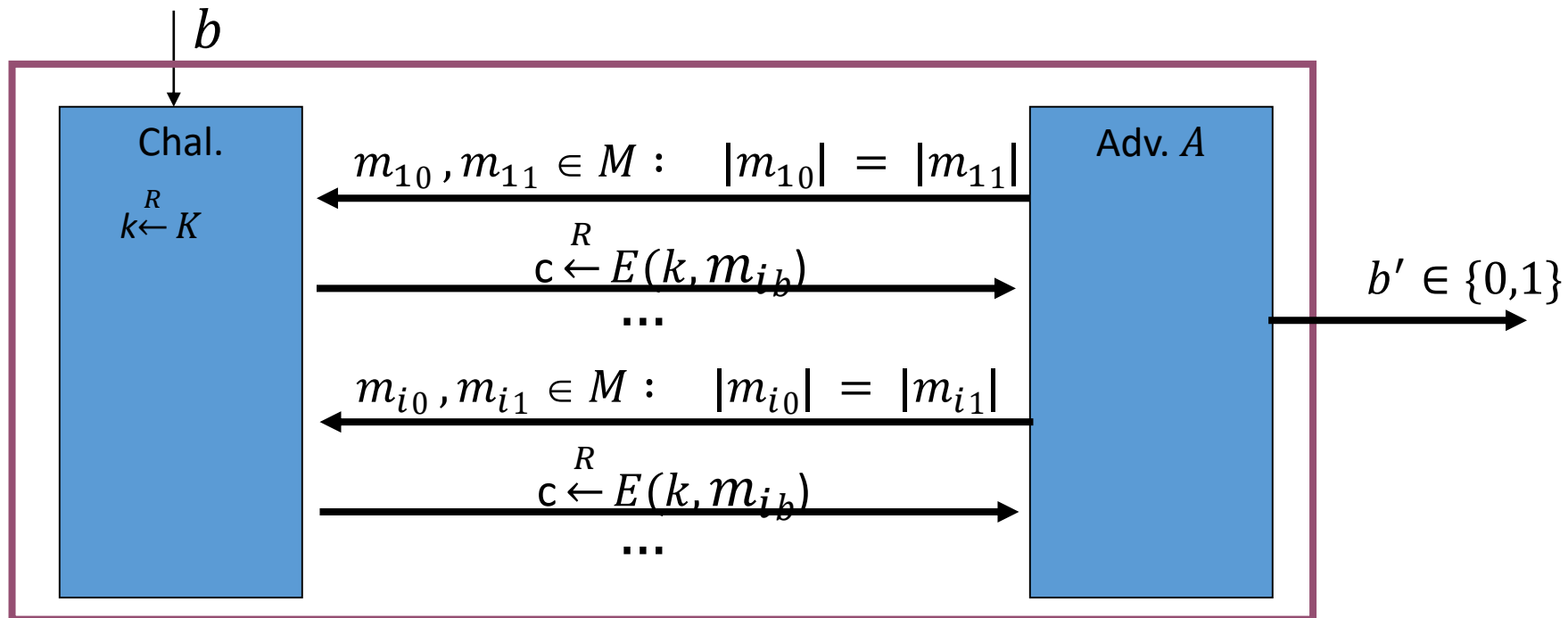


# Прикладная Криптография: Симметричные криптосистемы nonce CPA, det-CPA

Макаров Артём  
МИФИ 2018

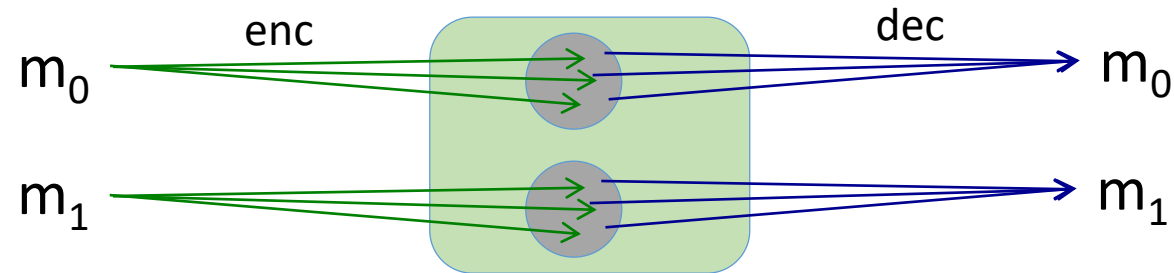
# CPA

- Шифр называется CPA стойким, если для любого противника  $A$  величина  $CPA_{adv}[A, E] = |\Pr[W_0] - \Pr[W_1]| \leq \epsilon$ ,  $\epsilon$  – пренебрежимо малая величина.
- Детерминированный шифр не может быть CPA стойким



# Вероятностное шифрование

- Как показано ранее, для СРА стойкости необходима «рандомизация» шифртекстов
- Подход 1 – рандомизация функции зашифрования



- Зашифрование одного и того же сообщения даст разные шифртексты
- Необходим внешний источник энтропии
- Шифртексты всегда длиннее открытых текстов, так как необходимо также передать энтропию, необходимую для восстановления открытого текста

# Вероятностное шифрование

- Подход 2 – использование уникальных, неповторяющихся величин (nonce)
- $m \rightarrow E(k, *, n) \rightarrow c \rightarrow D(k, *, m) \rightarrow m$
- Nonce должна быть уникально для каждого сообщения, пара (nonce, key) не должна повторяться при жизни ключа.
- В качестве nonce можно использовать счётчик или случайные величины
- Nonce может не пересылаться в явном виде, обе стороны могут синхронно обновлять его.
- Не любое использование nonce даёт стойкие схемы!

# CBC vs CTR

$$CPA_{adv}[A, E_{ctr}] \leq \frac{4Q^2l}{N} + 2 * PRF_{adv}[B, F]$$
$$CP_{adv}[A, E_{cbc}] \leq \frac{2Q^2l^2}{N} + 2 * BC_{adv}[B, E]$$

- CTR режим имеет большую стойкость для фиксированных параметров и блочного шифра
- CTR может использоваться в параллельном режиме, так как зашифрование блоков производит независимо
- Для коротких сообщений CTR может иметь длины шифртекстов значительно короче, чем CBC, так как нет необходимости в дополнении до длины блока.
- CTR использует только функцию зашифрования блочного шифра.
- **IV должны быть случайными!**

# Nonce based encryption

- Для всех рассмотренных ранее схем СРА шифрования длина результирующего шифртекста была больше длины открытых тестов из за добавления вектора инициализации.
- Длина вектора инициализации не зависит от длины сообщения
- Для больших сообщений не является проблемой (добавление 16 байт к мегабайту несущественно)
- Может являться проблемой для небольших шифртекстов, сравнимых с длиной блока (добавление 16 байт к сообщению длинны меньше 16 байт)
- Возможно ли уйти от случайных векторов инициализации?

