Структурная атака на криптосистемы типа Мак-Элиса-Сидельникова, построенной на основе комбинирования случайных кодов с кодами Рида-Маллера

И. В. Чижов, Е. А. Попова

.

***Аннотация*— В работе рассматривается модификация криптосистемы Мак-Элиса–Сидельникова, построенная на комбинировании случайных кодов с кодами Рида-Маллера. Суть данной криптосистемы с открытым ключом состоит в маскировании кода с эффективным алгоритмом декодирования под код, не обладающий видимой алгебраической и комбинаторной структурой, называемый кодом общего положения. Фактически стойкость этой криптосистемы основывается на сложности задачи декодирования кодов общего положения. Интерес к кодовым криптосистемам объясняется тем, что в отличие от классических криптосистем, стойкость которых определяется сложностью решения задач факторизации или дискретного логарифмирования в конечных группах, кодовые криптосистемы останутся стойкими в эру появления многокубитного квантового компьютера.**

Первоначально криптосистема Мак-Элиса–Сидельникова строилась на основе комбинирования кодов Рида-Маллера друг с другом. Однако на этот вариант криптосистемы была построена эффективная атака. В работе Г. Кабатянского и С. Таверньe предлагается вместо комбинированных кодов Рида–Маллера использовать комбинирование кодов из разных классов, например, кодов Рида-Маллера и кодов Гоппы. По многим характеристикам коды Гоппы похожи на случайные коды, поэтому в работе рассматривается криптосистема Мак-Элиса–Сидельникова, в которой коды Рида–Маллера комбинируется со случайными кодами. Изучается стойкость новой криптосистемы в модели, в которой противнику кроме открытого ключа и, соответственно, параметров кода Рида-Маллера, на которых построен открытый ключ, известна порождающая матрица случайного линейного кода, и целью является восстановление секретного ключа криптосистемы.

С использованием метода разделения носителя Н. Сендрие построена атака, восстанавливающая секретный ключ в описанной модели. Для работы метода Н. Сендрие необходимо выбрать некоторый инвариант кода относительно группы подстановок координат кодовых слов. В качестве такого инварианта предлагается использовать спектр оболочки кода. Важно также, чтобы инвариант был эффективно вычислимым. В работе была изучена размерность оболочки кода, получающегося комбинированием кода Рида–Маллера кода с некоторым случайным кодом. Вычислено математическое ожидание размера оболочки, которое показывает эффективную вычислимость рассматриваемого инварианта.

Предложенная атака была реализована на языке C++. В работе представлены экспериментальные данные по применению атаки к кодам разной размерности, из которых можно судить об её эффективности даже на достаточно больших размерах секретных ключей.

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

I. Введение

Криптосистема Мак-Элиса была предложена в 1978 году Робертом Мак-Элисом [1] практически в одно время с широко известной криптосистемой RSA. Однако она не получила в то время большой известности и до недавнего времени оставалась объектом исследований только узкого числа специалистов по алгебраической теории кодирования. Ситуация поменялась в последние 5 лет.

В настоящее время почти все криптографические протоколы и механизмы, используемые на практике, построены основе теоретико-числовых моделях. Их стойкость, в основном, определяется сложностью задачи факторизации целых чисел или дискретного логарифмирования в конечных группах. В 1997 году П. Шор предложил [2] полиномиальные алгоритмы решения этих классических теоретико-числовых задач для квантового вычислителя. Таким образом, появление многокубитного квантового компьютера поставит вопрос о смене практически всех криптографических протоколов и механизмов современного Интернета.

В последнее время интерес к исследованию криптографических механизмов также подкреплялся достижения в области, так называемой, криптоэкономики. Исследователями были предложены концепции криптовалюты и разработана технология распределённого электронного реестра «Блокчейн». Кроме того, криптовалюта биткоин настолько приобрела популярность, что авторы смогли привлечь финансирование и открыть биржу, на которой стало

Статья получена 9 апреля 2013. (укажите здесь дату, когда статья была загружена на сайт журнала).

Работа поддержана грантом РФФИ №18-29-03124

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

Е. А. Попова, МГУ имени М. В. Ломоносова (email: lady-lizochka@yandex.ru)

возможным использование криптовалюты в реальной экономике через механизм её конвертации в национальные валюты государств. Фактически сейчас любой желающий может приобрести криптовалюту, обменять её на валюту своей страны. Таким образом, криптовалюты стали полноценным финансовым активом.

В технологии «Блокчейн», как и в криптовалютах, используются схемы подписи и протоколы обмена ключами. Причём от стойкость этих криптографических механизмов зависит надёжность всей технологии и доверие к криптовалютам. После появления квантового компьютера с большим количеством кубитов всем криптографические валюты мира 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 подогрело Национальное агентство стандартов и технологий США (NIST USA), которое объявило конкурс на серию стандартов эра квантового компьютера. На уровне целого государства была признана угроза появления многокубитного квантового компьютера.

Теория кодов, исправляющих ошибки, является одним из основных источников конструкций постквантовых криптографических механизмов.

Самым важным криптографическим протоколом с открытым ключом является схема электронной подписи. Она используется повсеместно в прикладных системах, в том числе и в технологии «Блокчейн». В настоящее время было предложено много подобных протоколов, основанных на теории кодов, исправляющих ошибки. На основе криптосистемы Мак-Элиса построена одна из важнейших схем подписи – CFS[3]. Тактико-технические характеристики этой схемы существенно зависят от используемых кодов, исправляющих ошибки. Оригинальная схема как криптосистема Мак-Элиса была построена на основе кодов Гоппы. Однако такой вариант обладает достаточно низкой скоростью формирования подписи. Для исправления этого недостатка предпочтительно использовать коды, обладающие большей исправляющей способностью.

Например, в работе В. М. Сидельникова [4] предлагается использовать коды Рида–Маллера. Однако относительно недавно появилась серия структурных атак [5], [6] на криптосистему Мак-Элиса, построенную на таких кодах. Поэтому, из-за схожей природы CFS с криптосистемой Мак-Элиса, эти коды нельзя использовать для построения схемы подписи.

