О свойстве конфиденциальности AEAD-режима MGM

Л. Р. Ахметзянова

***Аннотация*—В настоящей работе исследуется стойкость режима аутентифицированного шифрования MGM к угрозе нарушения конфиденциальности. Данный режим впервые был представлен на конференции CTCrypt в 2017 году и позже был стандартизирован в Российской Федерации. Шифрование открытых текстов в режиме MGM происходит путем выработки маскирующей последовательности в режиме счетчика. Основным элементом процедуры формирования имитовставки является мультилинейная функция с секретными коэффициентами, которые аналогично процедуре шифрования вычисляются в режиме счетчика. Данная конструкция позволяет обеспечить такие эксплуатационные свойства как распараллеливание процедур шифрования и выработки имитовставки, возможность использования предвычислений и «online»-обработки данных.**

В докладе, представленном на конференции, были приведены принципы построения режима MGM с точки зрения обеспечения защиты информации. В рамках настоящей работы проводится анализ свойства конфиденциальности режима MGM в парадигме доказуемой стойкости. Была доказана нижняя оценка уровня стойкости в формальной модели противника IND-CPA как функция от параметров режима и объема доступной противнику информации. Доказанная оценка демонстрирует, что исследуемый режим обеспечивает стандартный для AEAD-режимов уровень стойкости в части конфиденциальности (в предположении стойкости базового блочного шифра).

***Ключевые слова*—MGM, AEAD-режим, модель противника, оценки стойкости, конфиденциальность**

I. Введение

Схемы аутентифицированного шифрования с ассоциированными данными (или AEAD-схемы), обеспечивающие конфиденциальность и целостность информации, используются во многих прикладных системах, таких как службы электронной почты, системы асинхронной передачи сообщений (мессенджеры), хранение данных на носителях. Также данные схемы являются обязательными для использования в протоколе TLS 1.3 [1].

Основными преимуществами AEAD-схем являются прозрачность и простота их использования. Действительно, явно определенная конструкция и унифицированный интерфейс способствуют их корректной реализации и прозрачному встраиванию в высокоуровневые схемы. Кроме того, использование одного ключа для обеспечения как конфиденциальности, так и целостности устраняет необходимость в дополнительной диверсификации ключей. Данная операция обычно используемой для создания пары независимых ключей для схем, являющихся композицией схемы симметричного шифрования и схемы выработки имитовставки. Отсутствие диверсификации позволяет утверждать (в дополнение к очевидному улучшению производительности), что AEAD-схемы обеспечивают более гарантированную безопасность по сравнению с универсальными композициями, поскольку их безопасность обеспечивается при меньшем количестве предположений о стойкости используемых примитивов.

AEAD-режим MGM впервые был предложен на конференции CTCrypt в 2017 (см. [2]). Процедура шифрования открытых текстов в режиме MGM аналогична процедуре в режиме CTR2 [3]. Основным элементом процедуры формирования имитовставки в режиме MGM является мультилинейная функция, секретные коэффициенты которой вырабатываются способом, аналогичным процедуре выработки секретных масок для шифрования открытых текстов.

В докладе, представленном на конференции, были приведены принципы построения режима MGM с точки зрения обеспечения защиты информации. Исследование свойств, соответствующих формальной модели противника IND-CPA (конфиденциальность), впервые проведено в рамках настоящей работы.

В настоящей работе доказывается содержательная верхняя оценка преимуществ всех возможных противников, определенных рассматриваемой моделью, как функции от размера блока n, количества обрабатываемых сообщений q, максимальной длины открытых текстов и ассоциированных данных в блоках l (Теорема V.1). Полученные в настоящей работе результаты о стойкости режима MGM использовались при стандартизации данного режима в Российской Федерации [4].

II. Обозначения

Обозначим через $\mathbb{B}^u$ множество всех u-битовых строк, и через $\mathbb{B}^*$ — множество всех битовых строк конечной длины, в том числе пустую строку. Битовую строку, состоящую из u нулей, будем обозначать через 0^u. Длину битовой строки U будем обозначать через $|U|$. Через $|U|_u = \lceil |U|/u \rceil$ обозначим длину битовой строки U в u-битовых блоках. Через $\mathbb{B}^{\leqslant u}$ обозначим множество всех строк, длина которых меньше или равна u.

Обозначим через $\mathbb{B}^{n\times m}$ множество всех упорядоченных наборов из m элементов, где каждый элемент является n-битовой строкой, значение m будем называть длиной набора. Для набора $X \in \mathbb{B}^{n\times m}$ через $\{X\}$ будем обозначать множество всех его элементов. Для упрощения формул для $x \in \mathbb{B}^n$, $X, Y \in \mathbb{B}^{n\times m}$ будем использовать обозначения $x \in X$, $x \notin X$, и $X \cap Y$ вместо $x \in \{X\}$, $x \notin \{X\}$, и $\{X\} \cap \{Y\}$.

Для битовой строки U и целого числа $0 < l \leqslant |U|$ через $\mathrm{msb}_l(U)$ ($\mathrm{lsb}_l(U)$) будем обозначать строку, состоящую из l крайних левых (правых) бит строки U.

Статья получена 4 февраля 2022

Лилия Руслановна Ахметзянова, МГУ им. М.В. Ломоносова, ООО "КРИПТО-ПРО", (email: lah@cryptopro.ru).

Для целых чисел $l > 0$ и $2^l > i \geqslant 0$ через $\mathrm{str}_l(i)$ будем обозначать l-битовое представление числа i, в котором наименее значащий бит находится справа. Для целого числа $l \geqslant 0$ и битовой строки $U \in \{0,1\}^l$ через $\mathrm{int}(U)$ будем обозначать целое число $i < 2^l$, такое что $\mathrm{str}_l(i) = U$. Для битовых строк $a \in \{0,1\}^n$ и $b \in \{0,1\}^n$ через $a \otimes b$ будем обозначать строку, которая является результатом их перемножения в поле $GF(2^n)$ (здесь строки кодируют полиномы стандартным образом). Факт того, что значение s выбрано из некоторого множества S в соответствии с равномерным распределением $\mathcal{U}$, будем обозначать через $s \xleftarrow{\mathcal{U}} S$.

Через $inc_r(U)$ ($inc_l(U)$) будем обозначать функцию, которая принимает на входу строку $L\|R$, где $L, R \in \mathbb{B}^{n/2}$, и возвращает строку $L\|\mathrm{str}_{n/2}(\mathrm{int}(R) + 1 \bmod 2^{n/2})$ ($\mathrm{str}_{n/2}(\mathrm{int}(L) + 1 \bmod 2^{n/2})\|R$).

Для множества $\mathbb{B}^n$ через $Perm(n)$ обозначим множество всех биективных отображений множества $\mathbb{B}^n$ в себя, а через $Func(n)$ — множество всех отображений множества $\mathbb{B}^n$ в себя. Под блочным шифром E (или просто шифром) с длиной блока n и длиной ключа k будем понимать произвольное семейство перестановок на множестве $\mathbb{B}^n$, параметризованное параметром $K \in \mathbb{B}^k$, т.е. $E = \left\{E_K \in Perm(n) \mid K \in \mathbb{B}^k\right\}$.

III. Модель противника

В данном разделе вводятся формальные определения схем аутентифицированного шифрования и стандартной модели IND-CPA для анализа уровня стойкости данных схем с точки зрения конфиденциальности.

