Эффективная структурная атака на криптосистему Мак-Элиса–Сидельникова

И. В. Чижов, С. А. Конюхов, А. М. Давлетшина

***Аннотация*—В работе предлагается алгоритм восстановления секретного ключа криптосистемы Мак-Элиса–Сидельникова в общем виде: с $u \in \mathbb{N}$ копиями кода Рида–Маллера. Задача восстановления секретного ключа криптосистемы Мак-Элиса–Сидельникова сводится к u задачам восстановления секретного ключа криптосистемы Мак-Элиса, построенной на кодах Рида–Маллера. Доказано, что предложенная атака является полиномиальной. Описано множество ключей, для которых алгоритм применим. Данному множеству дано название множество слабых ключей. Авторы считают, что подавляющее число ключей криптосистемы являются слабыми. Приведена мотивация, в соответствии с которой следует предполагать, что доля слабых ключей в пространстве ключей криптосистемы близка к единице. Описаны методы подсчёта числа слабых ключей и проведены вычислительные эксперименты подтверждающие это предположение.**

***Ключевые слова*—криптосистема Мак-Элиса, коды Рида–Маллера, постквантовая криптография, кодовые криптосистемы, полиномиальная атака.**

I. Введение

В 1978 году Робертом Мак-Элисом была предложена криптосистема с открытым ключом [1]. Главная её особенность состоит в том, что стойкость криптосистемы основана на сложности некоторых задач из теории кодов, исправляющих ошибки, а не, например, на сложности теоретико-числовых задач факторизации или дискретного логарифмирования. В связи с этим эта криптосистема устойчива к атакам, использующим алгоритм Шора [2], и поэтому представляет интерес в контексте возможности создания квантового компьютера.

Исследования криптосистемы Мак-Элиса важны в силу широкого распространения технологии «Блокчейн», в которой используются схемы подписи и протоколы обмена ключами для создания и хранения записей. Причём стойкость этих криптографических механизмов определяет надёжность всей технологии и доверие к динамически развивающейся области криптоэкономики. После появления квантового компьютера с большим количеством кубитов все криптографические валюты мира, которые используют механизмы на основе задач факторизации чисел и дискретного логарифмирования в циклической группе, станут ненадёжными. Поэтому разработка постквантовых криптографических механизмов, т. е. таких механизмов, которые останутся стойкими против атак с использованием квантового компьютера, позволит сохранить криптовалюты и технологию «Блокчейн».

В 2019 году интерес к постквантовым криптографическим механизмам был подогрет Национальным агентством стандартов и технологий США (NIST USA), которое объявило конкурс на серию стандартов постквантовой криптографии. Теория кодов, исправляющих ошибки, является одним из основных источников конструкций схем постквантовой подписи и протоколов передачи секретных ключей.

В алгоритме шифрования криптосистемы Мак-Элиса используются случайные векторы некоторого фиксированного веса, а значит, необходимо наличие хорошего датчика случайных чисел. В работе [1] рекомендовалось использовать коды Гоппы. В 1986 году Г. Нидеррайтер [3] предложил эквивалентный вариант криптосистемы Мак-Элиса, в котором нет необходимости в датчике случайных чисел. Однако в этой криптосистеме ниже скорость передачи, то есть отношение длины открытого текста к длине шифртекста.

В 1994 году В. М. Сидельников предложил использовать коды Рида–Маллера [4], так как это увеличило бы скорость передачи криптосистемы. Кроме того, в 2001 году группа французских исследователей на основе конструкции Нидеррайтера предложила схему постквантовой подписи CFS [5]. Оригинальная схема подписи строится на двоичных кодах Гоппы. Однако именно на кодах Рида–Маллера эта схема подписи будет иметь достаточно быстрый алгоритм формирования подписи, что делает такой вариант привлекательным для использования в различных приложениях.

В течении продолжительного времени не существовало эффективных атак на криптосистему Мак-Элиса, построенную на кодах Рида–Маллера. Ситуация изменилась в 2007 году, когда Л. Миндер и А. Шокроллахи построили на неё структурную атаку [6].

Окончательно вопрос о состоятельности предложения В. М. Сидельникова был решён в 2014 году. И. Чижов и М. Бородин построили атаку, которая для широкого класса параметров выполняется за полиномиальное число операций, а для оставшихся самая трудоёмкая операция — применение лишь одного шага из атаки Л. Миндера и А. Шокроллахи [7].

Вместе с тем в работе [4] предлагается более общая конструкция, которая использует не одну, а $u \in \mathbb{N}, u > 1$, копий кода Рида–Маллера. Далее эта конструкция будет называться криптосистемой Мак-Элиса–Сидельникова. На эту схему известные атаки Миндера–Шокроллахи и Чижова–Бородина уже работать не будут. Таким образом, до настоящего момента криптосисте-

Статья получена 29 июня 2020.

Иван Владимирович Чижов, МГУ им. М. В. Ломоносова, Федеральный исследовательский центр «Информатика и управление» РАН, АО «НПК «Криптонит» (email: ichizhov@cs.msu.ru).

Сергей Андреевич Конюхов, МГУ им. М. В. Ломоносова, ОАО «ИнфоТеКС» (email: koniukhov.serge@gmail.com).

Александра Маратовна Давлетшина, МГУ им. М. В. Ломоносова, ОАО «ИнфоТеКС» (email: sdav94@rambler.ru).

Работа частично поддержана грантом РФФИ №18-29-03124мк

ма Мак-Элиса–Сидельникова, как и соответствующая ей схема подписи CFS, могут рассматриваться в качестве кандидата в постквантовые криптографические алгоритмы. А значит она потенциально может заменить собой схемы, использующиеся во всех реализациях технологии «Блокчейн».

В настоящей работе строится эффективная атака на криптосистему Мак-Элиса–Сидельникова, в случае применения ключей специального вида.

В атаке из работы [7] в качестве базового инструмента используется операция произведения Адамара линейных кодов. Попытки использовать эту операцию в криптоанализе кодовых криптосистем можно встретить, например, в работе [8]. Также, используя произведение Адамара кодов, были достигнуты определённые успехи в анализе криптосистемы BBCRS [9]. Произведение Адамара кодов, и, в частости, квадрат Адамара кода, представляет большой исследовательский интерес. Так, в работе [10] рассмотрены вероятностные оценки размерности квадрата случайных кодов, исправляющих ошибки.

Отметим, что в работе [11], техника с использованием произведения Адамара применяется и к анализу криптосистемы Мак-Элиса, построенной на подкодах кодов Рида–Маллера, хоть и маленького коранга.

При построении структурных атак на криптосистему Мак-Элиса–Сидельникова, например, в работе [12], где была описана атака для специального подмножества ключей, стала очевидной необходимость исследования квадрата конкатенированного кода Рида–Маллера.

Основным результатом настоящей работы является расширение области применимости атаки по сравнению с существующей атакой [12]. Новая атака применима для всего множества доменных параметров криптосистемы. Кроме того, вводится понятие слабого ключа, как ключа, уязвимого к предложенному алгоритму атаки. Приводятся методы оценки количества слабых ключей. Проведены вычислительные эксперименты, дающие основания полагать, что множество слабых ключей представляет собой подавляющее большинство во множестве всех возможных ключей изучаемой криптосистемы.

II. Криптосистема Мак-Элиса–Сидельникова

Обозначим через V_n линейное векторное пространство длины n над полем $\mathbb{F}_2$. *Двоичным линейным* $[n,k]$*-кодом* называется линейное подпространство C пространства V_n размерности k. Матрица G размера $k \times n$, состоящая из базисных векторов C, называется *порождающей* матрицей кода C.

Под *кодовым расстоянием* линейного кода C обычно понимают минимальный вес ненулевых кодовых слов.

Дуальным кодом $\mathcal{C}^\perp$ к коду $\mathcal{C}$ с порождающей матрицей G называется линейный $[n, n-k]$-код, состоящий из векторов $x \in V_n$, для которых верно равенство $Gx^T = 0$.

Будем обозначать результат применения подстановки $\sigma \in S_n$ к вектору a длины n как a^σ, т. е. если $a = (a_1, \ldots, a_n)$, то $a^\sigma = (a_{\sigma(1)}, \ldots, a_{\sigma(n)})$. Далее, результат применения этой же подстановки к каждому вектору из некоторого множества векторов U длины n представляет собой множество $X^\sigma = \{x^\sigma | x \in X\}$. Очевидно, что X^σ является линейным кодом, если и только если X — линейный код.