В той же работе [4] В. М. Сидельников предлагает конструкцию, использующую несколько повторений матрицы кода Рида-Маллера. Такая схема строится на конкатенации нескольких различных порождающих матриц кода Рида–Маллера. Конструкцию В. М. Сидельникова принято в литературе называть криптосистемой Мак-Элиса–Сидельникова. Вместе с тем в работе [5] предлагается подход к взлому такой модификации криптосистемы (а значит и подписи CFS).

Идеи В. М. Сидельникова получили развития в работе Е. Егоровой, Г. Кабатянского, Е. Крука и С. Тавернье [7]. Они предлагают использовать конкатенацию не кодов Рида–Маллера, а двух различных кодов, имеющих разную алгебраическую структуру, например, кодов Гоппы и Рида-Маллера. В качестве обоснования авторы указывают, что для кодов Гоппы пока не существует эффективных структурных атак, поэтому они могут «спрятать» алгебраическую структуру кодов Рида–Маллера, которая значительно используется в атаках. В дальнейшем в работе эта криптосистема называется крптосистемой Кабатянского–Тавернье, т.к. она впервые была предложена именно в таком составе авторов на конференции «Sixteenth International Workshop on Algebraic and Combinatorial Coding Theory (ACCT-XVI)».

Настоящая работа посвящена исследованию стойкости постквантовой криптосистемы Мак-Элиса–Сидельникова, построенной на основе комбинирования кодов Рида–Маллера со случайными кодами. Замена кодов Гоппы случайным двоичным кодом объясняется тем, что случайные коды должны сильнее спрятать структуру кода, поэтому результаты анализа такой криптосистемы можно перенести и на конструкцию Кабатянского–Тавернье.

Главным результатом работы является построение структурной атаки на криптосистему Мак-Элиса–Сидельникова в модели, в которой противнику кроме открытого ключа известна порождающая матрица кода Рида–Маллера и случайного двоичного код, с которым проводится его комбинирование. Целью криптоаналитика является восстановление секретного ключа криптосистемы. Эксперименты показывают достаточную эффективность предложенной атаки не только для комбинации кодов Рида–Маллера со случайными кодами, но и для комбинирования кодов Рида–Маллера с кода Гоппы. Полученный результат переносится и на схему подписи CFS, которая строится на основе анализируемой конструкции.

II. Описание исследуемой криптосистемы типа Мак-Элиса-Сидельникова

Дадим краткое описание криптосистемы Мак-Элиса–Сидельникова, построенной на основе комбинирования случайных кодов с кодами Рида–Маллера, а также сформулируем задачу криптоаналитика в используемой модели нарушителя.

Обозначим через V_n линейное векторное пространство над полем $GF(2)$.

Определение 1. Подпространство C размерности k пространства V_n называется двоичным линейным $[n, k]$-кодом.

Двоичная $(k \times n)$-матрица, строками которой являются базисные векторы C, называется порождающей матрицей кода. Таким образом, $C = a \cdot G | a \in V_k$.

Скалярным произведением векторов $x = (x_1, \dots, x_n)$ и $y = (y_1, \dots, y_n)$ называется отображение $\langle x, y \rangle : V_n \to GF(2)$, определяемое соотношением

$$\langle x, y \rangle = x_1 y_1 \oplus \dots \oplus x_n y_n.$$

Векторы x и y называются ортогональными, если их скалярное произведение равно 0, т. е. $\langle x, y\rangle = 0$.

Определение 2. Дуальным кодом к $[n, k]$-коду C называется $[n, n-k]$-код $C^{\perp}$, который состоит из всех векторов V_n, ортогональных каждому кодовому слову кода C, т. е. $C^{\perp} = \{x \in V_n | \langle x, c\rangle = 0, \forall\, c \in C\}$.

Определение 3. Порождающая матрица H кода $C^{\perp}$ называется проверочной матрицей кода C. Матрица H имеет размер $\big((n-k)\times n\big)$. Фактически код C может быть задан как множество векторов $c \in V_n$, которые удовлетворяют следующему линейному уравнению

$$Hc^T = 0.$$

Определим теперь важный для дальнейшего изложения класс кодов – кодов Рида–Маллера.

Булевой функцией от m переменных называется произвольное отображение $f(v_1, \dots, v_m): V_m \to GF(2)$. Каждая булева функция может быть задана её вектором-значений Ω_f длины 2^m следующим образом:

$$\Omega_f = \big(f(0, \dots, 0), f(0, \dots, 1), \dots, f(1, \dots, 1)\big).$$

Мономом называется константа 1 и любая функция, представимая в виде произведения:

$$v_{i_1} v_{i_2} \dots v_{i_s},$$

здесь $1 \le i_1 < i_2 < \dots < i_s \le m$.

Известно, см. например [8], что любая булева функция от m переменных представима однозначным образом в виде некоторой суммы мономов. Такая сумма называется полиномом Жегалкина.

Степенью монома называется количество переменных, входящих в моном, так степень монома $v_{i_1} v_{i_2} \dots v_{i_s}$ равна s, а степень константы 1 равна нулю. Тогда будем называть степенью булевой функции максимальную степень мономов, входящих в её представление в виде полинома Жегалкина.

Определение 4. Для произвольных целых r и m, $0 \le r \le m$, двоичным кодом Рида-Маллера $RM(r, m)$ порядка r и длины 2^m называется множество векторов значений булевых функций $f = f(v_1, \dots, v_m)$, которые могут быть заданы полиномами Жегалкина степени не выше r, т. е.

$$RM(r, m) = \{\Omega_f \,|\, \deg f(v_1, \dots, v_m) \le r\}.$$

Известно [8], что длина кода $RM(r, m)$ равна $n = 2^m$, а его размерность равна $k = 1 + \binom{m}{1} + \binom{m}{2} + \dots + \binom{m}{r}$. Порождающая матрица $G(r, m)$ этого кода состоит из векторов-значений всех мономов степени не выше r.