Определение III.1. *AEAD-схемой* для множества ключей $\boldsymbol{K}$, множества векторов инициализации $\boldsymbol{N}$, множества открытых текстов $\boldsymbol{P}$, множества ассоциированных данных $\boldsymbol{A}$, множества шифртекстов $\boldsymbol{C}$ и множества имитовставок $\boldsymbol{T}$ является набор алгоритмов $\Pi = \{\Pi.\mathsf{Gen},\ \Pi.\mathsf{Enc},\ \Pi.\mathsf{Dec}\}$, где

- $\Pi.\mathsf{Gen}() \xrightarrow{\$} K$: Вероятностный алгоритм генерации ключа. Результатом работы данного алгоритма является ключ $K \in \boldsymbol{K}$.
- $\Pi.\mathsf{Enc}(K, N, A, P) \to (C, T)$: Детерминированный алгоритм аутентифицированного шифрования, принимающий на вход значения $K \in \boldsymbol{K}$, $N \in \boldsymbol{N}$, $A \in \boldsymbol{A}$ и $P \in \boldsymbol{P}$. Результатом работы данного алгоритма являются значения $C \in \boldsymbol{C}$ и $T \in \boldsymbol{T}$.
- $\Pi.\mathsf{Dec}(K, N, A, C, T) \to P$: Детерминированный алгоритм расшифрования с проверкой целостности, принимающий на вход значения $K \in \boldsymbol{K}$, $N \in \boldsymbol{N}$, $A \in \boldsymbol{A}$, $C \in \boldsymbol{C}$ и $T \in \boldsymbol{T}$. Результатом работы данного алгоритма является значение $P \in \boldsymbol{P}$ или символ ошибки $\perp$.

Для формализации модели противника используется алгоритмический подход (для деталей см. [5]). Данный подход заключается в построении вероятностного интерактивного алгоритма — экспериментатора, моделирующего работу схемы в присутствии противника, и определении количественной характеристики успешности противника по реализации угрозы — преимущества.

Ниже приведена стандартная модели IND-CPA, используемая для анализа AEAD-режимов с точки зрения конфиденциальности. В данной модели противник может подавать на шифрование сообщения только с уникальными векторами инициализации.

Определение III.2. [6] Преимущество противника $\mathcal{A}$ в модели IND-CPA для AEAD-схемы Π с длиной имитовставки s определяется следующим образом:

$$\mathsf{Adv}_{\Pi}^{\text{IND-CPA}}(\mathcal{A}) = \Pr\left[\mathbf{Exp}_{\Pi}^{\text{IND-CPA-1}}(\mathcal{A}) \to 1\right] - \Pr\left[\mathbf{Exp}_{\Pi}^{\text{IND-CPA-0}}(\mathcal{A}) \to 1\right],$$

где $\mathbf{Exp}_{\Pi}^{\text{IND-CPA-}b}(\mathcal{A}), b \in \{0,1\}$ описываются следующим образом:

$\mathbf{Exp}_{\Pi}^{\text{IND-CPA-}b}(\mathcal{A})$	Oracle $Encrypt\text{-}b(N, A, P)$
$K \xleftarrow{\$} \Pi.\mathsf{Gen}(\,)$	**if** $N \in \mathcal{L}$:
$\mathcal{L} \leftarrow \emptyset$	**return** $\perp$
$b' \xleftarrow{\$} \mathcal{A}^{Encrypt\text{-}b(\,)}$	**if** $b = 1$:
return b'	$(C, T) \leftarrow \Pi.\mathsf{Enc}(K, N, A, P)$
	else :
	$C \parallel T \xleftarrow{\mathcal{U}} \{0,1\}^{\lvert P\rvert + s}$
	$\mathcal{L} \leftarrow \mathcal{L} \cup \{N\}$
	return (C, T)

IV. Режим MGM

Параметрами режима $\mathrm{MGM}_{E,s}$ являются блочный шифр E с длиной ключа k и длиной блока n, а также длина имитовставки в битах s. Данные параметры фиксируются в рамках конкретного протокола.

Режим MGM определен для следующих множеств: $\boldsymbol{K} = \{0,1\}^k$, $\boldsymbol{N} = \mathbb{B}^{n-1}$, $\boldsymbol{A} = \boldsymbol{P} = \boldsymbol{C} = \{0,1\}^{\leqslant 2^{n/2}-1}$, $\boldsymbol{T} = \{0,1\}^s$. Дополнительно на длину открытого текста и ассоциированных данных накладывается следующее ограничение: $0 < |A| + |P| \leqslant n \cdot 2^{n/2}$. Алгоритм аутентифицированного шифрования режима определен на Рис. 1 и Рис. 2.

V. Оценка стойкости режима MGM в модели IND-CPA

Теорема V.1. *Для любого противника $\mathcal{A}$ в модели IND-CPA для схемы $\mathrm{MGM}_{Perm(n)}$, который делает не более q запросов, где максимальная суммарная длина открытого текста и ассоциированных данных не превосходит l блоков, выполнено следующее неравенство:*

$$\mathbf{Adv}^{\text{IND-CPA}}_{\mathrm{MGM}_{Perm(n),s}}(\mathcal{A}) \leqslant \frac{(2ql + 5q)^2}{2^n}.$$

Для доказательства данной теоремы сначала введем вспомогательную модель mIND-CPA для семейства функций $\mathcal{F} \subseteq Func(n)$. Далее докажем, что из стойкости в данной модели для семейства $Perm(n)$ следует стойкость схемы $\mathrm{MGM}_{Perm(n)}$ в модели IND-CPA (см. Утверждение V.2).

После этого докажем оценку для семейства $Perm(n)$ в модели mIND-CPA (см. Теорему V.3).

A. Вспомогательная модель mIND-CPA

Введем вспомогательную модель противника mIND-CPA с параметрами l и s для семейства функций $\mathcal{F} \subseteq Func(n)$. При инициализации эксперимента выбирается функция $f \xleftarrow{\mathcal{U}} \mathcal{F}$. Далее противнику предоставляется доступ к оракулу $Encrypt\text{-}b$, к которому он делает два последовательных запроса следующего вида:

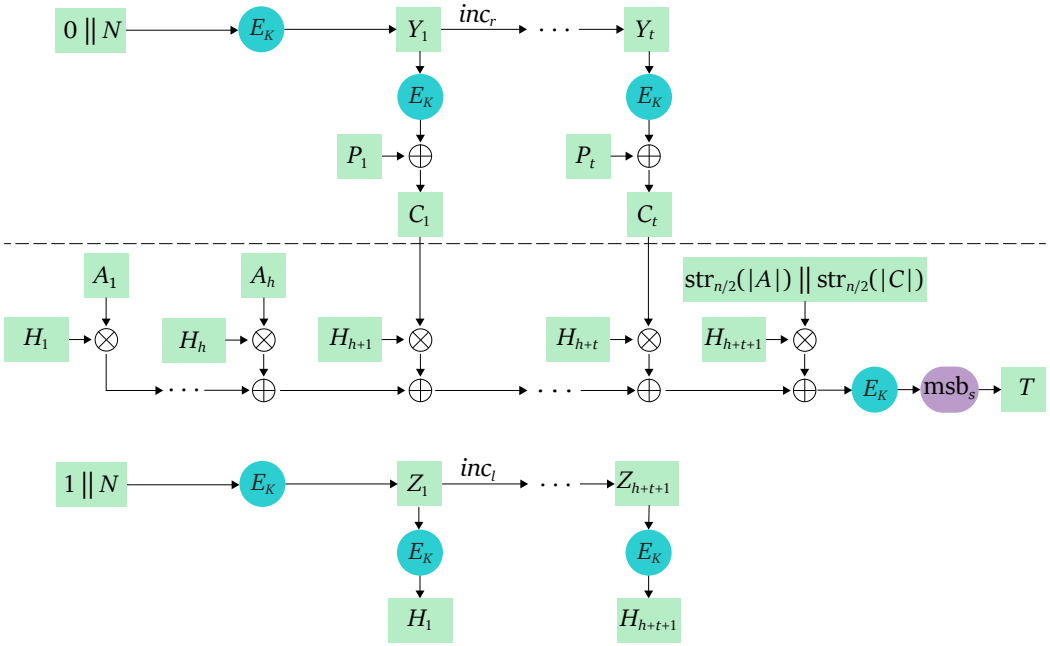

Рис. 1. Алгоритм аутентифицированного шифрования режима MGM

.