# Nonce based encryption

- Первый подход – хранить некоторое состояние на стороне получателя и отправителя, которое явно или не явно синхронизируется перед процедурой шифрования. Затем обновлять эти значения после приёма-отправления сообщений.
  - Необходима полная синхронизация, при рассинхронизации – необходимо заново проводить процедуру синхронизации
- Второй подход – использование nonce. Вместо использования внутренних состояний использовать уникальные неповторяющиеся величины (nonce).

# Nonce based encryption

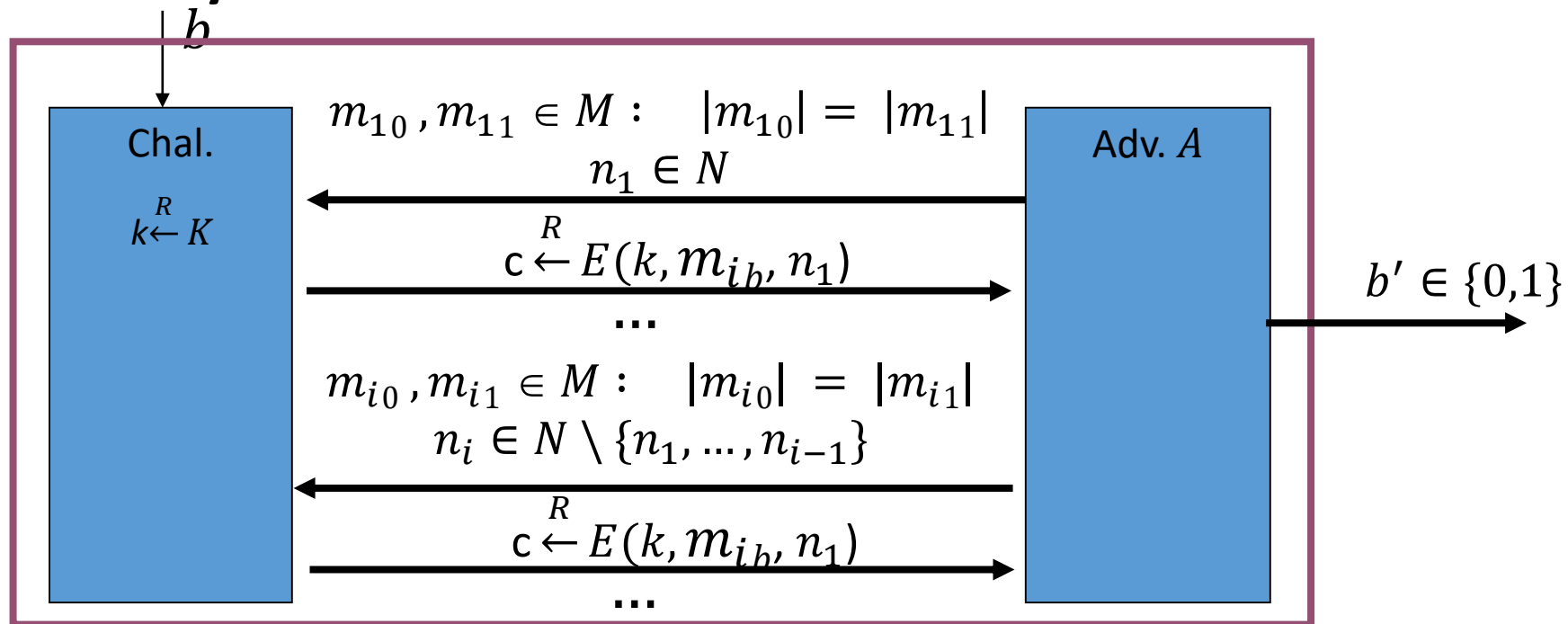
Для  $k \in K, t \in M, c \in C, n \in N$  шифром на основе nonce называется пара алгоритмов  $E = (E, D)$  на  $(K, M, C, N)$ :

- Зашифрование  $c = E(k, t, n)$
- Расшифрование  $t = D(k, c, n)$
- Корректность  $D(k, (E(k, t, n), n) = t$



# Nonce based CPA

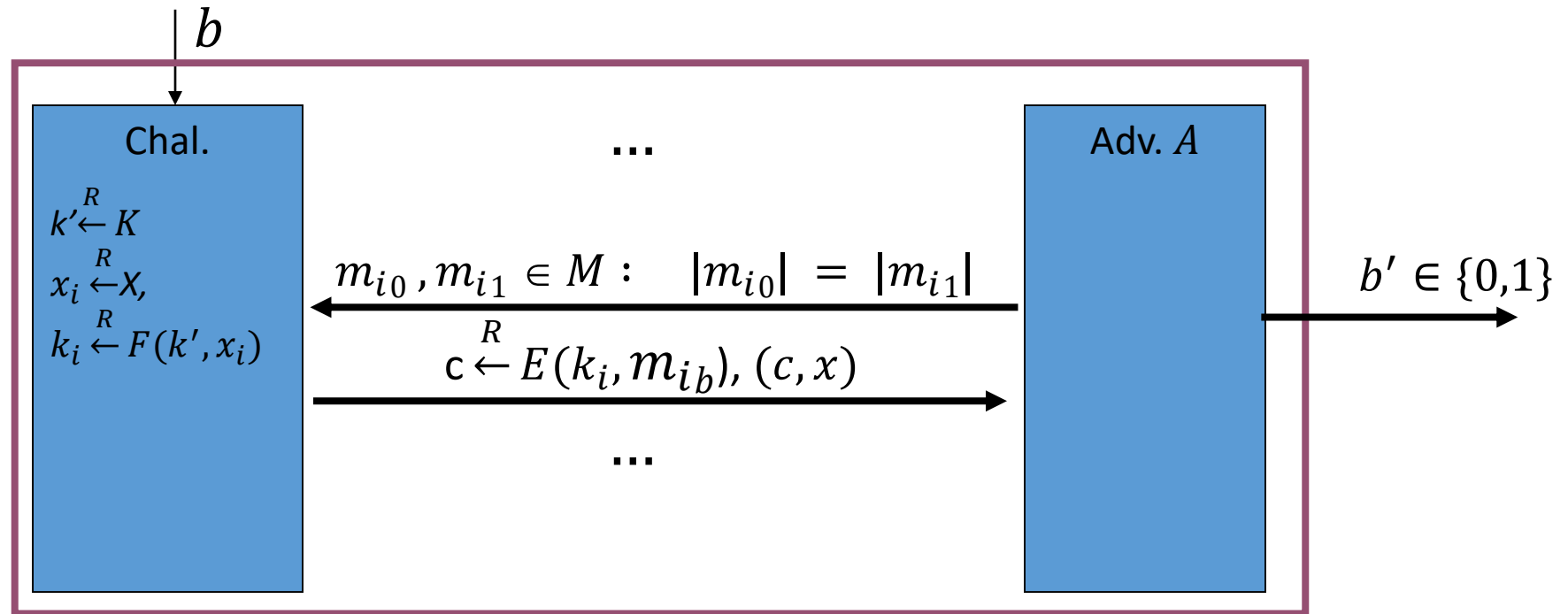
- Шифр на основе nonce называется nCPA стойким, если для любого противника  $A$  величина  $nCPA_{adv}[A, E] = |\Pr[W_0] - \Pr[W_1]| \leq \epsilon$ ,  $\epsilon$  – пренебрежимо малая величина.
- Заметим, что противник полностью выбирает nonce. Единственное требование – **уникальность**.



