0,1\}^k$, $X = \{0,1\}^n$, $Y = \{0,1\}^m$.

Опр.

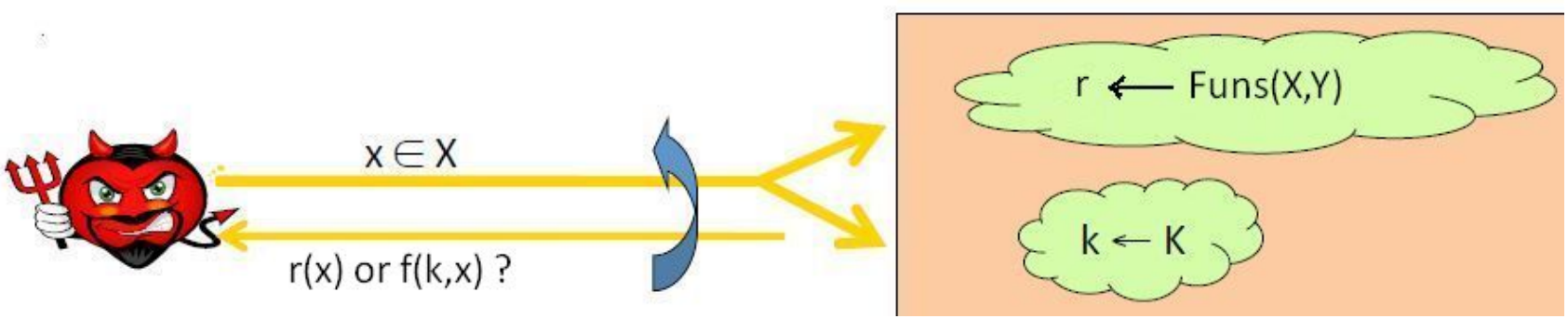
Функция $f : K \times X \to Y$ - это *псевдослучайная функция* [по второму аргументу], если
$f \in \mathrm{PT}$ и $\forall$ оракула $A \in \mathrm{PPT}$ при $k \xleftarrow{R} K$, $r \xleftarrow{R} \{\varphi : X \to Y\}$
величина
$Adv(A, f) = |P[A(f(k,\cdot)) = 1] - P[A(r(\cdot)) = 1]| < \varepsilon(n)$,
пренебрежимо малой функции.

Т.е. оракул из класса PPT не может отличить эту функцию от случайной.

$P[A(f(k,\cdot)) = 1]$:

Бросили монеты, получили k. Бросили монеты, получили r.
A вычисляет $f(k,\cdot)$ в $poly(n)$ числе точек x по своему выбору.
Выдает результат.
Этот результат усредняется по всем k и алгоритму A.

$P[A(r(\cdot)) = 1]$ - аналогично, в тех же точках x.



Задача

Пусть K, X, Y - конечные множества. Найти мощность множеств
$\{f(\cdot) : X \to Y\}$ и
$\{\text{ПСФ } f(k, \cdot) : X \to Y,\ k \in K\}$.

Задача

Пусть $f : K \times X \to \{0,1\}^{128}$ - псевдослучайная функция. Будет ли следующая функция псевдослучайной:

$$f'(k,x) = \begin{cases} 0^{128} & \text{если x=0} \\ F(k,x) & \text{иначе} \end{cases}$$

Построим криптографически стойкий ГПСЧ произвольной длины с помощью ПСФ.

Пусть $f : K \times \{0,1\}^n \to \{0,1\}^n$ - ПСФ. Тогда следующая конструкция с использованием ПСФ в режиме счетчика будет криптостойким ГПСЧ:

$G : K \to \{0,1\}^{nt}$
$G(k) = f(k,0)||f(k,1)||...||f(k,t-1)$

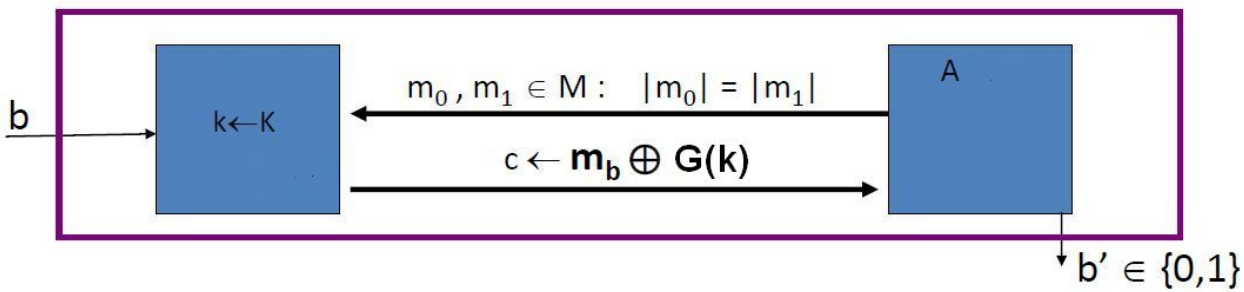
Шифр: $c = m \oplus G(k)$. Криптостойкий? В каком смысле?