MGM.Enc(K, N, A, P)

$h \leftarrow |A|_n, t \leftarrow |P|_n$

$\ell \leftarrow h + t + 1$

. . . . . Шифрование . . . . . .

$Y_1 \leftarrow E_K(0\|N),\ \Gamma_1 \leftarrow E_K(Y_1)$

for $i = 2 \ldots t$ **do**:

$Y_i \leftarrow inc_r(Y_{i-1}),\ \Gamma_i \leftarrow \mathsf{E}_K(Y_i)$

$C \leftarrow P \oplus \mathrm{msb}_{|P|}(\Gamma_1 \| \ldots \| \Gamma_t)$

. . . . . Дополнение . . . . . .

$a \leftarrow n|A|_n - |A|$

$c \leftarrow n|C|_n - |C|$

$len \leftarrow \mathrm{str}_{n/2}(|A|) \| \mathrm{str}_{n/2}(|C|)$

$M_1\| \ldots \|M_\ell \leftarrow A\|0^a\|C\|0^c\|len$

. . . . Вычисление имитовставки . . . .

$Z_1 \leftarrow E_K(1\|N),\ H_1 \leftarrow E_K(Z_1)$

for $i = 2 \ldots \ell$ **do**:

$Z_i \leftarrow inc_l(Z_{i-1}),\ H_i \leftarrow \mathsf{E}_K(Z_i)$

$\tau \leftarrow \bigoplus_{i=1}^{l} M_i \otimes H_i$

$T \leftarrow \mathrm{msb}_s(\mathsf{E}_K(\tau))$

return (C, T)

Рис. 2. Алгоритм аутентифицированного шифрования режима MGM

1) Первый запрос состоит из вектора инициализации $N \in \mathbb{B}^{n-1}$.

В случае оракула *Encrypt*-1 ответом на запрос является набор $\Gamma \in \mathbb{B}^{n \times l}$, состоящий из l блоков $\Gamma_k \in \mathbb{B}^n$, $k = 1, \ldots, l$. Также оракул сохраняет значения N и l, которые далее будут использоваться для обработки следующего запроса.

Обработка первого запроса:

$$\begin{cases} Y_1 = f(0\|N), \\ Y_k = inc_r(Y_{k-1}), & 2 \leqslant k \leqslant l, \\ \Gamma_k = f(Y_k), & 1 \leqslant k \leqslant l. \end{cases}$$

В случае оракула *Encrypt*-0 ответом является набор $\Gamma \in \mathbb{B}^{n \times l}$, состоящий из l блоков $\Gamma_k \xleftarrow{\mathcal{U}} \mathbb{B}^n$, каждый из которых выбирается случайно равновероятно из множества $\mathbb{B}^n$.

2) Второй запрос содержит набор $X \in \mathbb{B}^{n \times l}$, состоящий из l блоков $X_k \in \mathbb{B}^n$, $k = 1, \ldots, l$.

В случае оракула *Encrypt*-1 данный набор используется в качестве входа в мультилинейную функцию алгоритма вычисления имитовставки, который также принимает на вход вектор инициализации N, переданный на предыдущем запросе. Для предотвращения тривиальных атак введем следующее ограничение на набор X: $X_l \neq 0^n$. В качестве ответа противнику возвращается значение $T \in \mathbb{B}^s$.

Обработка второго запроса:

$$\begin{cases} Z_1 = f(1\|N), \\ Z_k = inc_l(Z_{k-1}), & 2 \leqslant k \leqslant l, \\ H_k = f(Z_k), & 1 \leqslant k \leqslant l, \\ \tau = \sum\limits_{k=1}^{l} H_k \cdot X_k, \\ T = \mathrm{msb}_s(f(\tau)). \end{cases}$$

В случае оракула *Encrypt*-0 противник получает значение $T \xleftarrow{\mathcal{U}} \mathbb{B}^s$, которое выбирается случайно равновероятно из $\mathbb{B}^s$.

Определение V.1. Преимущество противника $\mathcal{A}$ в модели mIND-CPA с параметрами l и s для семейства функций $\mathcal{F}$ определяется следующим образом:

$$\mathsf{Adv}_{\mathcal{F}}^{\text{mIND-CPA}_{l,s}}(\mathcal{A}) = \Pr\left[\mathbf{Exp}_{\mathcal{F}}^{\text{mIND-CPA}_{l,s}\text{-}1}(\mathcal{A}) \to 1\right] - \\ - \Pr\left[\mathbf{Exp}_{\mathcal{F}}^{\text{mIND-CPA}_{l,s}\text{-}0}(\mathcal{A}) \to 1\right],$$

где $\mathbf{Exp}_{\mathcal{F}}^{\text{mIND-CPA}_{l,s}\text{-}b}(\mathcal{A}), b \in \{0,1\}$, определены выше.

Легко показать, что верно следующее утверждение.

Утверждение V.2. Для любого противника $\mathcal{A}$ в модели IND-CPA для схемы $\text{MGM}_{\mathcal{F},f}$, делающего не более q запросов, где максимальная суммарная длина открытого текста и ассоциированных данных не превосходит l блоков, существует противник $\mathcal{A}'$ в модели mIND-CPA с параметрами $l+1$ и s для семейства $\mathcal{F}$, такой что

$$\mathbf{Adv}_{\text{MGM}_{\mathcal{F},f}}^{\text{IND-CPA}}(\mathcal{A}) = \mathbf{Adv}_{\mathcal{F}}^{\text{mIND-CPA}_{l+1,s}}(\mathcal{A}'),$$

где $\mathcal{A}'$ делает не более q пар связанных запросов.

Доказательство. Построим противника $\mathcal{A}'$ в модели mIND-CPA для семейства $\mathcal{F}$, использующего противника $\mathcal{A}$. Алгоритм $\mathcal{A}'$ запускает алгоритм $\mathcal{A}$, перехватывает его запросы и сам их обрабатывает. Алгоритм симулирует оракул *Encrypt-b* для $\mathcal{A}$, делая специальные запросы к своему оракулу *Encrypt-b*. Перехватывая запросы (N, A, P) от $\mathcal{A}$ противник $\mathcal{A}'$ делает пару следующих связанных запросов к своему оракулу. Первый запрос состоит из вектора инициализации N. Получая в ответ набор $\Gamma = (\Gamma_1, \ldots, \Gamma_l)$, противник $\mathcal{A}'$ формирует значение $C = P \oplus \text{msb}_{|P|}(\Gamma_1 \| \ldots \| \Gamma_{|P_n|})$ (заметим, что $l \geqslant |P|_n$). После этого $\mathcal{A}'$ делает второй запрос — набор X, состоящий из блоков строки $A\|0^{n-a}\|C\|0^{n-c}\|(\text{str}_{n/2}(|A|)\|\text{str}_{n/2}(|C|))\|0^{n(l-|A|_n-|P|_n)}$. Заметим, что длина набора X в точности равна $l+1$ блоков, а дополнение нулями не влияет на процесс вычисления T. Получая в ответ на второй запрос значение T, алгоритм $\mathcal{A}'$ возвращает значение (C, T) алгоритму $\mathcal{A}$. В качестве результата своей работы алгоритм $\mathcal{A}'$ возвращает то же самое, что и алгоритм $\mathcal{A}$.

Заметим, что противник $\mathcal{A}'$, взаимодействующий с оракулом *Encrypt-b* в модели mIND-CPA, в точности симулирует для противника $\mathcal{A}$ работу оракула *Encrypt-b* в модели IND-CPA. Поэтому

$$\mathbf{Adv}_{\mathcal{F}}^{\text{mIND-CPA}}(\mathcal{A}') = \mathbf{Adv}_{\text{MGM}_{\mathcal{F}}}^{\text{IND-CPA}}(\mathcal{A}).$$

□

B. Оценка в модели mIND-CPA для семейства $Perm(n)$

Рассмотрим детерминированный алгоритм-противник $\mathcal{A}$ с неограниченными вычислительными ресурсами, который делает в точности q пар связанных запросов. Тогда алгоритм определяется $2q$ функциями:

- q функций $N_i^{\mathcal{A}}$, которые определяют выбор параметра N из первой части i-го запроса к оракулу *Encrypt-b*.