Всё готово для описания криптосистемы Кабатянского–Тавернье. В силу того, что цель работы — построение структурной атаки, т. е. атаки, которая восстанавливает секретный ключ криптосистемы, ограничимся описанием только процедуры генерации ключей криптосистемы.

A. Общедоступные параметры криптосистемы

На основе заданного уровня стойкости в системе выбираются параметры (r, m) кода Рида–Маллера. Выбранные значения определяют размерность $k = \sum_{i=0}^{r} \binom{m}{i}$ и длину $n = 2^m$ кода Рида–Маллера.

Далее в системе выбирается случайная $(k \times n)$-матрица R максимального ранга k.

Заметим, что в оригинальной работе [7] авторы вместо случайного кода предлагают использовать коды Гоппы. Однако для упрощения анализа криптосистемы предлагается заменить коды Гоппы случайными кодами. Заметим, что безусловно, рассматриваемая конструкция является модельной и фактически не реализуемой на практике, т. к. для построения криптосистемы необходимо использовать коды, обладающие эффективными алгоритмами декодирования. Известно [9], что задача декодирования для случайных кодов является NP-полной, поэтому случайный код, скорее всего, не обладает эффективными алгоритмами декодирования.

Результаты анализа для криптосистемы, построенной на случайных кодах, будут перенесены на оригинальную идею работы [7].

B. Генерация ключей

Вход: Параметры (r, m) кода Рида-Маллера и матрица R.

Выход: K_{pub}, K_{sec} – открытый и секретный ключи криптосистемы.

1. Используя параметры (r, m), построить порождающую $(k \times n)$ −матрицу $G(r, m)$ кода Рида-Маллера.
2. Построить матрицу G_{pub} как конкатенацию матриц $G(r, m)$ и R, т. е. $G_{pub} = (G(r, m)|R)$.
3. Сгенерировать невырожденную двоичную $(k \times k)$-матрицу S.
4. Сгенерировать перестановочную $(n \times n)$-матрицу P.
5. Открытым ключом K_{pub} объявляется произведение $SG_{pub}P$.
6. Секретным ключ K_{sec} является кортеж матриц (S, P).

Сделаем небольшое замечание по поводу описанного алгоритма генерации ключей. Предполагается, что порождающая матрица R случайного линейного кода зафиксирована, поэтому матрица G_{pub} из алгоритма является общедоступной. Можно рассматривать криптосистему, в которой эта матрица не фиксируется, а выбирается в процессе генерации ключей. Вместе с тем необходимо потребовать существование эффективного алгоритма декодирования такого кода. При замене случайных кодов на коды Гоппы существование такого алгоритма известно, поэтому для этих кодов необходимо изучить стойкость не только описанной криптосистемы, но и её модификации, при которой матрица R выбирается абонентом. В работе такой вариант не рассматривается. Он является темой дальнейших исследований.

Сформулируем **задачу криптоаналитика**.

Вход: Матрицы $\big(G_{pub}, K_{pub}\big)$

Найти: Невырожденную двоичную матрицу S и перестановочную матрицу P, что выполняется равенство $K_{pub} = SG_{pub}P$.

Заметим, что если криптоаналитику удастся найти такую перестановочную матрицу P, что матрицы K_{pub} и $G_{pub}P$ порождают один и тот же код, тогда матрицу S можно найти, решая систему линейных алгебраических уравнений $K_{pub} = SG_{pub}P$.

Таким образом, задача криптоаналитика сводится к

поиску перестановочной матрицы, удовлетворяющей указанному свойству.

III. Атака на криптосистему

Построенная атака основывается на сигнатурном методе, предложенном в работе [10] Н. Сендрие. Суть метода состоит в построении некоторой эффективно вычислимой сигнатуры инварианта кода относительно группы подстановок координат кода, и восстановлении перестановки координат кода по значению сигнатуры.

Определение 5. Пусть $I_n = 1, \dots, n$. Сигнатура S на множестве F отображает код C длины n и элемент $i \in I_n$ в элемент F таким образом, что для любой подстановки σ на множестве I_n выполняется равенство

$$S(C, i) = S\big(C^\sigma, \sigma(i)\big),$$

здесь C^σ – код, который получается из кода C применение к каждому кодовому слову $c \in C$ подстановки σ следующим образом:

$$c^\sigma = \big(c_{\sigma(1)}, \dots, c_{\sigma(n)}\big).$$

Сигнатура называется полностью различающей, если для любых $i \neq j$ справедливо равенство $S(C, i) \neq S(C, j)$.

Если удастся построить эффективно вычислимую полностью различающую сигнатуру S для кода C, тогда по коду C^σ можно восстановить значение подстановки σ. Действительно, для этого нужно вычислить значения сигнатуры $S(C, i), i \in I_n$ для кода C и такие же значения $\{S(C^\sigma, j), j \in I_n\}$ для кода C^σ. Так как в сигнатура полностью различающая, то эти множества состоят из одних и тех же не повторяющихся элементах, поэтому $\sigma(i) = j$, если $S(C, i) = S(C^\sigma, j)$.

В работе [10] доказано, что полностью различающая сигнатура для кода C существует, если и только если код имеет тривиальную группу автоморфизмов, т. е. равенство $C = C^\sigma$ может быть выполнено только для тождественной подстановки σ. В реальных приложениях это, конечно, не так.

Если сигнатура не является полностью различающей, тогда множество $S(C, i), i \in I_n$ будет состоять из повторяющихся элементов, а значит его можно будет разбить на блоки, состоящие из одних и тех же значениях сигнатуры. Если таких блоков несколько, то сигнатура может дать некоторую информацию о подстановке σ. Действительно, можно утверждать, что $\sigma(i) \in j_1, \dots, j_p$, если $S(C, i) = S(C^\sigma, j_1) = \dots = S\big(C^\sigma, j_p\big)$. В этом случае можно сократить возможный перебор подстановок для поиска σ.