# Построение CPA шифров из семантически стойких шифров

- Пусть  $E = (E, D)$  – семантически стойкий шифр на  $(K, M, C)$ . Попробуем построить CPA стойкий шифр  $E'$  на  $(K', M, X \times C)$  используя PRF  $F$  на  $(K', X, K)$ .
- Ключом  $k'$  для  $E'$  будет ключ для PRF  $F$ . Для шифрования сообщения  $m$  выбирается случайный вход для PRF -  $x$ . Далее вычисляется ключ для  $E$   $k \leftarrow F(k', x)$ . Затем  $m$  шифруется с использованием ключа  $k$ :  $c \leftarrow E(k, m)$ . Шифр текстом является пара  $c' = (c, x)$ .
- $E(k', m) = [x \xleftarrow{R} X, k \leftarrow F(k', x), c \leftarrow E(k, m), \text{output}(x, c)]$
- $D(k', c') = [k \leftarrow F(k', x), m \leftarrow D(k, c), \text{output } m]$
- Называется – **гибридная конструкция**.

# Игра на СРА стойкость гибридной конструкции



# Построение CPA шифров из семантически стойких шифров

**Теорема 7.1.** Если  $F$  – стойкая PRF,  $E$  – семантически стойкий шифр,  $N = |X|$  - суперполиномиальная, то введённый ранее шифр  $E'$  - CPA стойкий шифр. В частности для любого противника в CPA игре, делающим не более  $Q$  запросов к претенденту существует противник  $B_F$  в игре на стойкость PRF и противник  $B_E$  в игре на семантическую стойкость, причём

$$CPA_{adv}[A, E'] \leq \frac{Q^2}{N} + 2 * PRF_{adv}[B_F, F] + Q * SS_{adv}[B_E, E]$$

# Гибридная конструкция на основе nonce

Модифицируем гибридную конструкцию, заменив случайный элемент  $x \in X$  на nonce.

Пусть  $E = (E, D)$  – семантически стойкий шифр на  $(K, M, C)$ .

Для ключа  $k' \in K, m \in M, c \in C, x \in X$  определим  $E'(k', m, x) = E(k, m), k = F(k', x)$

- $E(k', m) = [x \xleftarrow{R} X, k \leftarrow F(k', x), c \leftarrow E(k, m), \text{output}(x, c)]$
- $D(k', c') = [k \leftarrow F(k', x), m \leftarrow D(k, c), \text{output } m]$

# Детерминированная гибридная конструкция

**Теорема 8.1.** Если  $F$  – стойкая PRF,  $E$  – семантически стойкий шифр,  $N = |X|$  – суперполиномиальная, то введённый ранее шифр  $E'$  – CPA стойкий шифр. В частности для любого противника в nCPA игре, делающим не более  $Q$  запросов к претенденту существует противник  $B_F$  в игре на стойкость PRF и противник  $B_E$  в игре на семантическую стойкость, причём

$$nCPA_{adv}[A, E'] \leq 2 * PRF_{adv}[B_F, F] + Q * SS_{adv}[B_E, E]$$

▷ Аналогично **Теореме 7.1**, без необходимости добавления слагаемого  $Q^2/N$ , т.к. коллизии не возможно из за требования уникальности nonce◁

# Рандомизированный CTR режим

Рассмотрим ещё один способ построения – на основе CTR режима.

Пусть  $F$  PRF на  $(K, X, Y)$ . Пусть  $X = \{0, \dots, N - 1\}$ ,  $Y = \{0, 1\}^n$ . Для полиномиально ограниченной величины  $l \geq 1$  определим шифр  $E = (E, D)$  на  $(K, Y^{\leq l}, X \times Y^{\leq l})$  следующим образом:

Для  $k \in K, m \in Y^{\leq l}, v = |m|, = |c|, \quad c' = (x, c) \in X \times Y^{\leq l}$

$E(k, m) :=$

$x \xleftarrow{R} \mathcal{X}$

compute  $c \in \mathcal{Y}^v$  as follows:

for  $j \leftarrow 0$  to  $v - 1$  do

$c[j] \leftarrow F(k, x + j \bmod N) \oplus m[j]$

output  $(x, c)$ ;