  Функция $N_1^{\mathcal{A}}$ является константной, т.е. $N_1^{\mathcal{A}} = N_1$. Следующие функции $N_i^{\mathcal{A}}$, $i = 2, \ldots, q$, определяются следующим образом:

$$N_i = N_i^{\mathcal{A}}(\Gamma^1, T_1, \ldots, \Gamma^{i-1}, T_{i-1}) : \\ \underbrace{\mathbb{B}^{n\times l} \times \mathbb{B}^s \times \ldots \times \mathbb{B}^{n\times l} \times \mathbb{B}^s}_{2(i-1)} \to \mathbb{B}^{n-1}.$$

  Данные функции для $\forall\ \Gamma^1, T_1, \ldots, \Gamma^q, T_q$ должны удовлетворять следующему условию:

$$\forall\ 1 \leqslant i, j \leqslant q,\ i \neq j, \\ N_i^{\mathcal{A}}(\Gamma^1, \ldots, T_{i-1}) \neq N_j^{\mathcal{A}}(\Gamma^1, \ldots, T_{j-1}).$$

- q функций $X_i^{\mathcal{A}}$, $i = 1, \ldots, q$, которые определяют значение набора X из второй части i-го запроса к оракулу *Encrypt-b*:

$$X^i = X_i^{\mathcal{A}}(\Gamma^1, T_1, \ldots, \Gamma^{i-1}, T_{i-1}, \Gamma^i) : \\ \underbrace{\mathbb{B}^{n\times l} \times \mathbb{B}^s \times \ldots \times \mathbb{B}^{n\times l} \times \mathbb{B}^s}_{2(i-1)} \times \mathbb{B}^{n\times l} \to \mathbb{B}^{n\times l}.$$

  Данные функции для $\forall\ \Gamma^1, T_1, \ldots, \Gamma^{q-1}, T_{q-1}, \Gamma^q$ должны удовлетворять следующему условию:

$$\forall\ 1 \leqslant i \leqslant q,\ X_l^i \neq 0^n\ .$$

Далее через $\mathbf{D}^i$, $i = 1, \ldots, q$, будем обозначать набор, состоящий из всех значений, поступивших на вход функции f после обработки первых i запросов в модели mIND-CPA. В данный набор входят значения $0\|N_j$, $1\|N_j$, τ_j и значения элементов наборов $Y^j = (Y_1^j, \ldots, Y_l^j)$, $Z^j = (Z_1^j, \ldots, Z_l^j)$, $j = 1, \ldots, i$.

Докажем вспомогательную комбинаторную лемму V.1, в рамках которой оценивается вероятность коллизии в наборе $\mathbf{D}^q$. Основная идея доказательства заключается в подсчете количества различных размещений с пересечением в пространстве $\mathbb{B}^n$, представленном в виде тора $\mathbb{B}^{n/2} \times \mathbb{B}^{n/2}$, объектов разного типа: «точки», «горизонтальные отрезки» и «вертикальные отрезки» длины l (см. Рис. 3).

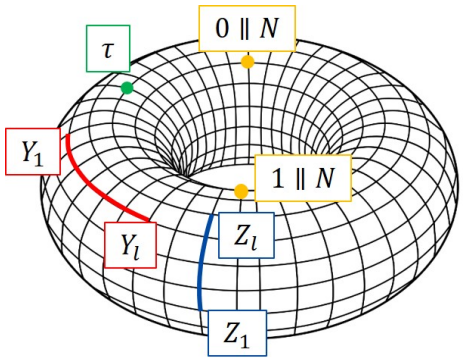

Рис. 3. Размещение объектов на торе $\mathbb{B}^{n/2} \times \mathbb{B}^{n/2}$: «точками» являются значения $\tau, 0\|N, 1\|N$, «горизонтальными отрезками» являются наборы $Y = \{Y_1, \ldots, Y_l\}$, т.к. $Y_i = inc_{\mathbf{r}}(Y_{i-1})$, «вертикальными отрезками» являются наборы $Z = \{Z_1, \ldots, Z_l\}$, т.к. $Z_i = inc_{\mathbf{l}}(Z_{i-1})$

Дополнительные обозначения. Далее будем обозначать случайную величину с использованием символа тильда (например, $\widetilde{\lambda}$), а конкретные ее значения – без символа тильда (λ). Пусть $\widetilde{\Lambda} = (\widetilde{\lambda}_1, \ldots, \widetilde{\lambda}_m)$ – набор из m случайных величин $\widetilde{\lambda}_i$: $Func(n) \to \mathbb{B}^n$, $i = 1, \ldots, m$, при этом на множестве $Func(n)$ задано равновероятное

распределение. Для $\omega \in Func(n)$, через $\widetilde{\Lambda}(\omega)$ обозначим набор $(\widetilde{\lambda}_1(\omega), \ldots, \widetilde{\lambda}_1(\omega))$.

Пусть Λ — некоторое значение набора $\widetilde{\Lambda}$. Через $\widetilde{\Lambda} \setminus \{\widetilde{\lambda}_i\}$ будем обозначать набор $\widetilde{\Lambda}$, из которого убрали величину $\widetilde{\lambda}_i$. Тогда через $\Lambda \setminus \{\widetilde{\lambda}_i\}$ обозначим конкретное значение набора $\widetilde{\Lambda} \setminus \{\widetilde{\lambda}_i\}$.

Через $\Lambda\, coll$ обозначим предикат, который выполняется тогда и только тогда, когда произошла коллизия среди элементов набора Λ. Обозначим через $\Lambda\ \overline{coll}$ отрицание предиката $\Lambda\ coll$. Для набора $\widetilde{\Lambda}$ через $\widetilde{\Lambda}\, coll$ обозначим событие $\{\omega \in Func(n) \mid \widetilde{\Lambda}(\omega)\ coll = 1\}$, т.е. набор Λ содержит одинаковые элементы. Через $\widetilde{\Lambda}\ \overline{coll}$ будем обозначать отрицание события $\widetilde{\Lambda}\ coll$.

Лемма V.1. *Для любого противника $\mathcal{A}$ в модели mIND-CPA с параметрами l и s для семейства $Func(n)$, который делает q пар связанных запросов, выполнено следующее неравенство:*

$$\Pr\left[\widetilde{\mathbf{D}}^q\ coll\right] \leqslant \frac{(2ql+3q)^2}{2^n}.$$

Доказательство. По формуле полной вероятности имеем:

$$\begin{aligned}\Pr\left[\widetilde{\mathbf{D}}^q\ coll\right] &= \Pr\left[\widetilde{\mathbf{D}}^q\ coll \cap \widetilde{\mathbf{D}}^{q-1}\ \overline{coll}\right] + \\ &+ \Pr\left[\widetilde{\mathbf{D}}^q\ coll \cap \widetilde{\mathbf{D}}^{q-1}\ coll\right] = \\ &= \Pr\left[\widetilde{\mathbf{D}}^q\ coll \cap \widetilde{\mathbf{D}}^{q-1}\ \overline{coll}\right] + \Pr\left[\widetilde{\mathbf{D}}^{q-1}\ coll\right].\end{aligned}$$

Заметим, что аналогичная формула верна и для $\Pr\left[\widetilde{\mathbf{D}}^{q-1}\ coll\right]$. Поэтому

$$\Pr\left[\widetilde{\mathbf{D}}^q\ coll\right] = \sum_{i=1}^{q} \Pr\left[\widetilde{\mathbf{D}}^i\ coll \cap \widetilde{\mathbf{D}}^{i-1}\ \overline{coll}\right].$$

Здесь для удобства записи формулы через $\widetilde{\mathbf{D}}^0\ \overline{coll}$ обозначен истинный предикат.