Для построения сигнатуры можно использовать инвариант ν кода, т. е. некоторое такое отображение кодов во множество F, что для любой подстановки σ выполнено равенство $\nu(C) = \nu(C^\sigma)$.

Имея эффективно вычислимый инвариант, можно попытаться построить сигнатуру, использую процедуру выкалывания координат кода: $S(C, i) = \nu(C_i)$.

Определение 6. Для $[n, k]$-кода C и координаты $i \in I_n$ код, получающийся из кода C выбором тех кодовых слов, у которых в координате с номером i содержится 0, называется выколотым кодом C_i.

Предлагаемая атака строит итерационно сигнатуру для комбинированного кода, порождаемого открытым ключом криптосистемы.

Для построения сигнатуры был выбран в качестве инварианта спектр оболочки кода.

Определение 7. Спектром кода C называется вектор $a_C = (a_0, \dots, a_n)$, в котором a_i – число кодовых слов веса Хэмминга i.

Определение 8. Оболочкой $\mathcal{H}(C)$ кода C называется код $C \cap C^\perp$.

На начальном этапе атаки строится сигнатура $S(C, i) = a_{\mathcal{H}(C_i)}$. Возможно, что она не будет полностью различающей. Однако, скорее всего будет существовать ряд координат кодовых слов, имеющих уникальное значение сигнатуры. В этом случае на этих координатах можно будет найти значение искомой подстановки – части секретного ключа.

Далее строится сигнатура для пары координат (i, j), в которой для одной из координаты, например i, известно значение подстановки. Эта сигнатура строится следующим образом: $S(C, \{i, j\}) = a_{\mathcal{H}(C_{i,j})}$, здесь $C_{i,j}$ – выколотый код одновременно по двум координатам, т. е. $C_{i,j} = (C_i)_j$. По уникальным значениям этой сигнатуры на парах координат можно восстановить значение подстановки σ на тех координатах j, для которых $S(C, \{i, j\})$ уникально для некоторой i.

Если не удастся восстановить всю подстановку, то процедура построения сигнатуры повторяется сначала на одной координате, а потом на двух, но уже для некоторого произведения Адамара кода, порождаемого матрицей открытого ключа.

Определение 9. Произведением Адамара кодов B и C одной и той же длины называется код, получающийся линейным замыканием множества $\{b \circ c | b \in B, c \in C\}$, здесь $b \circ c$ – покоординатное произведение векторов b и c.

Заметим, что произведение Адамара кодов может быть построено за $O(n^4)$ битовых операций, где n – длина кодов в произведении, на основе порождающих матриц кодов. Более подробно см. [6].

Опишем алгоритм атаки более подробно.

Алгоритм атаки.

Вход: матрицы K_{pub} и $G_{pub} = (G(r, m)|R)$.

Выход: подстановка σ, заданная перестановочная матрицей *P*.

1. Пусть C – код, порождаемый матрицей G_{pub}. Вычислить сигнатуру $S_1(C, i) = a_{\mathcal{H}(C_i)}$ для всех координат кодовых слов $1 \leq i \leq n$. И составить таблицу $T_1 = s_1 : i : S_1(C, i) = s_1$, здесь s_1 пробегает множество уникальных значений сигнатуры S_1.
2. Вычислить сигнатуру $S_2(C, i, j) = a_{\mathcal{H}(C_{i,j})}$ для всех координат кодовых слов $1 \leq i < j \leq n$. И составить таблицу $T_2 = s_2 : (i, j) : S_2(C, i, j) = s_2$, здесь s_2 пробегает множество уникальных значений сигнатуры S_2.
3. Пусть C^σ – код, порождаемый матрицей K_{pub}. Положим $U = \emptyset$. Для каждого номера $1 \leq i \leq n$ вычислить $s = S_1(C^\sigma, i)$. Если $T_1[s]$ содержит единственный номер j, тогда положить $\sigma(i) = j$, а номер i поместить во множество U.
4. Для всех пар $(i' \in U, i \notin U)$ вычислить $s = S_2(C^\sigma, i', i)$. Если $T_2[s]$ содержит единственную пару координат (j', j), тогда положить $\sigma(i) = j$, а номер i поместить во множество U.
5. Если $U \neq I_n$, то построить порождающую

матрицу кода $qC^\sigma = \underbrace{C^\sigma \circ C^\sigma \circ \ldots \circ C^\sigma}_{q}$, здесь $q = \lfloor m/r \rfloor$.

Повторить шаги 3 и 4, но уже для кода qC^σ. Заметим, что множество U при этом не опустошается, а постоянно пополняется. Если на каком-то шаге $U = I_n$, то необходимо перейти к шагу 7.

6. Пусть $\overline{U} = I_n \setminus U$. Не ограничивая общности, будем считать, что $\overline{U} = 1,2,\ldots,u$ для некоторого целого $u \geq 1$. Выбрать перестановку $(i_1, \ldots, i_u)$ элементов этого множества, положить $\sigma(j) = i_j$, $j = 1, \ldots, u$. Теперь подстановка σ определена на всех номерах множества I_n и по ней можно построить перестановочную матрицу P. Составить линейное уравнение $K_{pub} = \mathrm{S}G_{pub}\mathrm{P}$ относительно матрицы S, если оно разрешимо, то перейти к шагу 7. В противном случае выбирается другая перестановка элементов $1,2,\ldots,u$ и повторяется попытка найти матрицу S.
7. Выдать подстановку σ.

Ясно, что сложность предложенной атаки зависит от размерности оболочки кода с порождающей матрицей G_{pub}, от размерности оболочки кода qG_{pub}, который получается является произведением Адамара на шаге 5 алгоритма, а также от размера множества U после шага 5. Чем его размер ближе к размеру множества I_n, тем меньше подстановок опробывается на шаге 6 алгоритма. Проведём подробный анализ сложности каждого шага алгоритма атаки.