$D(k, c') :=$

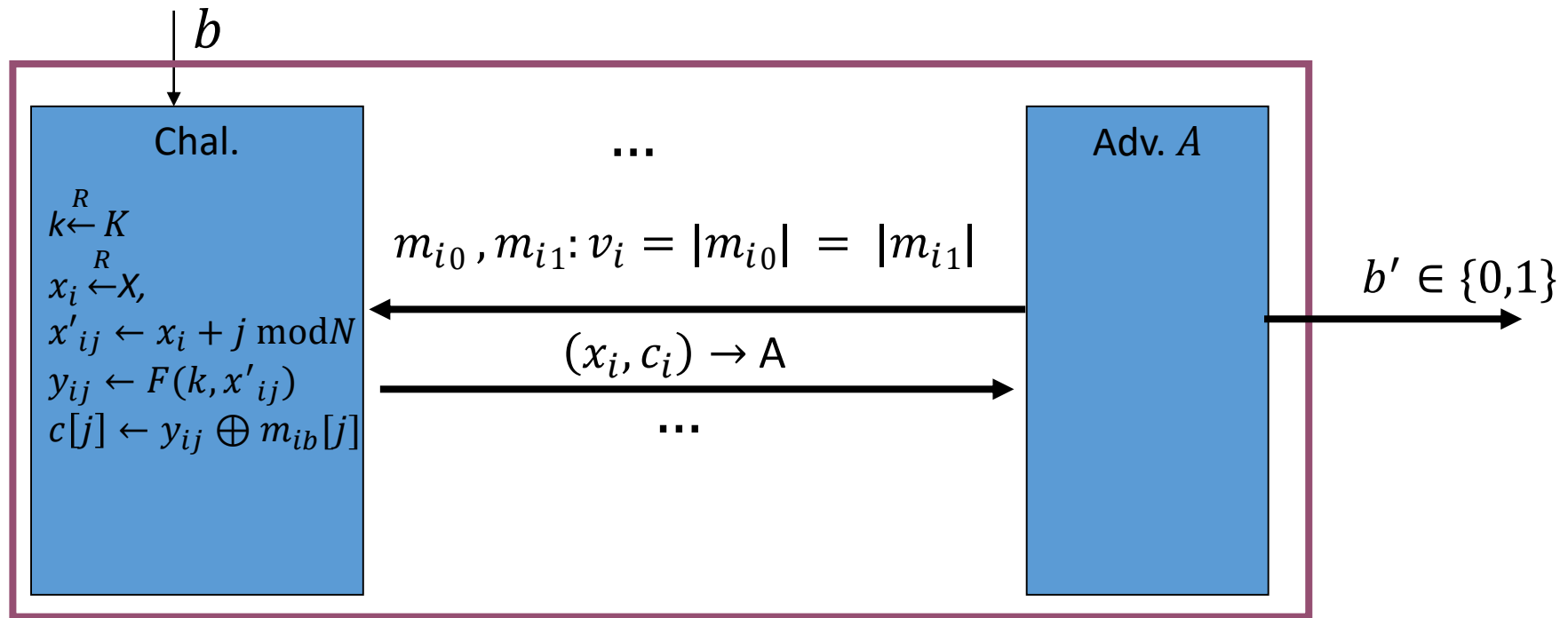
compute  $m \in \mathcal{Y}^v$  as follows:

for  $j \leftarrow 0$  to  $v - 1$  do

$m[j] \leftarrow F(k, x + j \bmod N) \oplus c[j]$

output  $m$ .

# Игра на СРА стойкость рандомизированного CTR режима





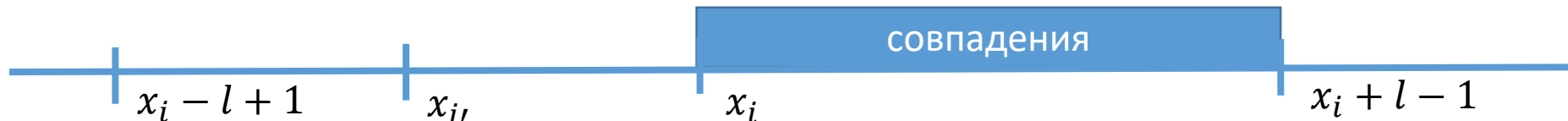
# Рандомизированный CRT режим

**Теорема 7.2.** Если  $F$  – стойкая PRF,  $N$  - суперполиномиальная,  $l$  – полиномиально ограниченная, то введённый ранее шифр  $E$  - CPA стойкий шифр. В частности для любого противника в CPA игре, делающим не более  $Q$  запросов к претенденту существует противник  $B$  в игре на стойкость PRF причём

$$CPA_{adv}[A, E'] \leq \frac{4Q^2l}{N} + 2 * PRF_{adv}[B, F]$$

# Nonce based CTR

- Можно ли построить CTR режим, заменив случайный элемент на nonce?
- Нет! В отличие от гибридной конструкции, где нам была важна уникальность nonce, здесь нам важна не только уникальность «начальных состояний», но и уникальность «отрезков». (См **лемму** из **Теоремы 7.2**).
- Иными словами, если заменить  $x_i \in X$  на nonce, то противник может выбрать такие  $x_i \neq x_{i'}: \{x_i, \dots, x_i + l - 1\} \cap \{x_{i'}, \dots, x_{i'} + l - 1\} \neq \emptyset$ , т.е. могут совпасть счётчики на каком то блоке для различных сообщений => имеем двухразовый блокнот.



# Nonce based CTR

- Введём nonce по другому. Пусть  $l|N$ . Пусть  $n \in \{0, \dots, N/l - 1\}$  – nonce,  $x = nl$ . Т.е. на вход PRF подаётся не nonce, а nonce умноженная на максимально допустимую длину сообщения в блоках.
- Т.е. два различных nonce  $n_1$  и  $n_2$  дают два входа для PRF  $x_1 = n_1l, x_2 = n_2l$  в интервалах  $\{x_1, \dots, x_1 + l - 1\}$  и  $\{x_2, \dots, x_2 + l - 1\}$ , которые не пересекаются.