Два линейных кода $\mathcal{A}$ и $\mathcal{B}$ длины n называются эквивалентными, если существует $\sigma \in S_n$ такая, что $\mathcal{A} = \mathcal{B}^\sigma$. Эквивалентные коды имеют одинаковую размерность, и одинаковое кодовое расстояние.

Так как в работе особое внимание уделяется кодам Рида–Маллера, приведём их определение.

Определение 1 ([13])**.** Для произвольного r, $0 \le r \le m$, *двоичный код Рида–Маллера* $\mathrm{RM}(r,m)$ порядка r и длины 2^m определяется как множество всех векторов-значений булевых функций $f(v_1, \ldots, v_m)$, задаваемых многочленом Жегалкина, степень которого не превосходит r. Известно [13], что код $\mathrm{RM}(r,m)$ имеет размерность $k = \sum_{i=0}^{r} \binom{m}{i}$ и кодовое расстояние $d = 2^{m-r}$.

Прежде чем перейти к описанию криптосистемы Мак-Элиса–Сидельникова и постановке задачи криптоаналитика приведём определение конкатенированного кода.

Определение 2. Пусть $\mathcal{C}$ — некоторый $[n,k]$-код с порождающей матрицей G, пусть также Δ — упорядоченный кортеж $(H_1, \ldots, H_{u-1})$ невырожденных $(k \times k)$-матриц. Тогда $[u \cdot n, k]$-код $\mathcal{C}_\Delta$ с порождающей матрицей вида

$$G_\Delta = (G \| H_1 G \| \ldots \| H_{u-1} G)$$

будем называть *конкатенированным кодом*, построенным на *базовом* коде C.

Договоримся в дальнейшем, что в случае, когда длина Δ равна 1, т. е. если $\Delta = (H)$, обозначать код C_Δ просто как C_H.

Криптосистема Мак-Элиса–Сидельникова — криптосистема с открытым ключом, предложенная В. М. Сидельниковым в работе [4]. Параметры криптосистемы $u \in \mathbb{N}, u \ge 2$, $m \in \mathbb{N}$, $r \in \mathbb{N}, 0 \le r \le m$ являются доменными и не включаются ни в закрытый, ни в открытый ключ.

Секретный ключ: набор матриц $\langle M_1, \ldots, M_u, P_\sigma \rangle$.
Абонент выбирает u невырожденных $(k \times k)$-матриц $M_1, \ldots, M_u$, и подстановку $\sigma \in S_{u \cdot n}$. Набор матриц $\langle M_1, \ldots, M_u, P_\sigma \rangle$ объявляется секретным ключом, здесь P_σ — перестановочная матрица размера $un \times un$, соответствующая подстановке $\sigma \in S_{un}$.

Открытый ключ: матрица G.
По закрытому ключу $\langle M_1, \ldots, M_u, P \rangle$ абонент строит матрицу:

$$G = (M_1 R \| \ldots \| M_u R) P,$$

где R — порождающая матрица кода $\mathrm{RM}(r,m)$. Матрица G объявляется открытым ключом.

В текущей работе операции шифрования и расшифрования никак не задействованы, поэтому опустим их описание, подробности см. [4].

Сформулируем теперь задачу криптоаналитика.

Задача криптоаналитика — найти по открытому ключу G такие матрицы $M_1, \ldots, M_u$, P, что будет выполнено равенство $G = (M_1 R \| \ldots \| M_u R) P$.

Матрица M_1 является невырожденной, поэтому последнее соотношение можно переписать в виде

$$G \cdot P^{-1} = M_1 \cdot (R \| M_1^{-1} M_2 R \| \| \ldots \| M_1^{-1} M_u R).$$

Матрица $M_1 \cdot (R \| M_1^{-1} M_2 R \| \| \dots \| M_1^{-1} M_u R)$ порождает некоторый код $RM_\Delta(r,m)$ для $\Delta = (H_1 = M_1^{-1}M_2, \dots, H_{u-1} = M_1^{-1}M_u)$. Так как матрица P является перестановочной, то такой будет и матрица P^{-1}. Пусть ей соответствует подстановка $\gamma \in S_{un}$. Тогда GP^{-1} порождает код $\mathcal{C}^\gamma$, если G порождает код $\mathcal{C}$.

Это рассуждение позволяет сформулировать следующую *кодовую* задачу криптоаналитика, полностью эквивалентную исходной задаче криптоаналитика.

Кодовая задача криптоаналитика — по коду $\mathcal{C}$, заданному порождающей матрицей G, являющейся открытым ключом, найти кортеж $\Delta = (H_1, \dots, H_{u-1})$ невырожденных двоичных матриц и подстановку $\gamma \in S_{u \cdot n}$, что выполнено следующее равенство кодов

$$\mathcal{C}^\gamma = RM_\Delta(r,m).$$

Далее в работе предложено решение кодовой задачи криптоаналитика в случае произвольного значения параметра $u > 1$. В конце работы даётся методика получения оценок количества ключей, для которых применима атака. Проведены вычислительные эксперименты.

Техника построения атаки существенно опирается на идеи работы [7].

III. Криптоанализ схемы

Основным инструментом атаки является произведение Адамара линейных кодов. Перейдём к описанию этой операции.

Определение 3. Пусть $x = (x_1, \dots, x_n)$ и $y = (y_1, \dots, y_n)$ — векторы над полем $\mathbb{F}_2$. Тогда произведением Адамара векторов x и y называется вектор

$$x \star y = (x_1 \cdot y_1, x_2 \cdot y_2, \dots, x_n \cdot y_n).$$

Замечание. В литературе так же встречается термин произведение Шура.

В силу того, что произведение Адамара применяется к векторам покоординатно, и, на основании свойств операций в поле, равенство

$$(x+y) \star z = z \star (x+y) = x \star z + y \star z$$

выполнено для любых векторов $x, y, z \in V_n$.

Произведение Адамара векторов можно обобщить на случай линейных кодов следующим образом.

Определение 4 ([7])**.** Пусть $\mathcal{A}$ и $\mathcal{B}$ — линейные коды длины n с базисами $\langle a_1, \dots, a_{k_1} \rangle$ и $\langle b_1, \dots, b_{k_2} \rangle$ соответственно. Тогда назовём произведением Адамара линейных кодов $\mathcal{A}$ и $\mathcal{B}$ линейный код, который является линейной оболочкой векторов $\{a_i \star b_j | 1 \le i \le k_1, 1 \le j \le k_2\}$; обозначим его $\mathcal{A} \star \mathcal{B}$. Для $q \in \mathbb{N}$ введём обозначение $\mathcal{A}^{\star q} = \underbrace{\mathcal{A} \star \dots \star \mathcal{A}}_{q \text{ множителей}}$, кроме того, будем считать, что $\mathcal{A}^{\star 0} = \mathcal{A}$.

В силу свойств произведения Адамара векторов, произведение Адамара кодов обладает свойством коммутативности с операцией суммы кодов (сумма линейных подпространств). Кроме того, для любых кодов $\mathcal{A}, \mathcal{B}$ и $\mathcal{C}$ длины n справедлив закон дистрибутивности

$$(\mathcal{A} + \mathcal{B}) \star \mathcal{C} = \mathcal{C} \star (\mathcal{A} + \mathcal{B}) = \mathcal{A} \star \mathcal{C} + \mathcal{B} \star \mathcal{C}.$$

В завершении сформулируем утверждение о применении к произведению Адамара кодов подстановки.

Утверждение 1. *Для кодов $\mathcal{A}$ и $\mathcal{B}$ одной длины n и подстановки $\sigma \in S_n$ выполнено:*

1) $\mathcal{A}^\sigma \star \mathcal{B}^\sigma = (\mathcal{A} \star \mathcal{B})^\sigma$,
2) $(\mathcal{A}^\sigma)^\perp = (\mathcal{A}^\perp)^\sigma$.

Интересующие нас свойства линейных кодов, которые будут приведены далее, сформулированы с использованием такого понятия, как декартова степень кода.