Раздел 2 - Поточные шифры и генераторы ПСЧ

Семантическая стойкость шифра (semantic security) с одноразовым ключом

Эксперимент SS:

1. Система выбирает случайное значение бита $b \xleftarrow{R} \{0,1\}$. Оно секретное.
2. Злоумышленник выбирает длину n и отправляет два сообщения $m_0 \neq m_1$ длины n.
3. Система вычисляет $c = E(k, m_b)$ и отправляет его злоум-ку.
4. Злоум-к $A \in \mathrm{PPT}$ анализирует c и выдает результат - бит b'.
5. Если $b' = b$, A достиг успеха.



Один запрос, ответ.

Опр.

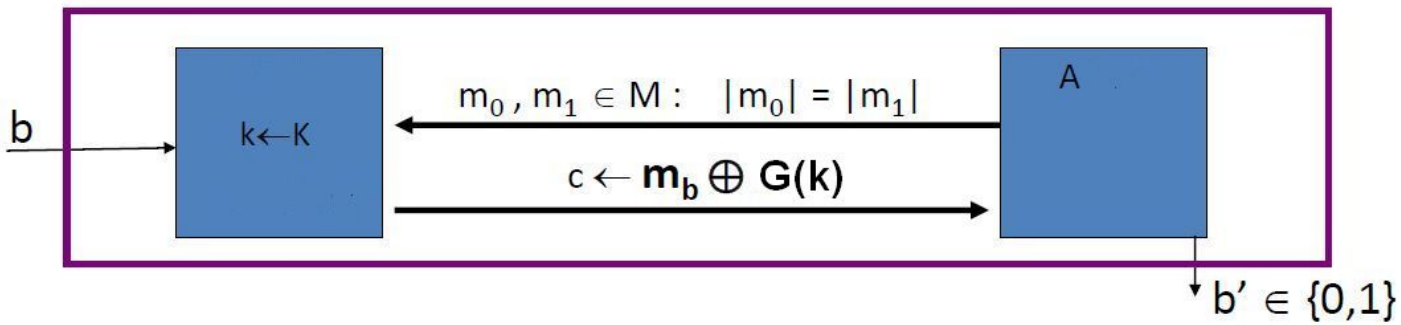
Шифр наз. *семантически стойким для одноразового ключа*, если в этом эксперименте вероятность успеха A не более $\frac{1}{2} + \varepsilon(n)$,
где $\varepsilon(n)$ - пренебр. малая,
вероятность вычисляется по случайным выборам: k, b, алгоритма A и шифрования E.

Эквивалентно:
$Adv_{SS}(A, E_{OTK}) = \left|P(b' = 1|b = 1) - P(b' = 1|b = 0)\right| < \varepsilon(n)$.

Т.е. семантическая стойкость шифра с одноразовым ключом: при однократном использовании ключа вероятность того, что “эффективный” злоум-к правильно укажет, ш/т какого из двух сообщений равной длины он получил, отличается от вероятности угадать (1/2) не более чем на пренебрежимо малую величину.

Пример

Пусть оракул A по шифротексту может точно определить значение бита открытого текста, например $m[1]$. Тогда строится эксперимент: $m_0[1] = 0, m_1[1] = 1$. Анализируем шифротекст...



$$Adv_{SS}(A, E_{OTK}) = \left|P(b' = 1|b = 1) - P(b' = 1|b = 0)\right| = |1 - 0| = 1$$

Пусть A может определить не бит $m[1]$, а значение некоторой функции от m. Тот же сценарий: m_0, m_1, где функция принимает разные значения.