# Рандомизированный CRT режим

**Теорема 8.2.** Если  $F$  – стойкая PRF,  $N$  – суперполиномиальная,  $l$  – полиномиально ограниченная, то введённый ранее шифр  $E$  – CPA стойкий шифр. В частности для любого противника в nCPA игре, делающим не более  $Q$  запросов к претенденту существует противник  $B$  в игре на стойкость PRF причём

$$\text{nCPA}_{adv}[A, E'] \leq 2 * \text{PRF}_{adv}[B, F]$$

▷ Аналогично **Теореме 7.2**, без необходимости добавления слагаемого  $\frac{4Q^2l}{N}$ , т.к. коллизии не возможно из за требования уникальности nonce◁

# CBC

Пусть  $E = (E, D)$  блочный шифр на  $(K, X)$  где  $X = \{0,1\}^n$ ,  $N = |X| = 2^n$ .  
Для полиномиально ограниченной величины  $l \geq 1$  определим шифр  $E = (E', D')$  на  $(K, X^{\leq l}, X^{\leq l+1} \setminus X^0)$ . Зашифрование и расшифрование определены следующим образом:

Для  $k \in K, m \in M, v = |m| = |c| - 1$

$E'(k, m) :=$

compute  $c \in \mathcal{X}^{v+1}$  as follows:

$c[0] \xleftarrow{\mathcal{R}} \mathcal{X}$

for  $j \leftarrow 0$  to  $v - 1$  do

$c[j + 1] \leftarrow E(k, c[j] \oplus m[j])$

output  $c$ ;

$D'(k, c) :=$

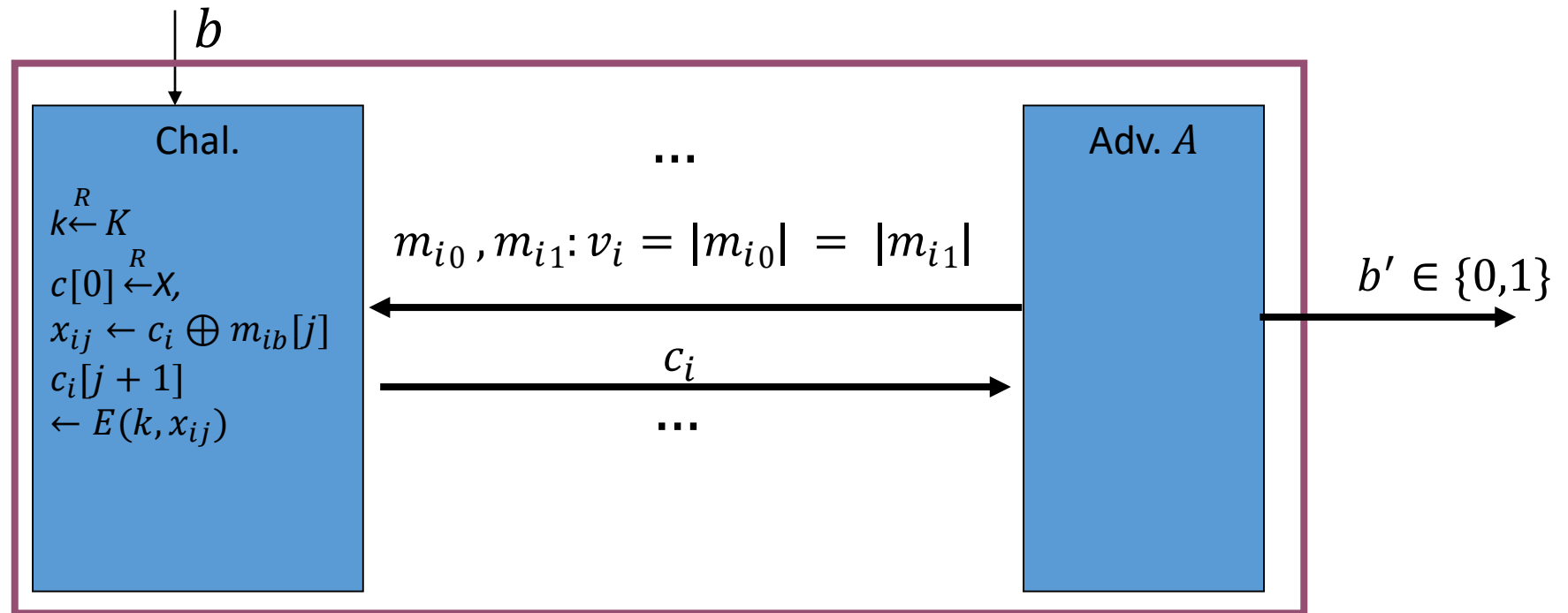
compute  $m \in \mathcal{X}^v$  as follows:

for  $j \leftarrow 0$  to  $v - 1$  do

$m[j] \leftarrow D(k, c[j + 1]) \oplus c[j]$

output  $m$ .

# Игра на СРА стойкость СВС



# CBC

**Теорема 7.3.** Пусть  $E = (E, D)$  – семантически стойкий шифр на  $(K, C)$ ,  $N = |X|$  - суперполиномиальная,  $l \geq 1$  – полиномиально ограниченная. Тогда введенный ранее CBC шифр является CPA стойким, причём для любого противника  $A$  в игре на CPA стойкость, делающим не более  $Q$  запросов к оракулу, существует противник  $B$  в игре на стойкость блочных шифров, при чём

$$CPA_{adv}[A, E'] \leq \frac{2Q^2 l^2}{N} + 2 * BC_{adv}[B, E]$$

# Nonce based CBC

- Можно ли построить CBC режим, заменив случайный элемент на nonce?
- Нет! Противник может сделать 2 запроса  $(m_{10}, m_{11}, n_1)$ ,  $(m_{20}, m_{21}, n_2)$ :  $m_{10} = n_1 \neq n_2 = m_{20}, m_{11} = m_{21}$ . В эксперименте 0 шифртексты будут одинаковые, в эксперименте 1 – разными.



# Nonce based CBC

- Идея – заменить случайный IV на псевдослучайный, полученный из nonce с помощью PRF.
- Пусть  $F$  – PRF на  $(K', N', X)$ , где  $X$  – множество блоков блочного шифра  $E = (E, D)$ , отпрядённого на  $(K, X)$ .
- Ключом является элемент из множества  $K \times K'$ , алгоритм зашифрования и расшифрования отличаются от CBC только в получении  $n[0] = F(k', n)$ .