Определение 5. Декартовой u-степенью $[n,k]$-кода $\mathcal{C}$ для некоторого натурального числа u называется $[un, uk]$-код $\mathcal{C}^{\times u}$, определяемый следующим образом:

$$\mathcal{C}^{\times u} = \{(c_1 \| \dots \| c_u) | \forall\ c_i \in C,\ i = 1, 2, \dots, u\}.$$

Заметим, что код $C^{\times u}$ порождается блочно–диагональной $(uk \times un)$-матрицей, у которой на диагонали стоит порождающая $(k \times n)$-матрица кода C.

A. Свойства произведения Адамара различных кодов

Сформулируем и докажем группу утверждений и следствий, необходимых для построения атаки на криптосистему.

Утверждение 2. *Пусть $\mathcal{A}$ и $\mathcal{B}$ — линейные коды длины n, тогда*

$$\mathcal{A}^{\times u} \star \mathcal{B}^{\times u} = (\mathcal{A} \star \mathcal{B})^{\times u}.$$

Доказательство. Пусть A — порождающая матрица кода $\mathcal{A}$ некоторой размерности k. Обозначим за $\mathcal{A}_i^{\times u}$, $0 \le i \le u$, код с порождающей матрицей вида

$$(\underbrace{\Theta \| \dots \| \Theta}_{i-1} \| A \| \underbrace{\Theta \| \dots \| \Theta}_{u-i}),$$

где Θ — нулевая $(k \times n)$-матрица. Ясно, что код $\mathcal{A}^{\times u}$ может быть выражен через сумму кодов:

$$\mathcal{A}^{\times u} = \mathcal{A}_1^{\times u} + \dots + \mathcal{A}_u^{\times u},$$

Очевидно, что код $\mathcal{A}_i^{\times u} \star \mathcal{B}_j^{\times u}$ состоит лишь из нулевого кодового слова при $i \ne j$. В силу базовых свойств произведения Адамара справедливо разложение

$$\mathcal{A}^{\times u} \star \mathcal{B}^{\times u} = \sum_{i=1}^{u} \mathcal{A}_i^{\times u} \star \mathcal{B}_i^{\times u}.$$

При этом порождающая матрица кода $\mathcal{A}_i^{\times u} \star \mathcal{B}_i^{\times u}$ имеет вид

$$(\underbrace{\Theta \| \dots \| \Theta}_{i-1} \| C \| \underbrace{\Theta \| \dots \| \Theta}_{u-i}),$$

где C — порождающая матрица кода $\mathcal{A} \star \mathcal{B}$. Таким образом, порождающая матрица кода $\mathcal{A}^{\times u} \star \mathcal{B}^{\times u}$ является блочно–диагональной матрицей с порождающими матрицами $\mathcal{A} \star \mathcal{B}$ на диагонали. □

Утверждение 3. *Пусть $\mathcal{A}$ и $\mathcal{B}$ — линейные коды длины n. Тогда для любого Δ длины $u-1$ справедливо свойство «поглощения»*

$$\mathcal{A}^{\times u} \star \mathcal{B}_\Delta = (\mathcal{A} \star \mathcal{B})^{\times u}.$$

Доказательство. По аналогии с тем, как это было сделано в утверждении 2, запишем

$$\mathcal{A}^{\times u} \star \mathcal{B}_\Delta = \bigoplus_{i=1}^{u} \mathcal{A}_i^{\times u} \star \mathcal{B}_\Delta.$$

При этом порождающая матрица кода $\mathcal{A}_i^{\times u} \star \mathcal{B}_\Delta$ имеет вид

$$(\underbrace{\Theta\| \cdots \|\Theta}_{i-1} \|C\| \underbrace{\Theta\| \cdots \|\Theta}_{u-i}),$$

где C — порождающая матрица кода $\mathcal{A} \star \mathcal{B}$. Это означает, что порождающая матрица кода $\mathcal{A}^{\times u} \star \mathcal{B}_\Delta$ является блочно–диагональной матрицей с порождающими матрицами $\mathcal{A} \star \mathcal{B}$ на диагонали. □

Утверждение 4. *Для любого линейного кода $\mathcal{A}$ выполнено*

$$\left(\mathcal{A}^{\times u}\right)^{\perp} = \left(\mathcal{A}^{\perp}\right)^{\times u}.$$

Доказательство. Рассмотрим блочно–диагональную матрицу с порождающими матрицами $\mathcal{A}^{\perp}$ на диагонали. Очевидно, что код с такой порождающей матрицей является подкодом $(\mathcal{A}^{\times u})^{\perp}$. При этом его размерность совпадает с размерностью указанного кода, а значит эти коды совпадают. □

Приведённые утверждения могут быть применены к кодам Рида–Маллера, которые представляют для нас наибольший интерес. Для этого нужно вспомнить следующие свойства кодов Рида–Маллера.

Утверждение 5 ([13])**.** *Для всех $0 \le r_{1,2} \le m$ произведение Адамара кодов* $\mathrm{RM}(r_1, m)$ *и* $\mathrm{RM}(r_2, m)$ *— код Рида–Маллера порядка* $\min(r_1 + r_2, m)$*, то есть*

$$\mathrm{RM}(r_1, m) \star \mathrm{RM}(r_2, m) = \mathrm{RM}(\min(r_1 + r_2, m), m).$$

Утверждение 6 ([13])**.** *Для всех $0 \le r \le m$ код* $\mathrm{RM}(m - r - 1, m)$ *дуален коду* $\mathrm{RM}(r, m)$*, то есть*

$$\mathrm{RM}^{\perp}(r, m) = \mathrm{RM}(m - r - 1, m).$$

Перейдём к следствиям из доказанных утверждений, которые проливают свет на структуру конкатенированных кодов Рида–Маллера.

Следствие 1. *Для всех $0 \le r_{1,2} \le m$ произведение Адамара кодов, являющихся декартовым произведением кода Рида–Маллера порядка r_1 и r_2 соответственно, есть декартово произведение кода Рида–Маллера порядка* $\min(r_1 + r_2, m)$*, то есть*

$$\mathrm{RM}^{\times u}(r_1, m) \star \mathrm{RM}^{\times u}(r_2, m) = \mathrm{RM}^{\times u}(\min(r_1 + r_2, m), m).$$

Доказательство. Непосредственно применим утверждения 2 и 5. □

Утверждение 7. *Для всех $0 \le r_{1,2} \le m$ и любого Δ длины u*

$$\mathrm{RM}^{\times u}(r_1, m) \star \mathrm{RM}_\Delta(r_2, m) = \mathrm{RM}^{\times u}(\min(r_1 + r_2, m), m).$$

Доказательство. Непосредственно применим утверждения 3 и 5. □

Утверждение 8. *Для всех $0 \le r \le m$ код* $\mathrm{RM}^{\times u}(m - r - 1, m)$ *дуален коду* $\mathrm{RM}^{\times u}(r, m)$*, то есть*

$$\left(\mathrm{RM}^{\times u}\right)^{\perp}(r, m) = \mathrm{RM}^{\times u}(m - r - 1, m).$$

Доказательство. Непосредственно применим утверждения 4 и 6. □

B. Атака на криптосистему

Определение 6. Пусть $\Phi = \{x, x \star y, x^{\perp}\}$, где x, y — символы переменных. Множество формул над Φ описывается следующими условиями.

1) Элемент $\varphi \in \Phi$ есть формула над Φ глубины 1.
2) Пусть $\varphi(x_1, \ldots, x_p)$ — формула над Φ глубины s. Тогда $x_{p+1} \star \varphi$ и $\varphi^{\perp}$, где x_{p+1} — символ переменной, являются формулами над Φ глубины $s + 1$.
3) Других формул нет.

Определение 7. Пусть $\mathcal{C}$ — линейный код. Замыканием $[\mathcal{C}]$ кода $\mathcal{C}$ называется множество всех кодов, которые могут представлены в виде $\varphi(\mathcal{C}, \ldots, \mathcal{C})$, где φ — формула над Φ, то есть

$$[\mathcal{C}] = \{\varphi(\mathcal{C}, \ldots, \mathcal{C}) | \varphi \in \Phi\}.$$

Теорема 1. *Пусть $0 < r < \frac{m}{2}$, а кортеж Δ длины u таков, что выполнено равенство*

$$\mathrm{RM}_\Delta^{\star 2}(r, m) = \mathrm{RM}^{\times u}(2r, m).$$

Тогда существует формула $\varphi(x)$ глубины $O(r \log_2 m)$, что

$$\mathrm{RM}^{\times u}(0, m) = \varphi\left(\mathrm{RM}_\Delta(r, m)\right).$$

Доказательство. При построении формулы требуется различать два случая: когда r делит m, и когда нет. И в том, и в другом случае применяются операции над кодом, которые могут быть записаны в виде формулы над Φ.