Теорема 2

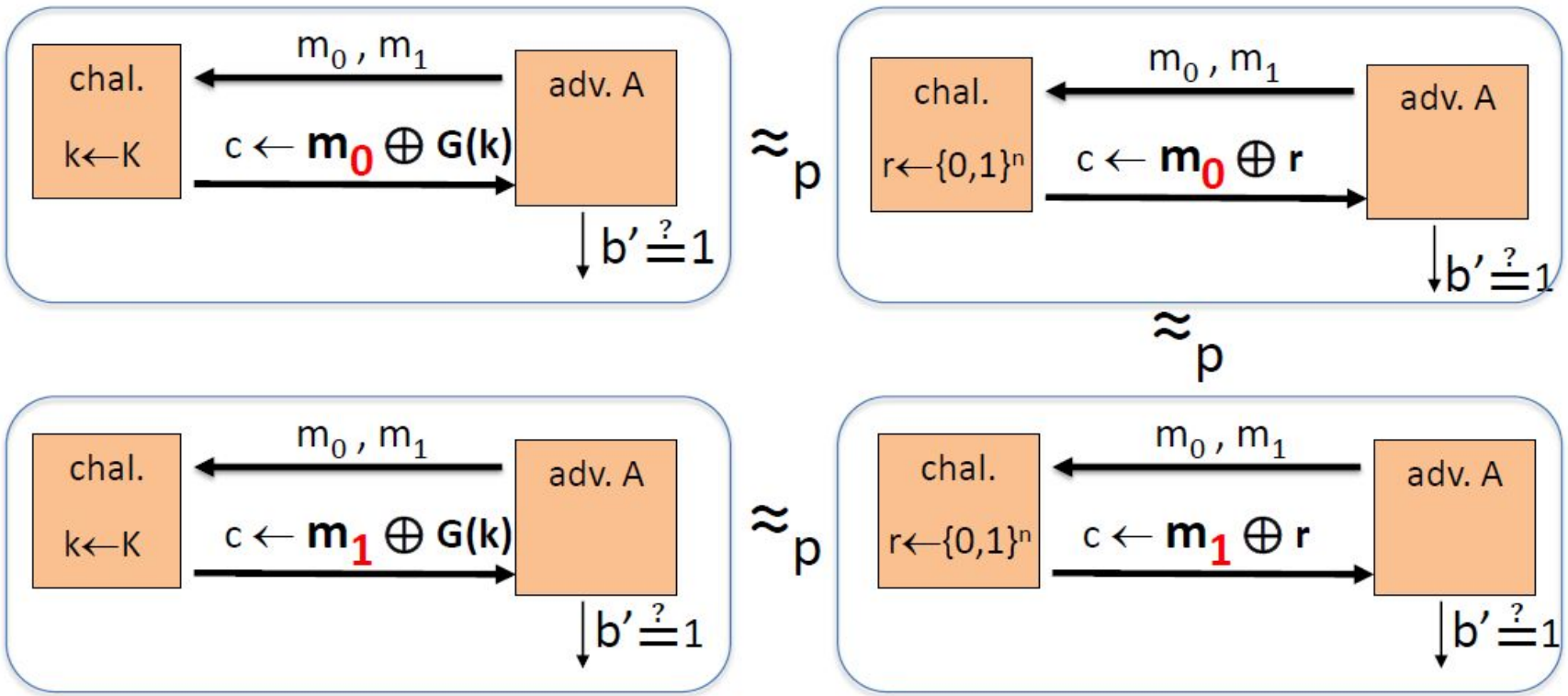
Пусть $G : K \to \{0,1\}^*$ - криптостойкий ГПСЧ. Тогда поточный шифр на основе $G(k)$ семантически стойкий при одноразовом ключе.

И $\forall$ алгоритма A, пытающегося взломать шифр,

$\exists$ алгоритм B атаки на ГПСЧ:

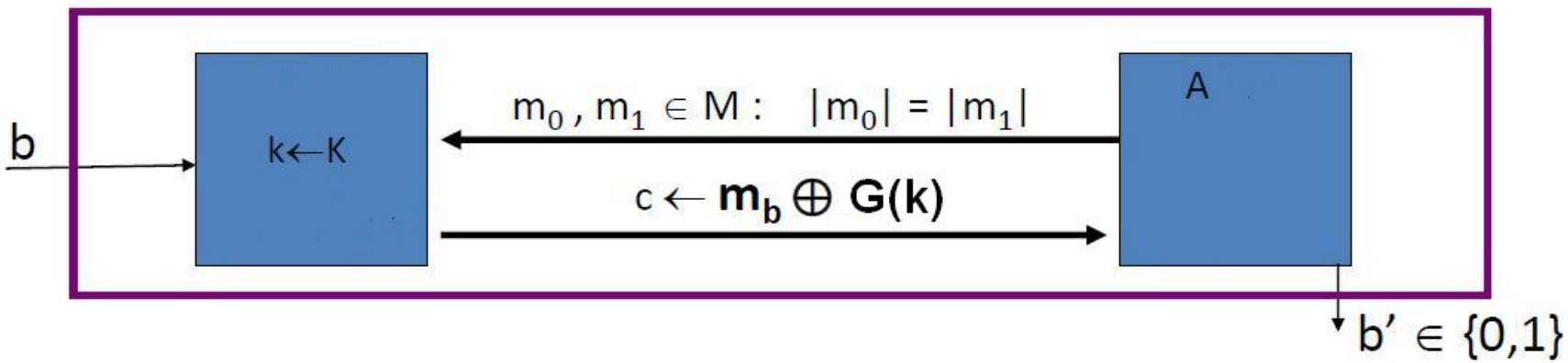
$Adv_{SS}(A, E) \leq 2 \cdot Adv_{PRG}(B, G)$.