1. Вычисление $S_1(C, i)$ состоит из $2^{dim(\mathcal{H}(C_i))}$ подсчетов веса двоичного вектора длины n. Для построения таблицы T_1 необходимо вычислить n значений $S_1(C, i)$. Итого, сложность вычислений на первом шаге равна $2^{dim(\mathcal{H}(C_i))}n^2$ битовых операций.
2. Вычисление $S_2(C, i, j)$ состоит из $2^{dim(\mathcal{H}(C_{i,j}))}$ подсчетов веса двоичного вектора длины n. Для построения таблицы T_2 необходимо вычислить $\binom{n}{2}$ значений $S_2(C, i, j)$. Итого, сложность вычислений на первом шаге равна $\binom{n}{2}2^{dim(\mathcal{H}(C_{i,j}))}$ битовых операций.
3. На третьем шаге происходят вычисления, аналогичные вычислениям первого шага, поэтому сложность равна $2^{dim(\mathcal{H}(C_i))}n^2$ битовых операций.
4. На четвертом шаге происходят вычисления, аналогичные вычислениям второго шага, поэтому сложность равна $\binom{n}{2}2^{dim(\mathcal{H}(C_{i,j}))}$.
5. Построение qC^σ состоит из вычисления $q-1$ произведения Адамара, сложность каждого из которых $O(n^4)$ битовых операций. С учетом вычисленных сложностей третьего и четвертого шагов получаем сложность этого шага равной $O(qn^4) + 2^{dim(\mathcal{H}((qC^\sigma)_i))}n^2 + \binom{n}{2}2^{dim(\mathcal{H}((qC^\sigma)_{i,j}))}$ битовых операций.
6. Сложность последнего шага атаки определяется числом возможных подстановок на множестве U. При этом для каждой выбранной подстановки решается система линейных уравнений методом Гаусса, сложность которого равна $O(n^3)$. Итого, сложность данного шага равна $O(u!\,n^3)$.

Теорема 1. Сложность предложенной атаки равна $O\big((2^{h_0} + q)n^4 + u!\,n^3 + 2^{h_1}n^2\big)$ где h_0 – максимум размерности оболочки кодов C_i, $i = 1, \ldots, n$, и кодов $C_{i,j}$, $1 \leq i < j \leq n$, h_1 – максиму размерности оболочки кодов $(qC)_i$, $i = 1, \ldots, n$, и кодов $(qC)_{i,j}$, $1 \leq i < j \leq n$.

Доказательство. Пусть

$$h_0 = \max\left\{\max_{1\leq i\leq n} \dim \mathcal{H}(C_i), \max_{1\leq i<j\leq n} \dim \mathcal{H}(C_{i,j})\right\}$$

и

$$h_1 = \max\left\{\max_{1\leq i\leq n} \dim \mathcal{H}((qC)_i), \max_{1\leq i<j\leq n} \dim \mathcal{H}((qC)_{i,j})\right\}.$$

Тогда, складывая оценки сложности каждого шага алгоритма, с учётом введённых обозначений, а также с учётом того, что применение подстановки к коду не изменяет размерности оболочки, получим, что сложность предложенного алгоритма равна

$$2^{h_0}n^2(n^2 - n + 2) + O(qn^4) + 2^{h_1}(1.5n^2 - 0.5n) + O(u!\,n^3)$$

битовых операций. Далее, $n^2(n^2 - n + 2) = O(n^4)$ и $1.5n^2 - 0.5n = O(n^2)$, значит окончательно получим

$$O\big((2^{h_0} + q)n^4 + u!\,n^3 + 2^{h_1}n^2\big).$$

□.

Как видно, сложность атаки существенно зависит от размера оболочки кода C и его произведения Адамара qC.

Для начала обсудим вопрос размерности оболочки кода qC. Пусть $\mathcal{R}$ - код, порождаемый матрицей R. В работе [6] доказано, что $qRM(r, m) = RM(qr, m)$. Значит, в силу того, что $C \subseteq RM(r, m) \times \mathcal{R}$, где $\times$ - декартово произведение кодов. Тогда из свойств произведения Адамара следует, что $qC \subseteq RM(qr, m) \times q\mathcal{R}$. В работе [11] доказано, что с большой вероятностью $2\mathcal{R} = V_n$. Если $2r < m$, то $q > 1$, поэтому с большой вероятностью $q\mathcal{R} = V_n$. А значит с большой вероятностью $qC = RM(qr, m) \times V_n$, тогда $\mathcal{H}(qC) = RM^\perp \times \{0\}$. Верно равенство [8]

$$RM^\perp(qr, m) = RM(m - qr - 1, m).$$

Таким образом, с большой вероятностью $\dim \mathcal{H}(qC) = \dim RM(m - qr - 1, m)$. Учитывая, что $q = \lfloor m/r \rfloor$, получим, что $m - qr$ – остаток от деления m на r. Окончательно получаем, что

$$\dim \mathcal{H}(qC) = \dim RM(m \bmod r - 1, m).$$

Чем меньшее остаток от деления m на r тем меньше эта размерность.

В следующем разделе обсуждается размерность оболочки кода, с порождающей матрицей G_{pub}.

IV. Оболочка комбинированного кода

В основе криптосистемы, как видно из предыдущего пункта, лежит код, порождающая матрица которого равна G_{pub}. Этот код получен путем конкатенации порождающих матриц кода Рида–Маллера и случайного кода.

Исследуем размерность оболочки данного кода. Фактически результат этого пункта является естественным обобщением результата Н. Сендрие [12] о

размерности оболочки случайного линейного кода.

A. Число слабо самодуальных кодов

Определение 10. Код C называется слабо самодуальным, если $C \subseteq C^{\perp}$.