Пусть r делит m, и $m = qr$. Заметим, что $q > 2$, так как по условию $m > 2r$. Код $\mathrm{RM}^{\times u}(0, m)$ можно получить, выполнив следующие действия:

1) $\mathrm{RM}_\Delta^{\star(q-1)}(r, m) = \mathrm{RM}^{\times u}((q-1)r, m)$. Действительно, пусть $q - 1 = 2q' + \varepsilon$, где $\varepsilon \in \{0, 1\}$. Тогда

   $$\mathrm{RM}_\Delta^{\star(q-1)}(r, m) = \left(\mathrm{RM}_\Delta^{\star 2}(r, m)\right)^{\star q'} \star \mathrm{RM}_\Delta^{\varepsilon}(r, m).$$

   По условию $\mathrm{RM}_\Delta^{\star 2}(r, m) = RM^{\times u}(2r, m)$. Тогда на основании утверждения 7

   $$\left(\mathrm{RM}_\Delta^{\star 2}(r, m)\right)^{\star q'} = \mathrm{RM}^{\times u}(2q'r, m).$$

   Если $\varepsilon = 0$, то $2q' = q - 1$. В случае, когда $\varepsilon = 1$, применяя утверждения 3 и 7, получим требуемый код.
2) $\left(\mathrm{RM}^{\times u}((q-1)r, m)\right)^{\perp} = \mathrm{RM}^{\times u}(r-1, m)$. Справедливость этого соотношения следует из утверждения 8 и того, что $m = qr$ и $m - (q-1)r - 1 = m - qr + r - 1 = r - 1$.
3) $\mathrm{RM}^{\times u}(r-1, m) \star \mathrm{RM}^{\times u}((q-1)r, m) = \mathrm{RM}^{\times u}(m - 1, m)$. Следует из утверждения 7 с учётом равенства $m = qr$.
4) $(\mathrm{RM}^{\times u}(m-1, m))^{\perp} = \mathrm{RM}^{\times u}(0, m)$. Следует из утверждения 8.

Глубина подформулы в таком случае $O(\log_2(q-1)) + 3 = O(\log_2 m)$, т. к. для возведения в степень можно использовать, например, метод аддитивных цепочек.

Пусть теперь $m = qr + p$, где $p \in \{1, \ldots, r-1\}$. Код $\mathrm{RM}^{\times u}(p-1, m)$ можно получить, выполнив следующие действия:

1) $(\mathrm{RM}_\Delta(r, m))^{\star q} = \mathrm{RM}^{\times u}(qr, m)$,
2) $(\mathrm{RM}^{\times u}(qr, m))^{\perp} = \mathrm{RM}^{\times u}(p-1, m)$.

Положим $r' \leftarrow p - 1 < r$. Если $r > 0$, то повторим алгоритм для нового r. Глубина подформулы в таком случае $O(\log_2 q) + 1 = O(\log_2 m)$.

Обратим внимание на то, что последовательность значений параметра r' строго убывает, так как $r' = p - 1 < p < r$. Это значит, что последовательность рекурсивного построения формулы конечна. Из этого же соотношение видно, что количество рекурсивных построений в худшем случае равно r. Таким образом, глубина формулы равна $O(r \log_2 m)$. □

Доказательство теоремы 1 является конструктивным и на его основе формулируется алгоритм 1, который по коду $\mathrm{RM}^{\sigma}_{\Delta}(r, m)$ строит код $\left(\mathrm{RM}^{\times u}(0, m)\right)^{\sigma}$.

Algorithm 1 Построение кода $\left(\mathrm{RM}^{\times u}(0, m)\right)^{\sigma}$

Input: Порождающая матрица кода $\mathrm{RM}^{\sigma}_{\Delta}(r, m)$.
Output: Порождающая матрица кода $\left(\mathrm{RM}^{\times u}(0, m)\right)^{\sigma}$.
  $\mathcal{C} \leftarrow \mathrm{RM}^{\sigma}_{\Delta}(r, m)$
  while $r \neq 0$ **and** r не делит m **do**
    $q \leftarrow \lfloor \frac{m}{r} \rfloor$, $r \leftarrow m - qr - 1$
    $\mathcal{C} \leftarrow (\mathcal{C}^{\star q})^{\perp}$
  end while
  if $r = 0$ **then**
    return G
  end if
  $q \leftarrow \frac{m}{r} - 1$
  $\mathcal{C} \leftarrow (\mathcal{C}^{\star(q-1)} \star (\mathcal{C}^{\star q})^{\perp})^{\perp}$
  return $\mathcal{C}$

Корректность алгоритма 1 следует из теоремы 1 с учётом того, что для любой формулы $\varphi(x) \in \Phi$, любого линейного кода $\mathcal{C}$ длины n и любой подстановки $\sigma \in S_n$ верно равенство

$$\varphi\left(\mathcal{C}^{\sigma}\right) = \mathcal{B}^{\sigma},$$

если $\varphi(\mathcal{C}) = \mathcal{B}$.

Получив порождающую матрицу кода $\left(\mathrm{RM}^{\times u}(0, m)\right)^{\sigma}$, можно классифицировать столбцы открытого ключа схемы по принадлежности к соответствующим копиям кода Рида–Маллера. Как это сделать?

Код $\mathrm{RM}(0, m)$ содержит единственное ненулевое кодовое слово, у которого все координаты равны 1. Значит код $\mathrm{RM}^{\times u}(0, m)$ будет иметь блочно-диагональную порождающую $(u \times un)$-матрицу G, у которой на диагонали стоят векторы длины n, состоящие из одних единиц:

$$G = \left(\begin{array}{ccc|ccc|c|ccc} 1 & \dots & 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & \dots & 0 & 1 & \dots & 1 & \dots & 0 & \dots & 0 \\ \vdots & \dots & \vdots & \vdots & \dots & \vdots & \dots & \vdots & \dots & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 & \dots & 1 & \dots & 1 \end{array}\right)$$

Ясно, что для получения порождающей матрицы такого вида из произвольной порождающей матрицы кода $\mathrm{RM}^{\times u}(0, m)$, достаточно выполнить над этой матрицей прямой и обратный ход Гауссова исключения.

Используя матрицу G, можно предложить следующий принцип разбиения координат: для i-ой строки матрицы G, где $i \in \{1, \dots, u\}$, множество координат $\{j | G_{i,j} = 1\}$ относится к одному сегменту и однозначно определяет координаты, в которых стоит одна копия кода $\mathrm{RM}(r, m)$.

Algorithm 2 Построение перестановочной матрицы P_0.

Input: Порождающая матрица G_0 кода $\left(\mathrm{RM}^{\times u}(0, m)\right)^{\sigma}$.
Output: Перестановочная матрица P.
  $G \leftarrow GaussianElimination(G_0)$
  $P \leftarrow 0 \in \mathbb{F}_2^{un \times un}$, $counter \leftarrow 0 \in \mathbb{N}^u$
  for $0 \le i \le u - 1$ **do**
    $counter[i] \leftarrow n \cdot i$
  end for
  for $0 \le j \le un - 1$ **do**
    for $0 \le i \le u - 1$ **do**
      if $G_{i,j} = 1$ **then**
        $P_{j, counter[i]} \leftarrow 1$
        $counter[i] \leftarrow counter[i] + 1$
      end if
    end for
  end for
  return P

Опишем теперь схему атаки при выполнении условия теоремы 1.

1. Используя алгоритм 1, построить порождающую матрицу G_0 кода $\left(\mathrm{RM}^{\times u}(0, m)\right)^{\sigma}$.
2. По матрице G_0, используя алгоритм 2, построить такую перестановочную матрицу P_0, что
$$G \cdot P_0 = (G_1 \| \dots \| G_u),$$
где $(k \times n)$-матрицы G_i являются порождающей матрицей кода $\mathrm{RM}^{\sigma_i}(r, m)$ для некоторой подстановки $\sigma_i \in S_n$, $i = 1, 2 \dots, u$.
3. Используя атаку И. Чижова и М. Бородина [7], по матрице G_i, $i = 1, 2, \dots, u$, найти такую невырожденную $(k \times k)$-матрицу M_i, и такую перестановочную матрицу P_i, что $G_i = M_i R P_i$, здесь R — порождающая матрица кода $\mathrm{RM}(r, m)$.
4. Составить следующую перестановочную $(un \times un)$-матрицу
$$P = \begin{pmatrix} P_1 & 0 & \dots & 0 \\ 0 & P_2 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & P_u \end{pmatrix} \cdot P_0^{-1}.$$
5. Выдать секретный ключ $\langle M_1, \dots, M_u, P \rangle$.