# CBC

**Теорема 8.3.** Пусть  $E = (E, D)$  – семантически стойкий шифр на  $(K, C)$ ,  $N = |X|$  - суперполиномиальная,  $l \geq 1$  – полиномиально ограниченная. Тогда введенный ранее CBC шифр является CPA стойким, причём для любого противника  $A$  в игре на  $n$ CPA стойкость, делающим не более  $Q$  запросов к оракулу, существует противник  $B$  в игре на стойкость блочных шифров, и  $B_F$  в игре на стойкость PRF, при чём

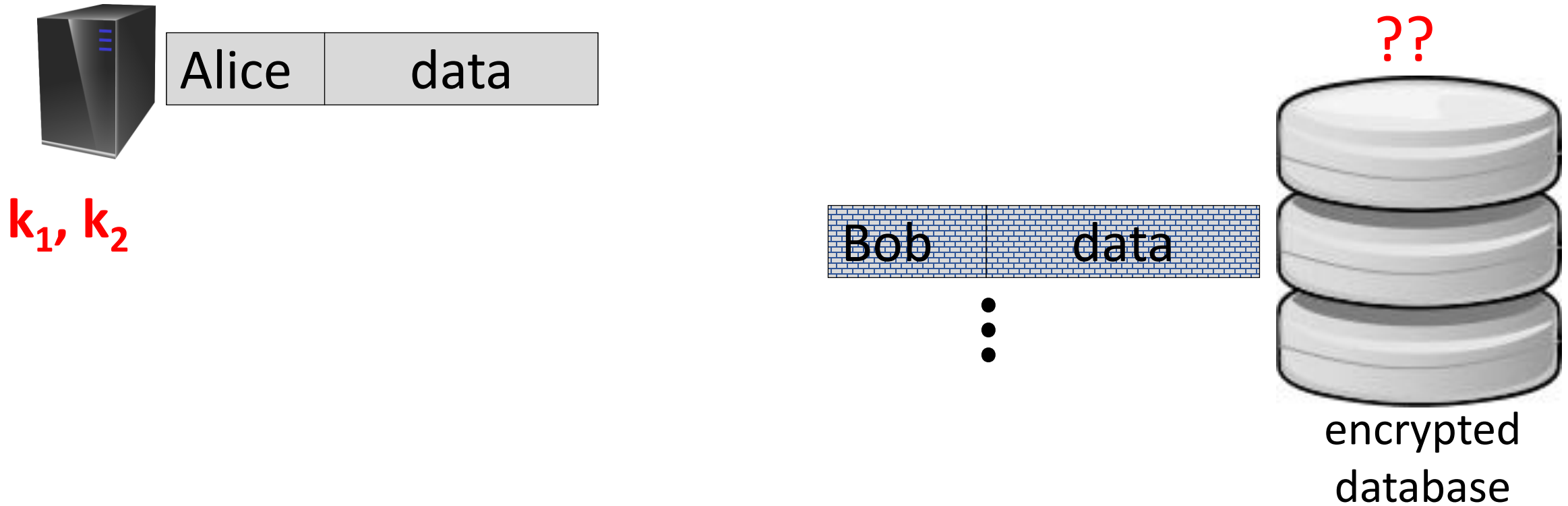
$$nCPA_{adv}[A, E'] \leq \frac{2Q^2l^2}{N} + 2 * BC_{adv}[B, E] + 2 * PRF[B_F, E]$$

▷Аналогично **Теореме 7.3**, но с учётом использовать не только блочного шифра, но и PRF◁

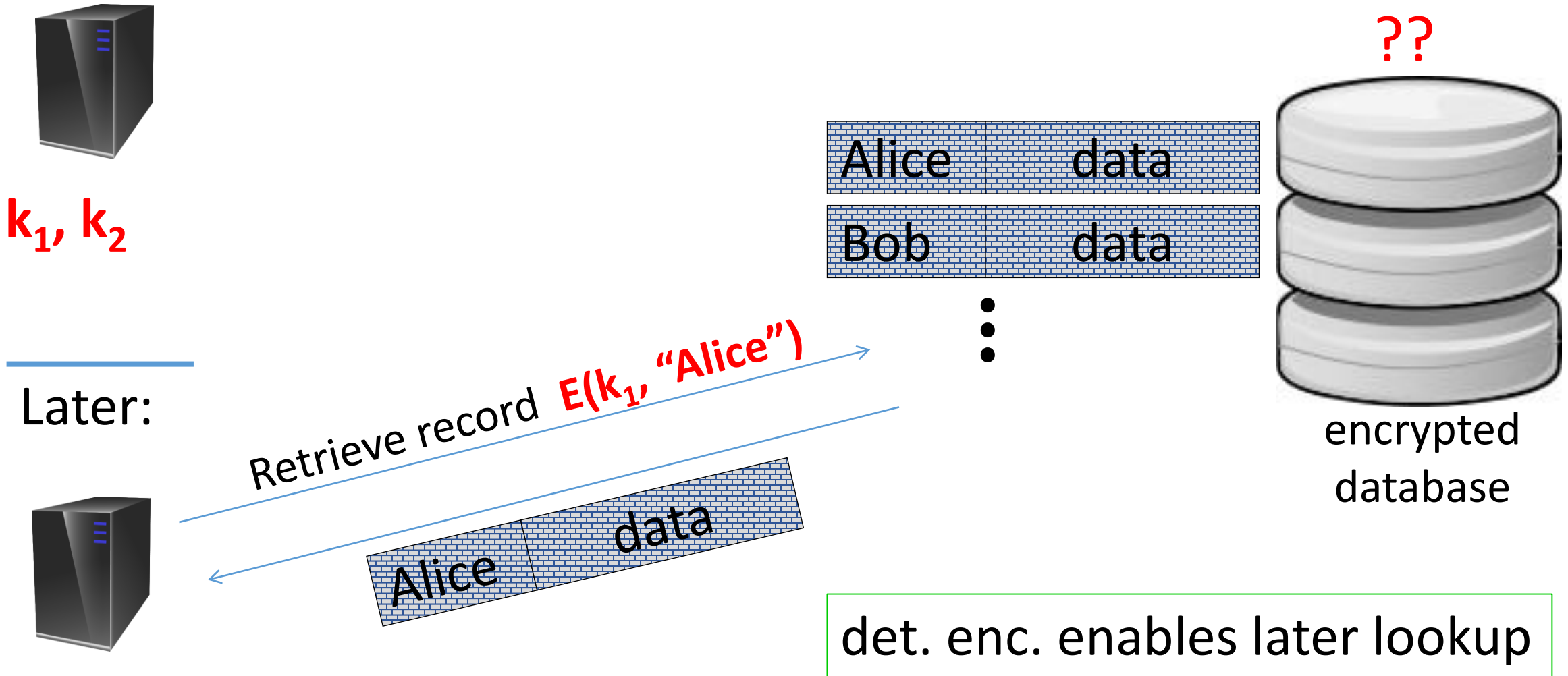
# Детерминированное шифрование

- Рассмотрим пример – хранение файлов на удалённом сервере.
- Пользователь отправляет зашифрованный файл на сервер, приписывая заголовок. Сервер записывает шифртекст без расшифровки
- Для получения файла из базы данных пользователь отправляет зашифрованный (тем же ключом) заголовок и получает шифртекст, который потом расшифровывает.
- Данная схема возможна только при детерминированном шифровании