Рассмотрим вероятность $\Pr\left[\widetilde{\mathbf{D}}^i\ coll \cap \widetilde{\mathbf{D}}^{i-1}\ \overline{coll}\right]$, $i = 1, \ldots, q$. Применим формулу полной вероятности по всем возможным значениям $\Gamma^j, T_j,\ j = 1, \ldots, i-1$:

$$\begin{aligned}&\Pr\left[\widetilde{\mathbf{D}}^i\ coll \cap \widetilde{\mathbf{D}}^{i-1}\ \overline{coll}\right] = \\ &= \sum_{\substack{\Gamma^1,\ldots,\Gamma^{i-1}\\ T_1,\ldots,T^{i-1}}} \Pr\left[\widetilde{\mathbf{D}}^i\ coll \cap \widetilde{\mathbf{D}}^{i-1}\ \overline{coll} \cap \left\{\begin{smallmatrix}\widetilde{\Gamma}^j=\Gamma^j\\ \widetilde{T}_j=T_j\end{smallmatrix}\right\}_{j=1}^{i-1}\right]. \quad (1)\end{aligned}$$

Рассмотрим слагаемое под знаком суммы для любых $\Gamma^j, T_j,\ j = 1, \ldots, i-1$,. Заметим, что после фиксации этих значений

1) в наборе $\widetilde{\mathbf{D}}^i$ случайные величины $\widetilde{N}_j$ принимают конкретные значения $N_j = N_j^{\mathcal{A}}(\Gamma^1, T_1, \ldots, \Gamma^{j-1}, T_{j-1})$, $j = 1, \ldots, i$, аналогично в наборе $\widetilde{\mathbf{D}}^{i-1}$ случайные величины $\widetilde{N}_1, \ldots, \widetilde{N}_{i-1}$ также принимают соответствующие конкретные значения.
2) случайные величины $\widetilde{X}^j$ принимают конкретные значения $X^j = X_j^{\mathcal{A}}(\Gamma^1, T_1, \ldots, \Gamma^{j-1}, T_{j-1}, \Gamma^j)$, $j = 1, \ldots, i-1$.

Также заметим, что после дополнительной фиксации значений $Y^j, Z^j, H^j,\ j = 1, \ldots, i-1$, все случайные величины в наборе $\widetilde{\mathbf{D}}^{i-1}$ принимают конкретные значения, в том числе случайные величины $\widetilde{\tau}_j$ принимают конкретные значения $\tau_j = \sum_{k=1}^{l} H_k^j \cdot X_k^j$, $j = 1, \ldots, i-1$. Отметим, что однозначная фиксация наборов Y^j, Z^j происходит после фиксации значений Y_1^j, Z_1^j, так как последующие значения в этих наборах однозначно определяются по первым блокам.

Таким образом, при фиксированных $\Gamma^j, T_j, Y^j, Z^j, H^j,\ j = 1, \ldots, i-1$, значение предиката $\mathbf{D}^{i-1}\ \overline{coll}$, становится определенным. Поэтому для каждого слагаемого из формулы (1) верно следующее равенство:

$$\begin{aligned}&\Pr\left[\widetilde{\mathbf{D}}^i\ coll \cap \widetilde{\mathbf{D}}^{i-1}\ \overline{coll} \cap \left\{\begin{smallmatrix}\widetilde{\Gamma}^j=\Gamma^j\\ \widetilde{T}_j=T_j\end{smallmatrix}\right\}_{j=1}^{i-1}\right] = \\ &= \sum_{\substack{Y^1,\ldots,Y^{i-1}\\ Z^1,\ldots,Z^{i-1}\\ H^1,\ldots,H^{i-1}:\\ \mathbf{D}^{i-1}\ \overline{coll}}} \Pr\left[\widetilde{\mathbf{D}}^i\ coll \cap \left\{\begin{smallmatrix}\widetilde{\Gamma}^j=\Gamma^j\\ \widetilde{T}_j=T_j\end{smallmatrix}\right\}_{j=1}^{i-1} \cap \left\{\begin{smallmatrix}\widetilde{Y}^j=Y^j\\ \widetilde{Z}^j=Z^j\\ \widetilde{H}^j=H^j\end{smallmatrix}\right\}_{j=1}^{i-1}\right].\end{aligned}$$

Далее для краткости событие

$$\left\{\begin{smallmatrix}\widetilde{\Gamma}^j=\Gamma^j\\ \widetilde{T}_j=T_j\end{smallmatrix}\right\}_{j=1}^{i-1} \cap \left\{\begin{smallmatrix}\widetilde{Y}^j=Y^j\\ \widetilde{Z}^j=Z^j\\ \widetilde{H}^j=H^j\end{smallmatrix}\right\}_{j=1}^{i-1}$$

будем обозначать через fixed. Вероятность данного события при условии выполнения предиката $\mathbf{D}^{i-1}\ \overline{coll}$ равна

$$\frac{1}{2^{n(i-1)(2l+2)+s(i-1)}}. \quad (2)$$

Последняя сумма может быть разбита на две подсуммы по событию, когда хотя бы одно из значений $0\|N_i$ или $1\|N_i$ попало в множество $\mathbf{D}^{i-1}$:

$$\begin{aligned}sum_1 &= \sum_{\substack{Y^1,\ldots,Y^{i-1}\\ Z^1,\ldots,Z^{i-1}\\ H^1,\ldots,H^{i-1}:\\ \mathbf{D}^{i-1}\ \overline{coll}\\ 0\|N_i \vee 1\|N_i \in \mathbf{D}^{i-1}}} \Pr\left[\text{fixed}\right];\\ sum_2 &= \sum_{\substack{Y^1,\ldots,Y^{i-1}\\ Z^1,\ldots,Z^{i-1}\\ H^1,\ldots,H^{i-1}:\\ \mathbf{D}^{i-1}\ \overline{coll}\\ 0\|N_i, 1\|N_i \notin \mathbf{D}^{i-1}}} \Pr\left[\widetilde{\mathbf{D}}^i\ coll \cap \text{fixed}\right].\end{aligned}$$

Оценка первой суммы. Рассмотрим слагаемое sum_1. Применяя формулу (2), получим следующее неравенство:

$$sum_1 \leqslant \underbrace{\#\left\{\begin{smallmatrix}Y^1,\ldots,Y^{i-1}\\ Z^1,\ldots,Z^{i-1}\\ H^1,\ldots,H^{i-1}:\\ 0\|N_i \vee 1\|N_i \in \mathbf{D}^{i-1}\end{smallmatrix}\right\}}_{A} \cdot \frac{1}{2^{n(i-1)(2l+2)+s(i-1)}}. \quad (3)$$

Оценим мощность множества A для фиксированных $\Gamma^j, T_j, j = 1, \ldots, i-1$ (как следствие, для фиксированных $N_j, X^j,\ j = 1, \ldots, i-1$, и N_i). Множество A принадлежит объединению следующих подмножеств:

$$\begin{aligned}A \subseteq \bigcup_{b\in\mathbb{B}} \bigcup_{j=1}^{i-1} \Bigg(&\left\{\begin{smallmatrix}Y^j:\\ b\|N_i \in Y^j\end{smallmatrix}\right\} \times \left\{\begin{smallmatrix}Y^1,\ldots,Y^{j-1},Y^{j+1},\ldots,Y^{i-1}\\ Z^1,\ldots,Z^{i-1}\\ H^1,\ldots,H^{i-1}\end{smallmatrix}\right\} \cup \\ &\cup \left\{\begin{smallmatrix}Z^j:\\ b\|N_i \in Z^j\end{smallmatrix}\right\} \times \left\{\begin{smallmatrix}Y^1,\ldots,Y^{i-1}\\ Z^1,\ldots,Z^{j-1},Z^{j+1},\ldots,Z^{i-1}\\ H^1,\ldots,H^{i-1}\end{smallmatrix}\right\} \cup \\ &\cup \left\{\begin{smallmatrix}H^j:\\ b\|N_i = \tau_j\end{smallmatrix}\right\} \times \left\{\begin{smallmatrix}Y^1,\ldots,Y^{i-1}\\ Z^1,\ldots,Z^{i-1}\\ H^1,\ldots,H^{j-1},H^{j+1},\ldots,H^{i-1}\end{smallmatrix}\right\}\Bigg).\end{aligned}$$

Здесь в первой строке выписано множество, при котором одна из двух «точек» $b\|N_i$ попала на один из $i-1$ «горизонтальных отрезков» Y^j длины l, во второй строке – на один из $i-1$ «вертикальных отрезков» Z^j длины l, в третьей строке – в одну из $i-1$ «точек» τ_j.