В связи с тем, что для множество ключей криптосистемы Мак-Элиса–Сидельникова, к которым применима теорема 1, была построена атака, есть смысл ввести следующее определение.

Определение 8. Пусть $K = \langle M_1, \dots, M_u, P \rangle$ — секретный ключ криптосистемы Мак-Элиса – Сидельникова. Будем говорить, что K является слабым ключом, если матрицы $M_1^{-1} M_2, \dots, M_1^{-1} M_u$ удовлетворяют условию теоремы 1.

Замечание. Обратим внимание, что перестановочная матрица P в определении не фигурирует.

C. Сложность атаки

Оценим сложность построенной атаки в битовых операциях.

Теорема 2. *Пусть НОД$(r, m - 1) = d > 1$, и $K = \langle M_1, \dots, M_u, P \rangle$ — слабый секретный ключ. Тогда сложность алгоритма решения задачи криптоаналитика для открытого ключа, соответствующего K, будет равна $u \cdot \exp\left(C_d (\log_2 n)^d (1 + o(1))\right)$ битовых операций.*

Таблица I
Время восстановления секретного ключа для $u = 3$.

r \ m	6	8	9	10
2	0,411 с	23,025 с	—	3 ч 18 м
3	—	51,891 с	10 м 41 с	2 ч 36 м
4	—	—	—	4 ч 36 м

Таблица II
Время восстановления секретного ключа для $u = 4$.

r \ m	6	8	9	10
2	0,719 с	40,094 с	—	5 ч 14 м
3	—	1 м 46 с	20 м 41 с	3 ч 40 м
4	—	—	—	7 ч 46 м

Доказательство. Наиболее трудоёмким является третий шаг схемы, в котором требуется u раз применить атаку И. Чижова и М. Бородина, сложность которой в условиях теоремы равна $\exp\big(C_d(\log_2 n)^d(1+o(1))\big)$ [7]. □

Теорема 3. *Пусть НОД$(r, m-1) = 1$, и $K = \langle M_1, \ldots, M_u, P\rangle$ — слабый секретный ключ. Тогда сложность алгоритма решения задачи криптоаналитика для открытого ключа, соответствующего K, равна $O(un^4 \log_2 n)$ битовых операций.*

Доказательство. Наиболее трудоёмким является третий шаг схемы, в котором требуется u раз применить атаку И. Чижова и М. Бородина, сложность которой в условиях теоремы равна $O(n^4 \log_2 n)$ [7]. □

D. Вычислительные эксперименты

Теоретические результаты атаки на криптосистему Мак-Элиса–Сидельникова, секретный ключ которой является слабым, подтверждаются практическими экспериментами. Алгоритм был реализован на языке C++ и исследован на ноутбуке с процессором Intel Core i5 dual-core 1.4GHz.

Результаты экспериментов для наиболее интересных параметров приведены в таблицах I и II. Отметим особо, что для параметров $r = 3, u = 4, m = 10$, предлагаемых в работе [4] использовать на практике, время восстановления секретного ключа составляет всего около 4 часов на современном ноутбуке.

IV. Оценка числа слабых ключей

При исследовании множества слабых ключей криптосистемы Мак-Элиса–Сидельникова удобной оказывается интерпретация квадрата Адамара линейного кода с точки зрения квадратичных форм над полем $\mathbb{F}_2$. Такой подход позволил авторам работы [10] установить поведение размерности квадрата Адамара случайного линейного кода.

Определение 9. Квадратичной формой над полем $\mathbb{F}_2$ называется однородный квадратичный многочлен над этим полем:

$$q(x_1, \ldots, x_k) = \sum_{1\le i<j\le k} a_{i,j}x_ix_j + \sum_{i=1}^{k} b_i x_i^2,$$

где $a_{i,j} \in \mathbb{F}_2, 1 \le i < j \le k$, а $b_i \in \mathbb{F}_2, 1 \le i \le k$.

Замечание. В поле $\mathbb{F}_2$ выполнено соотношение $x^2 = x$, поэтому квадратичная форма может быть представлена в виде

$$q(x_1, \ldots, x_k) = \sum_{1\le i<j\le k} a_{i,j}x_ix_j + \sum_{i=1}^{k} b_i x_i.$$

Обозначим через $\mathcal{Q}_k$ множество всех квадратичных форм над полем $\mathbb{F}_2$ от k переменных. Рассмотрим $(k\times n)$-матрицу G, пусть $g_i \in \mathbb{F}_2^k$ — столбец матрицы G с номером i. Определим отображение $\ell_G : \mathcal{Q}_k \to \mathbb{F}_2^n$ следующим образом:

$$\ell_G(q) = \big(q(g_1), \ldots, q(g_n)\big).$$

В таком случае квадрат линейного $[n, k]$-кода $\mathcal{C}$ с порождающей матрицей G может быть записан в терминах образа отображения [10]:

$$\mathcal{C}^{\star 2} = \operatorname{Im} \ell_G.$$

Этот подход приводит к следующей теореме, записанной в терминах ядра отображения.

Теорема 4. *Пусть $\mathcal{C}$ — линейный $[n, k]$-код с порождающей матрицей G, $\Delta = (H_1, \ldots, H_{u-1})$ — кортеж невырожденных двоичных $(k \times k)$-матриц Тогда*

$$\mathcal{C}_\Delta^{\star 2} = (\mathcal{C}^{\star 2})^{\times u},$$

если и только если для всех $1 \le j \le u-1$ выполнено равенство

$$\ker \ell_{H_{u-j}G} + \bigcap_{i=0}^{u-j-1} \ker \ell_{H_iG} = \mathcal{Q}_k,$$

здесь H_0 — единичная $(k \times k)$-матрица.

Доказательство. Будем обозначать как $\kappa = \dim \mathcal{C}^{\star 2}$ размерность квадрата кода $\mathcal{C}$. Для ядра отображения ℓ выберем следующее обозначение:

$$L_{E,H_1,\ldots,H_{u-1}} = \ker \ell_{(EG\|H_1G\|\ldots\|H_{u-1}G)}.$$

L_E — множество квадратичных форм, которые принимают значение ноль одновременно на всех столбцах матрицы G. Аналогично, для $1 \le i \le u-1$, L_{H_i} — на всех столбцах матрицы H_iG. Таким образом, имеет место быть следующее равенство:

$$L_{E,H_1,\ldots,H_{u-1}} = \bigcap_{i=0}^{u-1} L_{H_i}.$$

Условие

$$\mathcal{C}_\Delta^{\star 2} = (\mathcal{C}^{\star 2})^{\times u}$$

в указанных обозначениях эквивалентно выполнению равенства

$$\dim \operatorname{Im} \ell_{(EG\|H_1G\|\ldots\|H_{u-1}G)} = u \dim \operatorname{Im} \ell_G = u\kappa.$$

Далее,

$$\begin{aligned}\dim \operatorname{Im} \ell_{(EG\|H_1G\|\ldots\|H_{u-1}G)} &= \dim \mathcal{Q}_k - \dim L_{E,H_1,\ldots,H_{u-1}} \\ &= \dim \mathcal{Q}_k - \dim \bigcap_{i=0}^{u-1} L_{H_i}.\end{aligned}$$

Значит интересующее нас условие можно записать в виде

$$\dim \bigcap_{i=0}^{u-1} L_{H_i} = \dim \mathcal{Q}_k - u\kappa.$$

Для произвольных линейных подпространств U и V верно:

$$\dim(U+V) = \dim U + \dim V - \dim U \cap V.$$

Тогда для $1 \le j \le u-1$:

$$\dim \bigcap_{i=0}^{u-j} L_{H_i} = \dim L_{H_{u-j}} + \dim \bigcap_{i=0}^{u-j-1} L_{H_i} - \\ - \dim \left(L_{H_{u-j}} + \bigcap_{i=0}^{u-j-1} L_{H_i} \right),$$