Идея доказательства: $|m_0| = |m_1|$ и



Док-во

Цель: $Adv_{SS}(A, E) \leq 2 \cdot Adv_{PRG}(B, G) = 2\varepsilon(n)$.

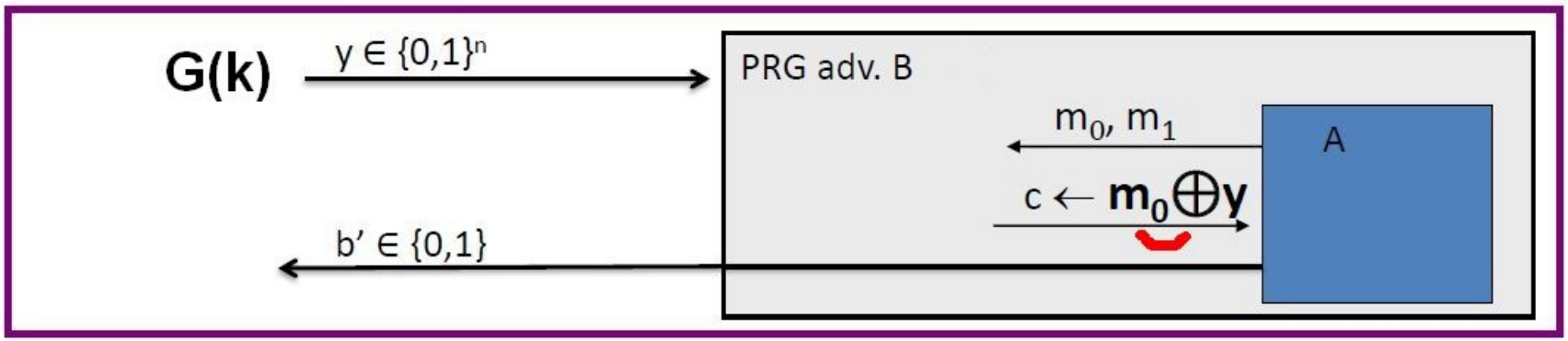


Обозн. $W_q :=$ событие: $b' = 1 | b = q$.
$Adv_{SS}(A, E) = \left| P_A(W_1) - P_A(W_0) \right|$.

Такая же атака на одноразовый блокнот.
Обозн. $R_q :=$ событие: $b' = 1 | b = q$ для нее.

Док-во (Продолжение)

Algorithm B:



$$Adv_{PRG}(B, G) = \left|P(B(r) = 1) - P(B(G(k)) = 1)\right| =$$

$$= \left|P_A(R_0) - P_A(W_0)\right|$$

Аналогично для $P(R_1), P(W_1)$ - подаем в A $m_1 \oplus y$.

Док-во (Продолжение)

$$Adv_{SS}(A, E) = \left|P_A(W_1) - P_A(W_0)\right| \leq$$

$$\leq \left|P_A(W_1) - P_A(R_1)\right| + \left|P_{OTP}(R_1) - P_{OTP}(R_0)\right| + \left|P_A(R_0) - P_A(W_0)\right| =$$

$$= 2 \cdot Adv_{PRG}(B, G)$$

Ч.т.д.

Некоторые итоги:

A. Стойкость шифров с одноразовым ключом:
A.1. Абсолютная стойкость к атаке с известным шифротекстом.
A.2. Семант. стойкость к атаке с известным шифротекстом.

B. -

C. Шифры:
C.1. Алфавитные.
C.2. Одноразовый блокнот.
C.3. Поточные шифры, ГПСЧ.

Литература к лекции

нет