Таким образом, мощность множества A можно оценить следующим образом:

$$\#A \leqslant \sum_{b\in\mathbb{B}}\sum_{j=1}^{i-1}\Big(\underbrace{\#\left\{Y^j : b\|N_i \in Y^j\right\}}_{=l}\cdot 2^{n(i-1)(l+2)-n} + \underbrace{\#\left\{Z^j : b\|N_i \in Z^j\right\}}_{=l}\cdot 2^{n(i-1)(l+2)-n} + \underbrace{\#\left\{H^j : b\|N_i = \tau_j\right\}}_{=2^{nl-n}}\cdot 2^{n(i-1)(l+2)-nl}\Big) = \underbrace{(i-1)(4l+2)}_{\omega_1}\cdot 2^{n(i-1)(l+2)-n}.$$

Подставив полученную оценку для мощности в формулу (3) и проведя простые арифметические действия, получим оценку для первого слагаемого:

$$sum_1 \leqslant \omega_1\cdot 2^{n(i-1)(l+2)-n}\cdot\frac{1}{2^{n(i-1)(2l+2)+s(i-1)}} = \frac{\omega_1}{2^{n(i-1)l+s(i-1)+n}}. \quad (4)$$

Оценка второй суммы. Рассмотрим второе слагаемое sum_2. Оно отражает вероятность коллизии за счет выбора «отрезков» Y^i, Z^i и за счет выбора «точек» τ_i. Для каждого слагаемого под знаком суммы верно

$$\Pr\left[\widetilde{\mathbf{D}}^i\ coll\cap\text{fixed}\right] = \underbrace{\sum_{\substack{Y^i,Z^i:\\ \mathbf{D}^i\setminus\{\widetilde{\tau}_i\}\ coll}}\Pr\left[\text{fixed}\cap\left\{\begin{smallmatrix}\widetilde{Y}^i=Y^i\\ \widetilde{Z}^i=Z^i\end{smallmatrix}\right\}\right]}_{sum_2^1} + \underbrace{\sum_{\substack{Y^i,Z^i:\\ \mathbf{D}^i\setminus\{\widetilde{\tau}_i\}\ \overline{coll}}}\Pr\left[\widetilde{\mathbf{D}}^i\ coll\cap\text{fixed}\cap\left\{\begin{smallmatrix}\widetilde{Y}^i=Y^i\\ \widetilde{Z}^i=Z^i\end{smallmatrix}\right\}\right]}_{sum_2^2}.$$

Рассмотрим слагаемое sum_2^1. В силу (2) и того, что вероятность события $\left\{\begin{smallmatrix}\widetilde{Y}^i=Y^i\\ \widetilde{Z}^i=Z^i\end{smallmatrix}\right\}$ при условии выполнения предиката $\mathbf{D}^i\setminus\{\widetilde{\tau}_i\}\ \overline{coll}$ равна $\frac{1}{2^{2n}}$, верно следующее неравенство:

$$sum_2^1 \leqslant \#\underbrace{\left\{\begin{smallmatrix}Y^i,Z^i:\\ \mathbf{D}^i\setminus\{\widetilde{\tau}_i\}\ coll\end{smallmatrix}\right\}}_{B}\cdot\frac{1}{2^{n(i-1)(2l+2)+s(i-1)+2n}}. \quad (5)$$

Оценим мощность множества B для фиксированных значений $\Gamma^j, T_j, Y^j, Z^j, H^j$, $j=1,\ldots,i-1$ (как следствие, для фиксированных N_i и всех значений из $\mathbf{D}^{i-1}$). Множество B принадлежит объединению следующих множеств:

$$B \subseteq \bigcup_{j=1}^{i-1}\left(\left\{\begin{smallmatrix}Z^i:\\ Z^i\cap Z^j\neq\emptyset\end{smallmatrix}\right\}\cup\left\{\begin{smallmatrix}Z^i:\\ \tau_j\in Z^i\end{smallmatrix}\right\}\right)\times\{Y^i\}\ \cup \bigcup_{j=1}^{i-1}\left(\left\{\begin{smallmatrix}Y^i:\\ Y^i\cap Y^j\neq\emptyset\end{smallmatrix}\right\}\cup\left\{\begin{smallmatrix}Y^i:\\ \tau_j\in Y^i\end{smallmatrix}\right\}\right)\times\{Z^i\}\ \cup \bigcup_{j=1}^{i-1}\bigcup_{k_j=1}^{l}\left(\left\{\begin{smallmatrix}Z^i:\\ Y^j_{k_j}\in Z^i\end{smallmatrix}\right\}\times\{Y^i\}\cup\left\{\begin{smallmatrix}Y^i:\\ Z^j_{k_j}\in Y^i\end{smallmatrix}\right\}\times\{Z^i\}\right)\ \cup \bigcup_{j=1}^{i}\bigcup_{b\in\mathbb{B}}\left(\left\{\begin{smallmatrix}Z^i:\\ b\|N_j\in Z^i\end{smallmatrix}\right\}\times\{Y^i\}\cup\left\{\begin{smallmatrix}Y^i:\\ b\|N_i\in Y^i\end{smallmatrix}\right\}\times\{Z^i\}\right)\ \cup \left\{\begin{smallmatrix}Y^i,Z^i:\\ Y^i\cap Z^i\neq\emptyset\end{smallmatrix}\right\}.$$

Здесь в первой (второй) строке выписано множество, в котором «вертикальный отрезок» Z^i («горизонтальный отрезок» Y^i) длины l пересекается с одим из $i-1$ «вертикальных отрезков» Z^j («горизонтальных отрезков» Y^j) длины l или с одной из $i-1$ «точек» τ_j. В третьей строке выписано множество, в котором «вертикальный отрезок» Z^i («горизонтальный отрезок» Y^i) длины l пересекается с одним из $i-1$ «горизонтальных отрезков» Y^j («вертикальных отрезков» Z^j) длины l. В четвертой строке выписано множество, в котором «отрезок» Z^i (Y^i) длины l пересекается с одной из $2i$ точек $b\|N_j$. В пятой строке выписано множество, в котором «отрезки» Z^i и Y^i пересекаются между собой.

Таким образом, $\forall\ \Gamma^j, T_j, Y^j, Z^j, H^j$, $j=1,\ldots,i-1$, мощность множества B может быть оценена следующим образом:

$$\#B \leqslant \sum_{j=1}^{i-1}\left(\underbrace{\#\left\{Z^i : Z^i\cap Z^j\neq\emptyset\right\}}_{=(l+l-1)} + \underbrace{\#\left\{Z^i : \tau_j\in Z^i\right\}}_{=l}\right)\cdot 2^n + \sum_{j=1}^{i-1}\left(\underbrace{\#\left\{Y^i : Y^i\cap Y^j\neq\emptyset\right\}}_{=(l+l-1)} + \underbrace{\#\left\{Y^i : \tau_j\in Y^i\right\}}_{=l}\right)\cdot 2^n + \sum_{j=1}^{i-1}\sum_{k_j=1}^{l}\left(\underbrace{\#\left\{Z^i : Y^j_{k_j}\in Z^i\right\}}_{=l}\cdot 2^n + \underbrace{\#\left\{Y^i : Z^j_{k_j}\in Y^i\right\}}_{=l}\cdot 2^n\right) + \sum_{j=1}^{i}\sum_{b\in\mathbb{B}}\left(\underbrace{\#\left\{Z^i : b\|N_j\in Z^i\right\}}_{=l}\cdot 2^n + \underbrace{\#\left\{Y^i : b\|N_j\in Y^i\right\}}_{=l}\cdot 2^n\right) + \sum_{Z_i}\underbrace{\#\left\{Y^i : Y^i\cap Z^i\neq\emptyset\right\}}_{l^2} = \underbrace{\left((i-1)(6l+2l^2-2)+4il+l^2\right)}_{\omega_2^1}\cdot 2^n.$$