# The need for det. Encryption (no nonce)



# The need for det. Encryption (no nonce)

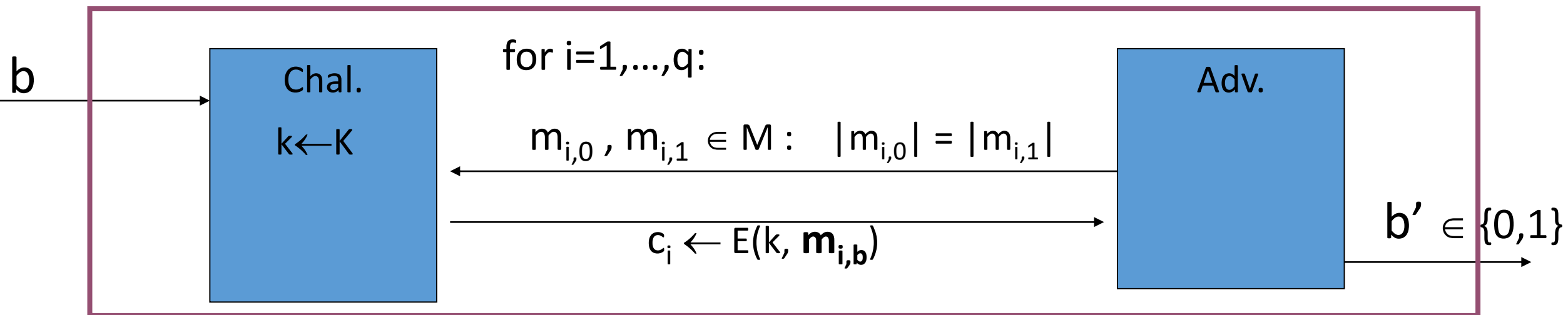


# Детерминированное шифрование

- Проблема – при детерминированном шифровании противник может проверять заголовки на равенство, т.к. одинаковые заголовки дают одинаковые зашифрования заголовков.
- Аналогично для шифртекстов. Если множество шифртекстов мало (например шифруются только слова, длины не более 6 символов), и распределение неравномерное, противник может провести частотный анализ и полностью расшифровать все шифртексты.
- Нужно новое определение. Основная идея – сообщения должны быть уникальными для фиксированного ключа.
  - Уникальные идентификаторы, которые не повторяются (номер в очереди, номер передаваемого пакета, уникальный для сессии id пользователя итд)
  - Сообщения выбранные случайно из большого множества (например ключи)

# Deterministic CPA security

Пусть  $E = (E, D)$  шифр на  $(K, M, C)$ . Введём игру на CPA стойкость, в которой противник запрашивает только уникальные сообщения, т.е.  $\mathbf{m}_{1,0}, \dots, \mathbf{m}_{q,0}$  и  $\mathbf{m}_{1,1}, \dots, \mathbf{m}_{q,1}$  различны.

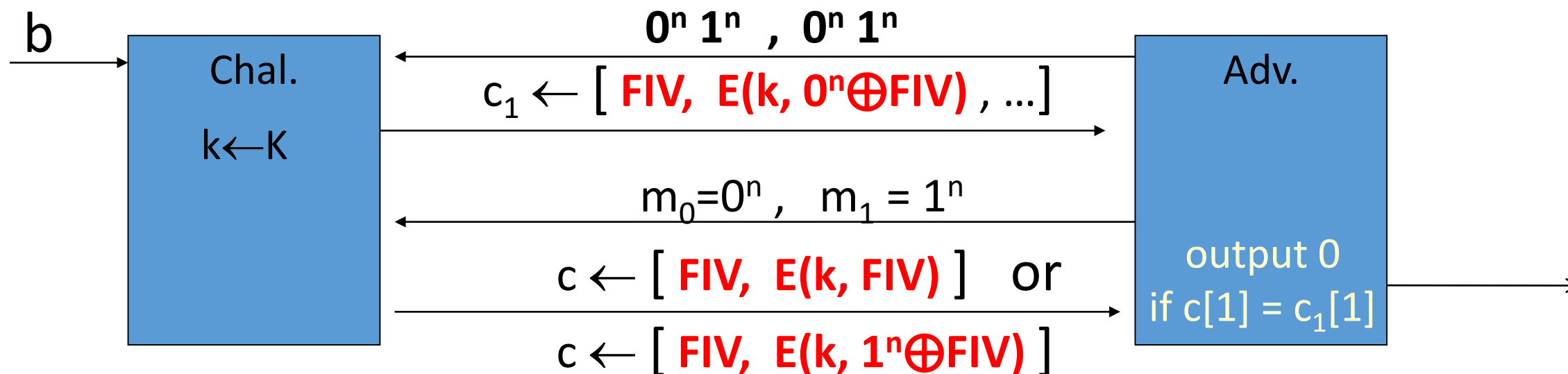


$E = (E, D)$ , определённый на  $(K, M, C)$ , называется детерминированно CPA стойким детерминированным, если  $\forall A$ :  $A$  – эффективный алгоритм в игре на стойкость Deterministic CPA величина  $dCPA_{adv}[A, E] = |\Pr[W_0] - \Pr[W_1]| \leq \epsilon$ , где  $\epsilon$  – пренебрежимо малая величина.