а значит, учитывая, что $\dim L_E = \dim \mathcal{Q}_k - \kappa$,

$$\dim \bigcap_{i=0}^{u-1} L_{H_i} = \\ = \sum_{j=0}^{u-1} \dim L_{H_j} - \sum_{j=1}^{u-1} \dim \left(L_{H_{u-j}} + \bigcap_{i=0}^{u-j-1} L_{H_i} \right) = \\ = u \dim L_E - \sum_{j=1}^{u-1} \dim \left(L_{H_{u-j}} + \bigcap_{i=0}^{u-j-1} L_{H_i} \right) = \\ = u \dim \mathcal{Q}_k - u\kappa - \sum_{j=1}^{u-1} \dim \left(L_{H_{u-j}} + \bigcap_{i=0}^{u-j-1} L_{H_i} \right) \ge \\ \ge \dim \mathcal{Q}_k - u\kappa.$$

При этом равенство выполняется тогда и только тогда, когда для всех $1 \le j \le u-1$

$$L_{H_{u-j}} + \bigcap_{i=0}^{u-j-1} L_{H_i} = \mathcal{Q}_k.$$

□

Следствие 2. *Пусть $\mathcal{C}$ — линейный $[n,k]$-код с порождающей матрицей G, H — произвольная невырожденная $(k \times k)$-матрица. Тогда*

$$\mathcal{C}_H^{\star 2} = (\mathcal{C}^{\star 2})^{\times 2},$$

если и только если $L_E + L_H = \mathcal{Q}_k$.

Доказательство. Применим теорему 4 при $u = 2$. □

А. Оценка числа подпространств специального вида пространства квадратичных форм

Рассмотрим линейное пространство квадратичных форм от k переменных $\mathcal{Q}_k$. Размерность $\dim \mathcal{Q}_k = \frac{k(k+1)}{2} = w$. Пусть $L_0 \subset \mathcal{Q}_k$ — линейное подпространство $\mathcal{Q}_k$ с размерностью равной $\dim L_0 = \dim L_E = w - \kappa = e$. Подсчитаем количество подпространств $L \subset \mathcal{Q}_k$ таких, что $\dim L = e$ и $L_0 + L = \mathcal{Q}_k$.

Введём обозначение $\delta = \dim L \cap L_0$. Тогда

$$\dim(L_0 + L) = \underbrace{\dim V}_{w} = \underbrace{\dim L_0}_{e} + \underbrace{\dim L}_{e} - \underbrace{\dim L \cap L_0}_{\delta}.$$

Таким образом, $\delta = w - 2\kappa$.

Теорема 5. *Пусть V — линейное пространство размерности $\delta + 2\kappa$, $L_0 \subset V$ — фиксированное линейное подпространство размерности $\delta + \kappa$. Тогда доля $\alpha(\delta, \kappa)$ всех подпространств $L \subset V$ размерности $\delta + \kappa$, для которых $L + L_0 = V$, к числу всех подпространств размерности $\delta + \kappa$ выражается по формуле:*

$$\alpha(\delta, \kappa) = \prod_{j=1}^{\kappa} \frac{2^{\delta+2\kappa+j} - 2^{\kappa}}{2^{\delta+2\kappa+j} - 1}.$$

Более того, $\alpha(\delta, \kappa) \to 1$ при $\kappa \to \infty$ для $\forall \delta$.

Доказательство. Пусть как и раньше $w = \delta + 2\kappa$, а $e = \delta + \kappa$. Число различных способов выбрать подпространство Π размерности δ такое, что $\Pi \subset L_0 \subset V$ выражается через биномиальный коэффициент Гаусса

$$\begin{bmatrix} e \\ \delta \end{bmatrix}_2 = \prod_{j=1}^{e} \frac{2^{e-\delta+j} - 1}{2^j - 1} \text{ для } \delta \le e.$$

Зафиксируем некоторый базис пространств Π и L_0.

$$\begin{aligned} \Pi: &\quad a_1, \ldots, a_\delta \\ L_0: &\quad a_1, \ldots, a_\delta, \quad b_1, \ldots, b_{e-\delta} \end{aligned}$$

Будем выбирать c_i, начиная с c_1 и заканчивая $c_{e-\delta}$, так, чтобы система векторов: $a_1, \ldots, a_\delta, b_1, \ldots, b_{e-\delta}, c_1, \ldots c_i$ являлась линейно независимой (иначе окажется, что $\dim L_0 \cap L > \delta$). Существует $2^w - 2^e$ способов выбрать c_1, $2^w - 2^{e+1}$ способов выбрать c_2 при фиксированном c_1, ..., $2^w - 2^{e+(e-\delta)-1} = 2^w - 2^{w-1}$ способов выбрать $c_{e-\delta}$ при фиксированных $c_1, \ldots, c_{e-\delta-1}$. Таким образом, существует

$$(2^w - 2^e)(2^w - 2^{e+1}) \ldots (2^w - 2^{w-1}) = \\ = 2^{e\kappa}(2^\kappa - 1)(2^\kappa - 2) \ldots (2^\kappa - 2^{\kappa-1})$$

способов выбрать $\langle c_1, \ldots, c_{e-\delta} \rangle$. При этом базисы

$$\langle a_1, \ldots, a_\delta, c_1, \ldots, c_{e-\delta} \rangle, \ \langle a_1, \ldots, a_\delta, c'_1, \ldots, c'_{e-\delta} \rangle$$

порождают одно и то же линейное подпространство, если существует такая невырожденная $((e-\delta) \times (e-\delta))$-матрица M и такая произвольная $((e-\delta) \times \delta)$-матрица, что выполнено соотношение

$$\begin{pmatrix} c'_1 \\ c'_2 \\ \vdots \\ c'_{e-\delta} \end{pmatrix} = \left(\begin{array}{c|c} A & M \end{array} \right) \cdot \left(\begin{array}{c} a_1 \\ \vdots \\ a_\delta \\ \hline c_1 \\ c_2 \\ \vdots \\ c_{e-\delta} \end{array} \right),$$

Количество различных матриц $S = (A|M)$ равно произведению количества различных матрицы A на число различных матриц M, т.е. равно

$$2^{(e-\delta)\delta}(2^{e-\delta} - 1) \ldots (2^{e-\delta} - 2^{e-\delta-1}) = \\ = 2^{\kappa\delta}(2^\kappa - 1) \ldots (2^\kappa - 2^{\kappa-1}).$$

Окончательно получаем, что число подпространств $L \subset V$ размерности e, для которых $L_0 + L = V$, равно

$$\begin{bmatrix} e \\ \delta \end{bmatrix}_2 \cdot 2^{\kappa(e-\delta)} = \begin{bmatrix} e \\ \delta \end{bmatrix}_2 \cdot 2^{\kappa^2}.$$

Деля это число на количество всех возможных подпространств $L \subset V$ размерности e, получим выражение искомой доли $\alpha(\delta, \kappa)$

$$\alpha(\delta,\kappa) = \frac{2^{\kappa^2} \cdot \begin{bmatrix} e \\ \delta \end{bmatrix}_2}{\begin{bmatrix} w \\ e \end{bmatrix}_2} = \frac{2^{\kappa^2} \cdot \begin{bmatrix} \delta+\kappa \\ \delta \end{bmatrix}_2}{\begin{bmatrix} \delta+2\kappa \\ \delta+\kappa \end{bmatrix}_2} = \prod_{j=1}^{\kappa} \frac{2^{\delta+2\kappa+j} - 2^{\kappa}}{2^{\delta+2\kappa+j} - 1}.$$

Откуда:

$$1 \geq \alpha(\delta,\kappa) = \prod_{j=1}^{\kappa} \left(1 - \underbrace{\frac{2^{\kappa}-1}{2^{\delta+2\kappa+j-1}}}_{\leq \frac{1}{2^{\kappa}}}\right) \geq \left(1 - \frac{1}{2^{\kappa}}\right)^{\kappa}.$$

Запишем следующее очевидное равенство

$$\left(1 - \frac{1}{2^{\kappa}}\right)^{\kappa} = \left(\left(1 - \frac{1}{2^{\kappa}}\right)^{2^{\kappa}}\right)^{\frac{\kappa}{2^{\kappa}}}.$$

Начиная с некоторого номера κ число $\left(1 - \frac{1}{2^{\kappa}}\right)^{2^{\kappa}}$ лежит в ε окрестности числа e. При этом $\frac{\kappa}{2^{\kappa}} \to 0$ при $\kappa \to \infty$. Значит $\left(1 - \frac{1}{2^{\kappa}}\right)^{\kappa} \to 1$ при $\kappa \to \infty$. Из того, что последовательность $\alpha(\delta, \kappa)$ ограничена сверху и снизу последовательностями, предел которых при $\kappa \to \infty$ равен 1, следует $\alpha(\delta, \kappa) \to 1$ при $\kappa \to \infty$. □

В случае, когда в качестве исходного кода берётся код $\mathrm{RM}(r, m)$, используемые параметры принимают следующие вид:

$$k = \sum_{j=0}^{r} \binom{m}{j}, \quad \kappa = \sum_{j=0}^{2r} \binom{m}{j},$$
$$\delta = \frac{k(k+1)}{2} - 2\sum_{j=0}^{2r} \binom{m}{j}.$$

Численные эксперименты для параметров (r, m), обозримо применимых в криптосистеме, показали, что доля таких пространств близка к 1.