Следуя [12], найдём число слабо самодуальных кодов среди кодов, порождаемых матрицей вида $(G(r,m)|R)$, где $G(r,m)$ – порождающая матрица код Рида–Маллера, R – порождающая матрица случайного линейного код

Утверждение 1. Число слабо самодуальных кодов с порождающей матрицей вида $(G(r,m)|R)$, где $G(r,m)$ – порождающая матрица кода Рида-Маллера $RM(r,m)$ и $2r < m$, R –порождающая матрица некоторого линейного $[n,k]$-кода, равно числу слабо самодуальных $[n,k]$-кодов.

Доказательство. Пусть $C = (G(r,m)|R)$. Рассмотрим дуальный код $C^{\perp}$. Его порождающая матрица имеет вид:

$$\begin{pmatrix} G(m-r-1,m) & 0 \\ 0 & R^{\perp} \\ A & B \end{pmatrix},$$

где $G(m-r-1,m)$ – порождающая матрица кода $RM(m-r-1,m)$, т.к. этот коду является дуальным к коду Рида–Маллера $RM(r,m)$[8], $R^{\perp}$ - порождающая матрица кода, дуального коду с порождающей матрицей R, A и B – $(k \times n)$-матрицы максимального ранга, зависящие от кода с порождающей матрицей R. Известно [8], что $RM(r,m)$ лежит в своем дуальном коде, если $2r < m$. Таким образом, код с порождающей матрицей $(G(r,m)|0)$ лежит в своем дуальном. Значит, если код с порождающей матрице R слабо самодуальный, то и код C также слабо самодуальный.

Докажем теперь обратное, что если код C слабо самодуален, то этим же свойством должен обладать и код порождающей матрицей R.

Рассмотрим вектор $(u|v) \in C$, т. к. по условию $C \subseteq C^{\perp}$, то в силу вида порождающей матрицы кода $C^{\perp}$ найдутся такие векторы $\alpha, \beta, \gamma \in V_k$, что

$$(u|v) = (\alpha G(m-r-1,m) + \gamma A | \beta R^{\perp} + \gamma B).$$

Вектор $u \in RM(r,m) \subseteq RM^{\perp}(r,m) = RM(m-r-1,m)$, поэтому $\gamma = 0$. Но это означает, что $v = \beta R^{\perp}$ для некоторого $\beta \in V_k$. Так как v пробегает все векторы кода с порождающей матрицей R, то этот код также слабо самодуален.

Таким образом, доказано, что условие слабой самодуальности кода C эквивалентно условию слабой самодуальности кода с порождающей матрицей R, а так как это матрица произвольного $[n,k]$-кода, то число слабо самодуальных кодов C равно общему числу слабо самодуальных $[n,k]$-кодов. □

Число слабо самодуальных линейных $[n,k]$-кодов над $GF(2)$ обозначим за $\sigma_{n,k}$.

Определение. Пусть n и k – два таких целых числа, что $n \geq k \geq 0$. Тогда биномиальным коэффициентом Гаусса называется число:

$$\begin{bmatrix} n \\ k \end{bmatrix} = \frac{(2^n-1)(2^{n-1}-1)\dots(2^{n-k+1}-1)}{(2^k-1)(2^{k-1}-1)\dots(2-1)}.$$

Далее будем считать, что, если $k < n$, то $\begin{bmatrix} n \\ k \end{bmatrix} = 0$.

Используя свойства биномиального коэффициента Гаусса, можно записать:

$$\begin{bmatrix} n \\ k \end{bmatrix} = 2^{k(n-k)} \frac{g_{2,n}}{g_{2,k} g_{2,n-k}}, \quad (1)$$

где последовательность $(g_{q,n})_{n \geq 0}$ определена для любого $q > 1$ следующим образом:

$$g_{q,n} = \prod_{i=1}^{n} \left(1 - \frac{1}{q^i}\right).$$

Утверждение 2 ([12]). (Формула обращения).

Пусть даны две последовательности $(u_i)_{i \geq 0}$и $(v_i)_{i \geq 0}$ целых чисел. Тогда для всех $k \geq 0$

$$\forall l, 0 \leq l \leq k, \qquad v_i = \sum_{i=0}^{l} \begin{bmatrix} l \\ i \end{bmatrix} u_i \Leftrightarrow$$

$$\forall l, \qquad 0 \leq l \leq k, \qquad u_i = \sum_{i=0}^{l} \begin{bmatrix} l \\ i \end{bmatrix} (-1)^{l-i} 2^{\binom{l-i}{2}} v_i$$

Утверждение 3 ([12]).

Последовательность $(g_{q,n})_{n \geq 0}$строго убывающая для всех. Обозначив за $g_{q,\infty}$ предел этой последовательности при n, стремящемся к бесконечности, получим

$$\frac{g_{q,\infty}}{g_{q,n}} = \sum_{i \geq 0} \frac{1}{q^{ni}} \frac{(-1)^i}{(q-1)(q^2-1)\dots(q^i-1)}.$$

Следствие 1 ([12]).

Для всех целых $n \geq 0$

$$1 - \frac{1}{q^n} \leq \frac{g_{q,\infty}}{g_{q,n}} \leq 1$$

Утверждение 4 ([12]).

Пусть $m = \lfloor n/2 \rfloor$Для всех $k \leq m$,

$$\sigma_{n,k} = s_{n,k} \frac{2^{k(n-k)}}{2^{k(k+1)/2}} \frac{g_{4,m}}{g_{4,m-k} g_{2,k}}, \text{где}$$

$$s_{n,k} = \begin{cases} 1, & \text{если } n \text{ нечётное} \\ \dfrac{2^n - 2^k}{2^n - 1}, & \text{если } n \text{ чётное} \end{cases}.$$

.

Утверждение 5. Для всех $k \leq n/4$,

$$1 - \frac{1}{2^{n/4-k}} \leq s_{n/2,k} \leq 1 + \frac{1}{2^{n/4-k}}.$$

Доказательство. Для $k \leq n/4$:

$$1 - \frac{1}{2^{\frac{n}{4}-k}} \leq \frac{2^{\frac{n}{4}}-2^k}{2^{\frac{n}{4}}-1} \leq \frac{2^{\frac{n}{2}}-2^k}{2^{\frac{n}{2}}-1} \leq 1 \leq \frac{2^{\frac{n}{4}}+2^k}{2^{\frac{n}{2}}+1} \leq 1 + \frac{1}{2^{\frac{n}{4}-k}}.\square$$

Таким образом, из утверждений 1 и 4 следует, что число слабо самодуальных кодов среди рассматриваемого класса кодов равно $\sigma_{n,k}$.