Подставляя полученную оценку для мощности в формулу (5), имеем

$$sum_2^1 \leqslant \omega_2^1\cdot 2^n\cdot\frac{1}{2^{n(i-1)(2l+2)+s(i-1)+2n}} \leqslant \frac{\omega_2^1}{2^{n(i-1)(2l+2)+s(i-1)+n}}. \quad (6)$$

Рассмотрим слагаемое sum_2^2. Оно отражает вероятность попадания «точки» τ_i на «отрезки» Y^i, Z^i, в «точки» $0\|N_1$ и $1\|N_1$ или в множество $\mathbf{D}^{i-1}$ (при условии,

что среди последних не было пересечений). Заметим, что после фиксации значений Γ^i, H^i случайные величины $\widetilde{\tau}_i$ принимают конкретные значения $\tau_i = \sum_{k=1}^{l} H_k^i \cdot X_k^i$, где $X^i = X_i^{\mathcal{A}}(\Gamma^1, T_1, \ldots, \Gamma^{i-1}, T_{i-1}, \Gamma^i)$. Поэтому,

$$sum_2^2 = \sum_{\substack{Y^i, Z^i: \\ \mathbf{D}^i \setminus \{\widetilde{\tau}_i\}\ \overline{coll}}} \sum_{\substack{\Gamma^i, H^i: \\ \mathbf{D}^i\ coll}} \Pr\left[\text{fixed} \cap \begin{Bmatrix} \widetilde{Y}^i = Y^i \\ \widetilde{Z}^i = Z^i \\ \widetilde{\Gamma}^i = \Gamma^i \\ \widetilde{H}^i = H^i \end{Bmatrix}\right] =$$

$$= \sum_{\substack{Y^i, Z^i: \\ \mathbf{D}^i \setminus \{\widetilde{\tau}_i\}\ \overline{coll}}} \underbrace{\#\left\{\begin{matrix}\Gamma^i, H^i: \\ \mathbf{D}^i\ coll\end{matrix}\right\}}_{C} \cdot \frac{1}{2^{n(i-1)(2l+2)+s(i-1)+n(2l+2)}}. \tag{7}$$

Оценим мощность множества C при фиксированных значениях $\Gamma^j, T_j, Y^j, Z^j, H^j$, $j = 1, \ldots, i-1$ (как следствие, для фиксированных значений N_i и всех значений из $\mathbf{D}^{i-1}$), Y^i, Z^i. Множество C принадлежит объединению следующих множеств:

$$C \subseteq \bigcup_{\Gamma^i} \left(\left\{ \begin{matrix} H^i: \\ \tau_i \in \mathbf{D}^{i-1} \end{matrix} \right\} \cup \left\{ \begin{matrix} H^i: \\ \tau_i \in Z^i \end{matrix} \right\} \cup \left\{ \begin{matrix} H^i: \\ \tau_i \in Y^i \end{matrix} \right\} \cup \bigcup_{b \in \mathbb{B}} \left\{ \begin{matrix} H^i: \\ \tau_i = b \| N_i \end{matrix} \right\} \right).$$

Здесь в первой фигурной скобке выписано множество, при котором «точка» τ_i попала в множество $\mathbf{D}^{i-1}$ мощности $(i-1)(2l+3)$. Во второй (третьей) скобке выписано множество, при котором «точка» τ_i попала на «отрезок» Z^i (Y^i) длины l. В четвертой скобке выписано множество, при котором «точка» τ_i попала в одну из двух «точек» $b \| N_i$.

Таким образом, мощность множества C может быть оценена следующим образом:

$$\#C \leqslant 2^{nl} \Big(\underbrace{\#\left\{ \begin{matrix} H^i: \\ \tau_i \in \mathbf{D}^{i-1} \end{matrix} \right\}}_{=(i-1)(2l+3)\cdot 2^{nl-n}} + \underbrace{\#\left\{ \begin{matrix} H^i: \\ \tau_i \in Z^i \end{matrix} \right\}}_{=l \cdot 2^{nl-n}} +$$
$$+ \underbrace{\#\left\{ \begin{matrix} H^i: \\ \tau_i \in Y^i \end{matrix} \right\}}_{=l \cdot 2^{nl-n}} + \sum_{b \in \mathbb{B}} \underbrace{\#\left\{ \begin{matrix} H^i: \\ \tau_i = b \| N_i \end{matrix} \right\}}_{2^{nl-n}} \Big) =$$
$$= (i-1)(2l+3) \cdot 2^{2nl-n} + (2l+2) \cdot 2^{2nl-n} =$$
$$= \underbrace{((i-1)(2l+3) + (2l+2))}_{\omega_2^2} \cdot 2^{2nl-n}.$$

Подставим полученную оценку для мощности в формулу (7):

$$sum_2^2 \leqslant \sum_{\substack{Y^i, Z^i: \\ \mathbf{D}^i \setminus \{\widetilde{\tau}_i\}\ \overline{coll}}} \frac{\omega_2^2 \cdot 2^{2nl-n}}{2^{n(i-1)(2l+2)+s(i-1)+n(2l+2)}} \leqslant$$
$$\leqslant \#\left\{Y^i, Z^i\right\} \cdot \frac{\omega_2^2}{2^{n(i-1)(2l+2)+s(i-1)+3n}} =$$
$$= 2^{2n} \cdot \frac{\omega_2^2}{2^{n(i-1)(2l+2)+s(i-1)+3n}} =$$
$$= \frac{\omega_2^2}{2^{n(i-1)(2l+2)+s(i-1)+n}}. \tag{8}$$

Просуммируем полученные оценки (6) и (8) и получим оценку для sum_2:

$$sum_2 = \sum_{\substack{Y^1, \ldots, Y^{i-1} \\ Z^1, \ldots, Z^{i-1} \\ H^1, \ldots, H^{i-1}: \\ \mathbf{D}^{i-1}\ \overline{coll} \\ 0 \| N_i, 1 \| N_i \notin \mathbf{D}^{i-1}}} \left(sum_2^1 + sum_2^2\right) \leqslant$$
$$\leqslant \#\begin{Bmatrix} Y^1, \ldots, Y^{i-1} \\ Z^1, \ldots, Z^{i-1} \\ H^1, \ldots, H^{i-1} \end{Bmatrix} \left(sum_2^1 + sum_2^2\right) =$$
$$= 2^{n(i-1)(l+2)} \cdot \frac{\omega_2^1 + \omega_2^2}{2^{n(i-1)(2l+2)+s(i-1)+n}}. =$$
$$= \frac{\omega_2^1 + \omega_2^2}{2^{n(i-1)l+s(i-1)+n}}.$$

Теперь просуммируем оценки для sum_1 (4) и sum_2:

$$\Pr\left[\widetilde{\mathbf{D}}^i\ coll \cap \widetilde{\mathbf{D}}^{i-1}\ \overline{coll}\right] = \sum_{\substack{\Gamma^1, \ldots, \Gamma^{i-1} \\ T_1, \ldots, T_{i-1}}} (sum_1 + sum_2) \leqslant$$
$$\leqslant \#\left\{\begin{matrix}\Gamma^1, \ldots, \Gamma^{i-1} \\ T_1, \ldots, T_{i-1}\end{matrix}\right\} (sum_1 + sum_2) =$$
$$= 2^{n(i-1)l} \cdot 2^{s(i-1)} \cdot \frac{\omega_1 + \omega_2^1 + \omega_2^2}{2^{n(i-1)l+s(i-1)+n}} =$$
$$= \frac{\omega_1 + \omega_2^1 + \omega_2^2}{2^n}.$$