Теорема 5 косвенно позволяет сформулировать гипотезу, что доля слабых ключей среди всех секретных ключей, является подавляющей. Вместе с тем, слабому секретному ключу соответствует линейное подпространство специального вида. Для доказательства выдвинутой гипотезы достаточно будет доказать теорему 5, но для класса подпространств L специального вида.

B. Достаточные условия, при выполнении которых секретный ключ не будет слабым

Сформулируем и докажем теоремы, непосредственное применение которых даёт достаточные условия на матрицы $H_1, \ldots, H_{u-1}$, при выполнении которых соответствующий секретный ключ не будет являться слабым. Эти условия могут быть полезны для получении оценок сверху количества слабых ключей, а также для установления структуры множества ключей криптосистемы Мак-Элиса–Сидельникова, которые не являются слабыми. Преимущество достаточного условия, которое даёт теорема 6, заключается в том, что оно применимо в общем случае, при $u \geq 2$.

Теорема 6. *Пусть $\mathcal{C}$ — $[n, k]$-код с порождающей матрицей G, и пусть $\Delta = (H_1, \ldots, H_{u-1})$ — кортеж неврожденных $(k \times k)$-матриц, $u \geq 2$. Если найдётся хотя бы одна матрица H_i такая, что выполнено соотношение $\mathrm{rk}(E + H_i) < k - \log_2(2^k - n)$ на ранг матрицы $E + H_i$, то*

$$\mathcal{C}_{\Delta}^{\star 2} \neq \left(\mathcal{C}^{\star 2}\right)^{\times u}.$$

Доказательство. Обозначим через $L_M^{(j)}$ ядро отображения ℓ, для которого в качестве параметра выступает матрица, состоящая из одного j-го столбца матрицы MG. Тогда $L_E = \bigcap_{1 \leq j \leq n} L_E^{(j)}$, а $L_{H_i} = \bigcap_{1 \leq j \leq n} L_{H_i}^{(j)}$. Если матрицы G и $H_i G$ будут иметь совпадающие столбцы, окажется, что $\exists j \in \{1, \ldots, n\}$ для которого $L_E, L_{H_i} \subset L_E^{(j)}$. Это, в с свою очередь, означает, что $L_E + L_{H_i} \subset L_E^{(j)} \subsetneq \mathcal{Q}_k$. Следовательно, условие теоремы 4 не выполнено.

Запишем матрицу H_i как сумму единичной матрицы и матрицы общего вида: $H_i = E + M_i$, $\mathrm{rk}(M_i) = s \in \{0, \ldots, k\}$. Тогда условие $H_i \cdot \alpha^T = \alpha^T \iff M_i \cdot \alpha^T = 0$. Количество таких $\alpha \in \mathbb{F}_2^k$ равно 2^{k-s}. Значит, по принципу Дирихле, если $n > 2^k - 2^{k-s}$, то среди столбцов матрицы G найдётся столбец π такой, что $H_i \cdot \pi = \pi$, а значит $L_E + L_{H_i} \neq \mathcal{Q}_k$. □

В случае когда $u = 2$, условие теоремы 6 можно ослабить. Обратим внимание, что $\mathcal{Q}_k$ является линейным пространством, изоморфным подкоду $\mathcal{R}$ кода Рида–Маллера второго порядка $\mathrm{RM}(2, k)$. Изоморфизм χ выглядит следующим образом:

$$\chi(q) = (q(0, \ldots, 0, 0), q(0, \ldots, 0, 1), \ldots, q(1, \ldots, 1, 1)).$$

Это позволяет воспользоваться результатом, полученным в работе [14]. Прежде чем перейти к теореме 7, опишем действие элемента полной линейной группы $H \in GL_k$ на код Рида–Маллера $\mathrm{RM}(2, k)$. Пусть f — булева функция k переменных, x — двоичный вектор аргументов. Тогда

$$f^H(x) = f(Hx),$$

а код, полученный под действием H, представляет собой

$$(E + H) \circ \mathrm{RM}(2, k) = \{f + f^H | f \in \mathrm{RM}(2, k)\}.$$

Заметим, что

$$\chi(L_{E+H}) = (E + H) \circ \chi(L_E).$$

Теорема 7. *Пусть $\mathcal{C}$ — $[n, k]$-код с порождающей матрицей G, матрица H — невырожденная двоичная $(k \times k)$-матрица, и пусть $\kappa = \dim \mathcal{C}^{\star 2}$. Дополнительно потребуем чтобы $\left(\frac{2k+1}{2}\right)^2 - 2\kappa \geq 0$. Если $\mathrm{rk}(E + H) < \frac{2k+1}{2} - \sqrt{\left(\frac{2k+1}{2}\right)^2 - 2\kappa}$, то $\mathcal{C}_H^{\star 2} \neq \left(\mathcal{C}^{\star 2}\right)^{\times 2}$.*

Доказательство. Очевидно, что $L_E + L_H = L_E + L_{E+H}$. Это можно видеть из того, что $\chi(L_E) + \chi(L_H) = \chi(L_E) + (E+H) \circ \chi(L_E)$. Известно [14], что для ранга $s = \mathrm{rk}(E + H)$ выполняется неравенство:

$$\dim\left((E + H) \circ \mathrm{RM}(2, k)\right) \leq sk - \frac{s(s-1)}{2},$$

Таблица III
Ограничения на ранг, полученные из теорем 6 и 7

(r, m)	$(2, 7)$	$(3, 7)$	$(2, 9)$	$(4, 9)$	$(3, 10)$
$\mathrm{rk}\, H$	8	29	10	130	56
$\mathrm{rk}(H + E)$	1	4	1	3	12

а значит тем более выполняется

$$\dim L_{E+H} = \dim\left((E + H) \circ \chi(L_E)\right) \leq sk - \frac{s(s-1)}{2}.$$

Запишем цепочку соотношений

$$\dim(L_E + L_{E+H}) \leq \dim L_E + \dim L_{E+H} = \\ = \frac{k(k+1)}{2} - \kappa + \dim L_{E+H}.$$

Значит, если $\dim L_{E+H} < \kappa$, то

$$\dim(L_E + L_{E+H}) < \frac{k(k+1)}{2} \iff L_E + L_H \neq \mathcal{Q}_k.$$

Из условия $sk - \frac{s(s-1)}{2} < \kappa$ следует утверждение теоремы. □

C. Вычислительные эксперименты для выяснения количества слабых ключей

Для получения оценки сверху числа слабых ключей необходимо подсчитать количество невырожденных матриц $H \in GL_k$ таких, что $\mathrm{rk}(H + E) \leq s$, для некоторого заданного s.

На основе теорем 6 и 7 были вычислены ограничения на ранг матрицы $H + E$ для параметров (r, m) криптосистемы Мак-Элиса–Сидельникова, наиболее интересных с точки зрения практического применения (см. таблицу III).

Авторам не удалось получить аналитического выражения числа таких матриц. Но, для указанных параметров, были проведены вычислительные эксперименты, которые показали, что их доля –крайне мала. В результате случайного выбора невырожденных матриц с объёмом выборки порядка 10^6 не было найдено ни одной удовлетворяющей данному условию матрицы. Максимальное падение ранга $\mathrm{rk}(H) - \mathrm{rk}(H + E)$ в ходе эксперимента равнялось 4 для всех параметров.