B. Асимптотические оценки

Пусть ℓ - размерность оболочки кода с порождающей матрицей $(G(r,m)|R)$. За $A_{n,k,\ell}$ обозначим количество $[2n,k]$-кодов, размерность оболочки которых равна ℓ.

Введём для всех ℓ, $0 \leq \ell \leq k$, две последовательности

$$b_{n,k,\ell} = \frac{\begin{bmatrix} n-2k+2\ell \\ \ell \end{bmatrix} \sigma_{n,k-\ell} 2^{k(k+1)/2}}{\begin{bmatrix} n \\ k \end{bmatrix} \begin{bmatrix} k \\ \ell \end{bmatrix}}$$

и

$$a_{n,k,\ell} = \frac{A_{n,k,k-\ell} 2^{k(k+1)/2}}{\begin{bmatrix} n \\ k \end{bmatrix} \begin{bmatrix} k \\ \ell \end{bmatrix}}.$$

Утверждение 6. Пусть $m = \lfloor n/4 \rfloor$. Для всех $k \leq m$ и для всех $\ell, 0 \leq \ell \leq k$:

$$b_{n,k,\ell} = 2^{\ell(\ell+1)/2} \frac{g_{4,m} g_{2,n-2k+2\ell} g_{2,n-k}}{g_{4,m-k+\ell} g_{q,n-2k+\ell} g_{q,n}} s_{n,k-\ell}.$$

Доказательство. Из утверждения 4 получаем, что

$$\sigma_{n,k-\ell} = \frac{2^{(k-\ell)(n-k+\ell)}}{2^{(k-\ell)(k-\ell+1)/2}} \frac{g_{4,m}}{g_{4,m-k+\ell} g_{2,k-\ell}} s_{n,k-\ell}.$$

Из (1) следует, что

$$\begin{bmatrix} n-2k+2\ell \\ \ell \end{bmatrix} = 2^{\ell(n-2k+\ell)} \frac{g_{2,n-2k+2\ell}}{g_{2,\ell} g_{2,n-2k+\ell}},$$

и

$$\begin{bmatrix} n \\ k \end{bmatrix} \begin{bmatrix} k \\ \ell \end{bmatrix} = 2^{k(n-k)+\ell(k-\ell)} \frac{g_{2,n} g_{2,k}}{g_{2,k} g_{2,n-k} g_{2,\ell} g_{2,k-\ell}},$$

и используя тот факт, что

$$\frac{k(k+1)}{2} - \frac{\ell(\ell+1)}{2} = \frac{(k-\ell)(k-\ell+1)}{2} + \ell(k-\ell),$$

получаем

$$\frac{\begin{bmatrix} n-2k+2\ell \\ \ell \end{bmatrix} \sigma_{n,k-\ell}}{\begin{bmatrix} n \\ k \end{bmatrix} \begin{bmatrix} k \\ \ell \end{bmatrix}} = \frac{2^{\ell(\ell+1)/2}}{2^{k(k+1)/2}} \frac{g_{4,,m} g_{2,n-2k+2\ell} g_{2,n-k}}{g_{4,m-k+\ell} g_{2,n-2k+\ell} g_{2,n}} s_{n,k-\ell}.$$

□

Лемма 1. Пусть $m = \lfloor n/4 \rfloor$, для всех $i \leq m$,

$$\frac{g_{4,m} g_{2,n-i}}{g_{4,m-i} g_{2,n}} \leq 1.$$

Доказательство. По определению последовательности $g_{2,n}$ и $g_{4,n}$ получим

$$\frac{g_{4,m} g_{2,n-i}}{g_{4,m-i} g_{2,n}} = \frac{\prod_{j=m-i+1}^{m} 1-\frac{1}{2^{2j}}}{\prod_{j=n-i+1}^{n} 1-\frac{1}{2^{j}}} = \prod_{j=1}^{i} \frac{1-1/2^{2m-2j+2}}{1-1/2^{n-j+1}}.$$

Для всех j, $1 \leq j \leq i$, вне зависимости от чётности n получим

$$1 - \frac{1}{2^{2m-2j+2}} \leq 1 - \frac{1}{2^{n-j+1}},$$

что завершает доказательство. □

Лемма 2. Пусть $m = \lfloor n/4 \rfloor$, для всех $i \leq m$,

$$\frac{g_{4,\infty} g_{2,n-2i}}{g_{4,m-i} g_{2,\infty}} \leq 1.$$

Доказательство. По определению получаем

$$\frac{g_{4,\infty} g_{2,n-2i}}{g_{4,m-i} g_{2,\infty}} = \frac{\prod_{j \geq m-i} 1-\frac{1}{2^{2j}}}{\prod_{j \geq n-2i} 1-\frac{1}{2^{j}}} = \prod_{j \geq 0} \frac{1-1/2^{2m-2i+2j}}{1-1/2^{n-2i+j}}.$$

Для всех $j > 0$ вне зависимости от четности n выполняются неравенства

$$1 - \frac{1}{2^{2m-2i+2j}} \geq 1 - \frac{1}{2^{n-2i+j}},$$

что завершает доказательство. □

Утверждение 7. Пусть $\delta_{n,k,\ell} = 2^{\ell(\ell+1)/2} - b_{n,k,\ell}$. Для всех $\ell, 0 \leq \ell \leq k$, верно:

$$-\frac{2^{\ell(\ell-1)/2}}{2^{n-k}} \leq \delta_{n,k,\ell} \leq 2\frac{2^{\ell(\ell-1)/2}}{2^{n-k}}.$$