# Фиксированный IV в CBC

**Фиксированный IV в CBC не даёт det-CRA стойкость!**

Пусть  $E: K \times \{0,1\}^n \rightarrow \{0,1\}^n$  стойкая  $PRP$  в CBC

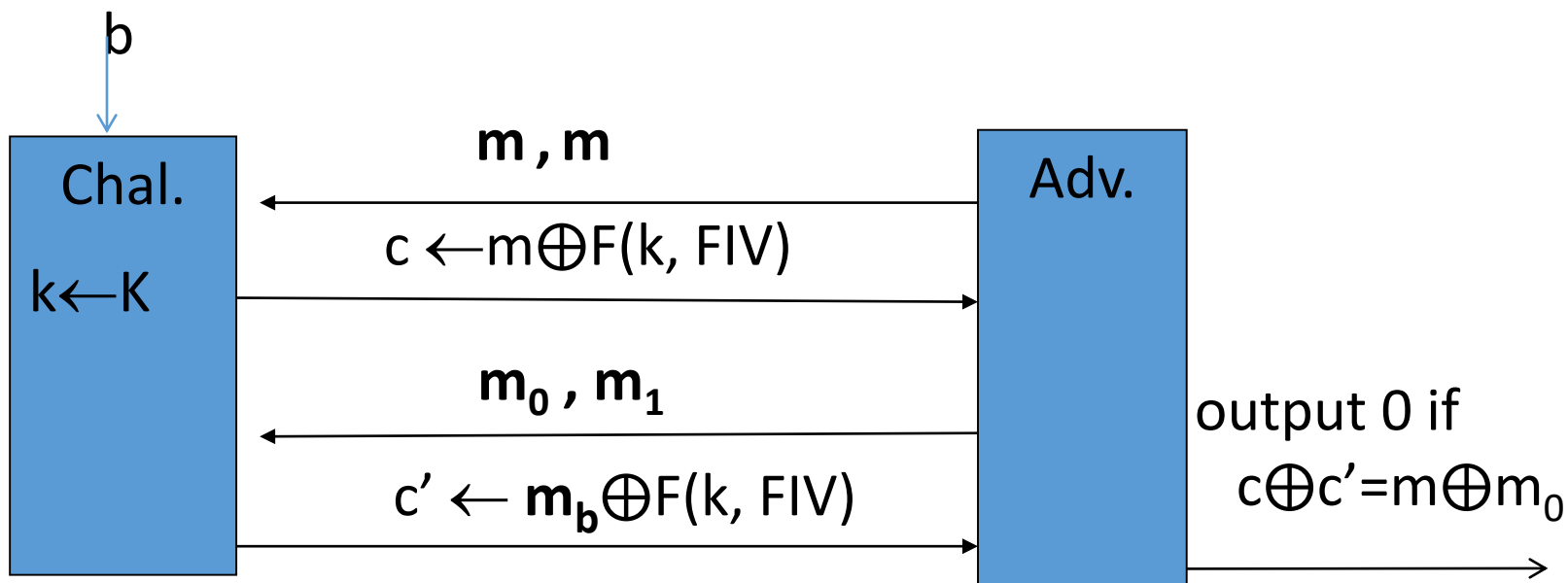
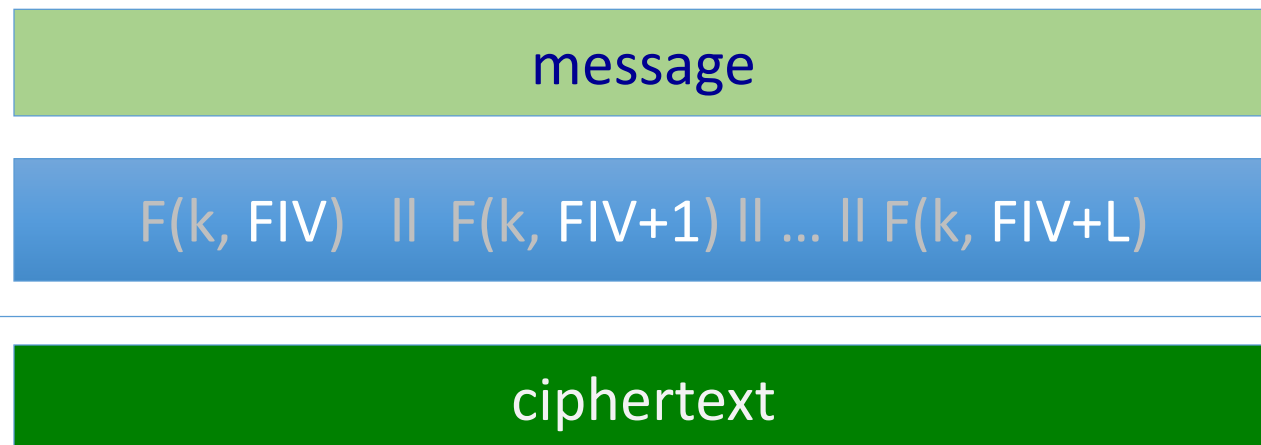




# Фиксированный IV в CTR

Фиксированный IV в CBC не даёт det-CRA стойкость!

Пусть  $F: K \times \{0,1\}^n \rightarrow \{0,1\}^n$   
стойкая PRF в CTR



# Синтетический IV

Пусть  $E = (E, D)$  – CPA стойкий шифр на  $(K, M, C)$ ,  $E = (k, m; r)$  – функция зашифрования, использующая случайный вход  $r \in_R R$ . Пусть  $F$  – стойкая PRF на  $(K', M, R)$ . Тогда детерминированный шифр  $E' = (E', D')$  на  $(K \times K', M, C)$ :

$$\begin{aligned} E'((k, k'), m) &= E(k, m; F(k', m)), \\ D'((k, k'), c) &= D(k, c) \end{aligned}$$

Называется детерминированным шифром, использующем синтетический IV.

NB: конструкция похожа на использование nonce в CTR и CBC, но случайность заменяется не шифрованием уникального nonce, а шифрованием уникального сообщения (сообщения уникальны для det-CPA).

**Теорема 8.4.** Описанный выше шифр является det-CPA стойким.

▷ без доказательства, или доказать самим ◁

# Выводы

- Шифры решают задачу конфиденциальности информации при пассивном противнике (противнике не влияющем на передаваемые сообщения)
- Абсолютная стойкость – достижимая, но не удобная для построения шифров модель
- Ослабленная версия абсолютной стойкости – семантическая стойкость (одноразовая семантическая стойкость) – используется для построения и анализа шифров при однократном использовании ключа
- При шифровании нескольких сообщений используется СРА стойкость (многократная семантическая стойкость), позволяющая противнику получать зашифрования нескольких сообщений на одном ключе

# Выводы

- Основные примитивы – псевдослучайные генераторы, поточные шифры, блочные шифры.
- Для построения семантических и СРА стойких шифров из блочных шифров используют режимы шифрования.
- При использовании режимов шифрования, требующих случайный IV, он должен быть случайным!
- Шифры не должны использоваться для обеспечения целостности или аутентичности!