V. Заключение

В работе предложен эффективный алгоритм атаки на криптосистему Мак-Элиса–Сидельникова, построенную на кодах Рида–Маллера. В отличие от существующих алгоритмов на множество доменных параметров криптосистемы ограничений не накладывается.

Стоит заметить, что алгоритм не применим в отдельно взятых случаях. В работе определён класс слабых ключей и описаны некоторые его характеристики, позволяющие дать оценки количества таких ключей, долю слабых ключей в пространстве ключей криптосистемы. Приведена мотивация, в связи с которой следует ожидать, что доля слабых ключей будет близка к единице. Проведены вычислительные эксперименты.

Становится понятно, что данная криптосистема не обеспечивает необходимый уровень стойкости. Для обозримо применимых параметров криптосистемы восстановить секретный ключ схемы возможно даже на персональной ЭВМ.

Библиография

[1] McEliece Robert J. A public-key cryptosystem based on algebraic coding theory // DSN Progress Report. — 1978. — Vol. 42–44. — P. 114–116.

[2] Shor Peter W. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer // SIAM Journal on Computing. — 1997. — Oct. — Vol. 26, no. 5. — P. 1484–1509.

[3] Niederreiter Harald. Knapsack-type cryptosystems and algebraic coding theory // Prob. Control and Inf. Theory. — 1986. — Vol. 15, no. 2. — P. 159–166.

[4] Сидельников В. М. Открытое шифрование на основе двоичных кодов Рида – Маллера // Дискретная математика. — 1994. — Vol. 6, no. 2. — P. 3–20.

[5] Courtois Nicolas T., Finiasz Matthieu, Sendrier Nicolas. How to achieve a mceliece-based digital signature scheme // Advances in Cryptology — ASIACRYPT 2001 / Ed. by Colin Boyd. — Lecture Notes in Computer Science. — Springer, 2001. — P. 157–174.

[6] Minder Lorenz, Shokrollahi Amin. Cryptanalysis of the sidelnikov cryptosystem // Annual International Conference on the Theory and Applications of Cryptographic Techniques / Springer. — 2007. — P. 347–360.

[7] Бородин М. А., Чижов И. В. Эффективная атака на криптосистему Мак-Элиса, построенную на основе кодов Рида–Маллера // Дискретная математика. — 2014. — Vol. 26, no. 1. — P. 10–20.

[8] Wieschebrink Christian. Cryptanalysis of the niederreiter public key scheme based on grs subcodes // International Workshop on Post-Quantum Cryptography / Springer. — 2010. — P. 61–72.

[9] A polynomial-time attack on the bbcrs scheme / Alain Couvreur, Ayoub Otmani, Jean-Pierre Tillich, Valérie Gauthier-Umana // IACR International Workshop on Public Key Cryptography / Springer. — 2015. — P. 175–193.

[10] Squares of random linear codes / Ignacio Cascudo, Ronald Cramer, Diego Mirandola, Gilles Zémor // IEEE Transactions on Information Theory. — 2015. — Vol. 61, no. 3. — P. 1159–1173.

[11] Классификация произведений Адамара подкодов коразмерности 1 кодов Рида-Маллера / Иван Владимирович Чижов, Ivan Vladimirovich Chizhov, Михаил Алексеевич Бородин, Mikhail Alekseevich Borodin // Дискретная математика. — 2020. — Vol. 32, no. 1. — P. 115–134.

[12] Давлетшина А. М. Поиск эквивалентных ключей криптосистемы Мак-Элиса–Сидельникова, построенной на двоичных кодах Рида–Маллера // Прикладная дискретная математика. Приложение. — 2019. — no. 12. — P. 98–100.

[13] Мак-Вильямс, Слоэн. Теория кодов, исправляющих ошибки. — Москва : Связь, 1979.

[14] Legeay Matthieu, Loidreau Pierre. Projected subcodes of the second order binary reed-muller code // 2012 IEEE International Symposium on Information Theory Proceedings / IEEE. — 2012. — P. 254–258.

Effective structural attack on McEliece-Sidelnikov public-key cryptosystem

Ivan Chizhov, Sergei Koniukhov, Alexandra Davletshina

***Abstract*—The authors propose an algorithm for recovering the secret key of the McEliece–Sidelnikov cryptosystem in general case: with $u \in \mathbb{N}$ copies of the Reed–Muller codes. Recovering the secret key of the McEliece–Sidelnikov cryptosystem is reduced to u problems of recovering the secret key of the McEliece cryptosystem based on the Reed–Muller codes. It is proved in the paper that the proposed attack is polynomial. A set of keys for which the algorithm is applicable is described. The set is called the set of weak keys. The authors believe that the most of the keys are weak and show that it should be assumed that the ratio of the weak keys in the cryptosystem's key space is close to one. Methods for calculating the number of the weak keys are described and computational experiments confirming it have been performed.**

***Keywords*—McEliece–Sidelnikov public key cryptosystem, Reed–Muller codes, postquantum cryptography, code-based cryptosystems, polynomial attack.**

References

[1] McEliece Robert J. A public-key cryptosystem based on algebraic coding theory // DSN Progress Report. — 1978. — Vol. 42–44. — P. 114–116.

[2] Shor Peter W. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer // SIAM Journal on Computing. — 1997. — Oct. — Vol. 26, no. 5. — P. 1484–1509.

[3] Niederreiter Harald. Knapsack-type cryptosystems and algebraic coding theory // Prob. Control and Inf. Theory. — 1986. — Vol. 15, no. 2. — P. 159–166.

[4] Sidelnikov V. M. Public encryption based on reed–muller codes // Discrete Math. Appl. — 1994. — Vol. 6, no. 2. — P. 3–20.

[5] Courtois Nicolas T., Finiasz Matthieu, Sendrier Nicolas. How to achieve a mceliece-based digital signature scheme // Advances in Cryptology — ASIACRYPT 2001 / Ed. by Colin Boyd. — Lecture Notes in Computer Science. — Springer, 2001. — P. 157–174.

[6] Minder Lorenz, Shokrollahi Amin. Cryptanalysis of the sidelnikov cryptosystem // Annual International Conference on the Theory and Applications of Cryptographic Techniques / Springer. — 2007. — P. 347–360.

[7] Borodin M. A., Chizhov I. V. Jeffektivnaja ataka na kriptosistemu makjelisa, postroennuju na osnove kodov rida–mallera // Discrete Math. Appl. — Vol. 26, no. 1. — P. 10–20.

[8] Wieschebrink Christian. Cryptanalysis of the niederreiter public key scheme based on grs subcodes // International Workshop on Post-Quantum Cryptography / Springer. — 2010. — P. 61–72.

[9] A polynomial-time attack on the bbcrs scheme / Alain Couvreur, Ayoub Otmani, Jean-Pierre Tillich, Valérie Gauthier-Umana // IACR International Workshop on Public Key Cryptography / Springer. — 2015. — P. 175–193.

[10] Squares of random linear codes / Ignacio Cascudo, Ronald Cramer, Diego Mirandola, Gilles Zémor // IEEE Transactions on Information Theory. — 2015. — Vol. 61, no. 3. — P. 1159–1173.

[11] Chizhov I. V., Borodin M. A. Hadamard products classification of subcodes of reed–muller codes codimension 1 // Discrete Math. Appl. — 2020. — Vol. 32, no. 1. — P. 115–134.

[12] Davletshina A. M. Search for equivalent keys of the mceliece - sidelnikov cryptosystem built on the reed - muller binary codes // Prikladnaya diskretnaya matematika. Prilozhenie. — no. 12. — P. 98–100. — URL: http://journals.tsu.ru/pdm2/&journal_page=archive&id=1897&article_id=42419.

[13] MacWilliams Florence Jessie, Sloane Neil J. A. The theory of error-correcting codes. North-Holland mathematical Library no. 16. — 2. print edition. — North-Holland. — ISBN: 978-0-444-85010-2 978-0-444-85009-6. — OCLC: 613292217.

[14] Legeay Matthieu, Loidreau Pierre. Projected subcodes of the second order binary reed-muller code // 2012 IEEE International Symposium on Information Theory Proceedings / IEEE. — 2012. — P. 254–258.