Итого,

$$\Pr\left[\widetilde{\mathbf{D}}^q\ coll\right] = \sum_{i=1}^{q} \Pr\left[\widetilde{\mathbf{D}}^i\ coll \cap \widetilde{\mathbf{D}}^{i-1}\ \overline{coll}\right] \leqslant$$
$$\leqslant \sum_{i=1}^{q} \frac{(i-1)(4l+2)}{2^n} +$$
$$+ \sum_{i=1}^{q} \frac{(i-1)(6l+2l^2-2)+4il+l^2}{2^n} +$$
$$+ \sum_{i=1}^{q} \frac{(i-1)(2l+3)+(2l+2)}{2^n} \leqslant$$
$$\leqslant \frac{4(ql)^2 + 12q^2l + 9q^2}{2^{n+1}} = \frac{(2ql+3q)^2}{2^{n+1}}.$$

□

Теорема V.3. *Для любого противника $\mathcal{A}$ в модели mIND-CPA с параметрами l и s для семейства $Perm(n)$, который делает в точности q пар связанных запросов, верно следующее неравенство:*

$$\mathbf{Adv}^{\text{mIND-CPA}_l}_{Perm(n)}(\mathcal{A}) \leqslant \frac{(2ql+3q)^2}{2^n}.$$

Доказательство. Заметим, что до наступления события $\widetilde{\mathbf{D}}^q\ coll$ оракул $Encrypt$-1 порождает такое же распределение на ответах, что и оракул $Encrypt$-0.

Тогда, согласно Лемме 2, [5], верно следующее соотношение:

$$\mathbf{Adv}^{\text{mIND-CPA}_{l,s}}_{Func(n)}(\mathcal{A}) \leqslant \Pr\left[\widetilde{\mathbf{D}}^q\ coll\right].$$

Используя Лемму «PRP-PRF Switching», см. [7], и Лемму V.1, мы получаем следующую оценку:

$$\mathbf{Adv}^{\text{mIND-CPA}_{l,s}}_{Perm(n)}(\mathcal{A}) \leqslant \mathbf{Adv}^{\text{mIND-CPA}_{l,s}}_{Func(n)}(\mathcal{A}) + \frac{(2ql+3q)^2}{2^{n+1}} \leqslant$$
$$\leqslant \Pr\left[\widetilde{\mathbf{D}}^q\ coll\right] + \frac{(2ql+3q)^2}{2^{n+1}} \leqslant \frac{(2ql+3q)^2}{2^{n+1}} +$$
$$+ \frac{(2ql+3q)^2}{2^{n+1}} \leqslant \frac{(2ql+3q)^2}{2^n}.$$

Второе неравенство верно в силу того, что в рамках эксперимента в модели mIND-CPA$_l$ противник делает не более $2ql + 3q$ запросов к функции f. □

VI. Заключение

В настоящей работе получена и доказана оценка уровня стойкости режима MGM в модели противника IND-CPA, в которой для обработки каждого сообщения используется уникальное значение вектора инициализации. Данная оценка демонстрирует, что режим MGM обеспечивает конфиденциальность при обработке данных, суммарный объем которых не превышает $2^{n/2}$ блоков. Данный уровень стойкости является стандартным для AEAD-режимов.

Отметим, что особенностью конструкции режима MGM является возможность возникновения коллизий между любыми возможными входами блочного шифра (вероятность данного события оценивалась в Лемме V.1). Данная особенность приводит к ухудшению уровня стойкости, что было показано в работе [8]. В ней была представлена атака на свойство целостности, существенно использующая возможность коллизий между входами.

Библиография

БИБЛИОГРАФИЯ

[1] E. Rescorla, «The Transport Layer Security (TLS) Protocol Version 1.3», *RFC 8446*, 2018.

[2] V. Nozdrunov, «Parallel and double block cipher mode of operation (PD-mode) for authenticated encryption», *In Proceedings of 6th Workshop on Current Trends in Cryptology (CTCrypt 2017)*, 2017.

[3] P. Rogaway, «Nonce-Based Symmetric Encryption», *Fast Software Encryption. FSE 2004. Lecture Notes in Computer Science*, т. 3017, 2004.

[4] «Информационная технология. Криптографическая защита информации. Режимы работы блочных шифров, реализующие аутентифицированное шифрование», *Федеральное агентство по техническому регулированию и метрологии (Росстандарт)*, 2019.

[5] M. Bellare и P. Rogaway, «The Security of Triple Encryption and a Framework for Code-Based Game-Playing Proofs», *Advances in Cryptology — EUROCRYPT 2006. Lecture Notes in Computer Science*, т. 4004, 2006.

[6] M. Bellare и C. Namprempre, «Authenticated encryption: Relations among notions and analysis of the generic composition paradigm», *Advances in Cryptology — ASIACRYPT 2000. ASIACRYPT 2000. Lecture Notes in Computer Science*, т. 1976, 2000.

[7] D. Chang и M. Nandi, «A Short Proof of the PRP/PRF Switching Lemma», *IACR Cryptology ePrint Archive*, т. 2008/078, 2008.

[8] A. Kurochkin и D. Fomin, «MGM Beyond the Birthday Bound», *8th Workshop on Current Trends in Cryptology (CTCrypt 2019)*, 2019.

On confidentiality property of MGM AEAD-mode

Liliya Akhmetzyanova

***Abstract*—In this paper the security of AEAD mode called the Multilinear Galois Mode (MGM) was analyzed regarding confidentiality property. This mode was originally proposed in CTCrypt 2017. Then it was adopted as a standard AEAD mode in the Russian Standardization system. The MGM plaintext encryption procedure is quite similar to encryption in the counter mode. The main element of the MGM authentication procedure is a multilinear function with secret coefficients produced in the same way as the secret masking blocks used for plaintext encryption. This construction allows keeping such advantages as parallelization, online and availability of precomputations.**

The report presented at the conference outlined the design principles of the MGM mode from the point of view of providing security. The analysis of the MGM mode was carried out in the paradigm of provable security, in other words, lower security bound was obtained for the IND-CPA notion as a function of the mode parameters and amount of data available to an adversary. This bound shows that the privacy of this mode is provably guaranteed (under security of the used block cipher) up to the birthday paradox bound.

***Keywords*—MGM, AEAD mode, security notion, security bounds, confidentiality**

References

REFERENCES

[1] E. Rescorla, «The Transport Layer Security (TLS) Protocol Version 1.3», *RFC 8446*, 2018.

[2] V. Nozdrunov, «Parallel and double block cipher mode of operation (PD-mode) for authenticated encryption», *In Proceedings of 6th Workshop on Current Trends in Cryptology (CTCrypt 2017)*, 2017.

[3] M. Bellare and C. Namprempre, «Authenticated encryption: Relations among notions and analysis of the generic composition paradigm», *Advances in Cryptology — ASIACRYPT 2000. ASIACRYPT 2000. Lecture Notes in Computer Science*, vol. 1976, 2000.

[4] «Information technology. Cryptographic data security. Authenticated encryption block cipher operation modes», *Federal Agency on Technical Regulating and Metrology*, 2019.

[5] P. Rogaway, «Nonce-Based Symmetric Encryption», *Fast Software Encryption. FSE 2004. Lecture Notes in Computer Science*, vol. 3017, 2004.

[6] M. Bellare and P. Rogaway, «The Security of Triple Encryption and a Framework for Code-Based Game-Playing Proofs», *Advances in Cryptology — EUROCRYPT 2006. Lecture Notes in Computer Science*, vol. 4004, 2006.

[7] D. Chang and M. Nandi, «A Short Proof of the PRP/PRF Switching Lemma», *IACR Cryptology ePrint Archive*, vol. 2008/078, 2008.

[8] A. Kurochkin and D. Fomin, «MGM Beyond the Birthday Bound», *8th Workshop on Current Trends in Cryptology (CTCrypt 2019)*, 2019.