Доказательство. Пусть $m = \lfloor n/4 \rfloor$, из утверждения 6 следует равенство

$$b_{n,k,\ell} = 2^{\ell(\ell+1)/2} \frac{g_{4,m} g_{2,n-2k+2\ell} g_{2,n-k}}{g_{4,m-k+\ell} g_{2,n-2k+\ell} g_{2,n}} s_{n,k-\ell},$$

Используя леммы 1 и 2 и тот факт, что последовательности $g_{2,n}$ и $g_{4,n}$ убывающие, получаем

$$\frac{g_{2,\infty}}{g_{2,n-2k+\ell}} s_{n,k-\ell} \leq \frac{b_{n,k,\ell}}{2^{\ell(\ell+1)/2}} \leq s_{n,k-\ell}.$$

Из следствия 1 и утверждения 5 получаем

$$L = \left(1 - \frac{1}{2^{n-2k+\ell}}\right)\left(1 - \frac{1}{q^{n/2-k+\ell}}\right) \leq \frac{b_{n,k,\ell}}{2^{\frac{\ell(\ell+1)}{2}}} \leq 1 + \frac{1}{2^{n-k+\ell}}.$$

Рассмотрим левую часть неравенства

$$L \geq 1 - \frac{1}{2^{\frac{n}{2}-2k+\ell}} - \frac{1}{2^{\frac{n}{2}-k+\ell}} \geq 1 - \frac{1}{2^{\frac{n}{2}-k+\ell}} - \frac{1}{q^{\frac{n}{2}-k+\ell}} =$$

$$= 1 - \frac{2}{q^{\frac{n}{2}-k+\ell}}$$

и в итоге

$$\left(1 - \frac{1}{(2^{n/2-k+\ell}}\right) \leq \frac{b_{n,k,\ell}}{2^{\frac{\ell(\ell+1)}{2}}} \leq 1 + \frac{1}{2^{\frac{n}{2}-k+\ell}},$$

что завершает доказательство. □

Утверждение 8 ([12]). Пусть последовательность $(u_\ell)_{\ell \geq 0}$ является решением уравнения

$$\sum_{i=0}^{\ell} \begin{bmatrix} \ell \\ i \end{bmatrix} u_i = 2^{\ell(\ell+1)/2},$$

и пусть $\gamma_{n,k,\ell} = u_\ell - a_{n,k,\ell}$. Тогда для всех ℓ, $0 \leq \ell \leq k$, выполнено

$$\sum_{i=0}^{\ell} \begin{bmatrix} \ell \\ i \end{bmatrix} \gamma_{n,k,i} = \delta_{n,k,\ell}.$$

Из формулы обращения (утверждение 2) в условиях утверждения 8 получим

$$u_\ell = \sum_{i=0}^{\ell} \begin{bmatrix} \ell \\ i \end{bmatrix} (-1)^{\ell-i} 2^{\binom{i+1}{2}+\binom{\ell-i}{2}}.$$

Лемма 3 ([12]). Для всех $\ell \geq 0$ справедливо введём последовательность

$$w_\ell = \sum_{i=0}^{\ell} \begin{bmatrix} \ell \\ i \end{bmatrix} (-1)^{\ell-i} 2^{\binom{i+1}{2}+\binom{\ell-i+1}{2}}.$$

Тогда

$$w_\ell = \begin{cases} 0, & \text{если } \ell \text{ нечетное} \\ u_{\ell,} & \text{иначе} \end{cases}.$$

Утверждение 9 ([12]). Пусть последовательность $(u_\ell)_{\ell \geq 0}$ является решением уравнения

$$\sum_{i=0}^{\ell} \begin{bmatrix} \ell \\ i \end{bmatrix} u_i = 2^{\ell(\ell+1)/2}.$$

Тогда для всех $\ell \geq 0$

$$w_l = \prod_{0 \leq i \leq \ell, i-\text{четное}} 2^i \prod_{0 \leq i \leq \ell, i-\text{нечетное}} (2^i - 1)$$

или, что эквивалентно, $u_0 = 1$ и для всех $\ell > 0$

$$u_\ell = \begin{cases} u_{\ell-1} 2^\ell & \text{если } \ell \text{ нечетное} \\ u_{\ell-1}(2^\ell - 1) & \text{иначе} \end{cases}.$$

Следствие 2 ([12]). Для всех $\ell \geq 0$

$$u_\ell = 2^{\ell(\ell+1)/2} \frac{g_{2,\ell}}{g_{4,\lfloor \ell/2 \rfloor}}.$$

Утверждение 10. Для всех $\ell, 0 \leq \ell \leq k$:

$$\left|\gamma_{n,k,\ell}\right| \leq 2(\ell+1) \frac{2^{\ell(\ell-1)/2}}{g_{2,\lfloor \ell/2 \rfloor} 2^{n/2-k}}.$$

Доказательство. Формула обращения дает

$$\gamma_{n,k,\ell} = \sum_{i=0}^{\ell} \begin{bmatrix} \ell \\ i \end{bmatrix} (-1)^{\ell-i} 2^{\binom{\ell-i}{2}} \delta_{n,k,i}.$$

Откуда из утверждения 7 получим неравенство:

$$\left|\gamma_{n,k,\ell}\right| \leq \sum_{i=0}^{\ell} \begin{bmatrix} \ell \\ i \end{bmatrix} 2^{\binom{\ell-i}{2}} \left|\delta_{n,k,i}\right| \leq \frac{2}{2^{\frac{n}{2}-k}} \sum_{i=0}^{\ell} \begin{bmatrix} \ell \\ i \end{bmatrix} 2^{\binom{\ell-i}{2}+\binom{i}{2}}.$$

В силу того, что $\binom{\ell-i}{2} + \binom{i}{2} = \binom{\ell}{2} - i(\ell-i)$

получаем

$$\left|\gamma_{n,k,\ell}\right| \leq 2\frac{2^{\binom{\ell}{2}}}{2^{n/2-k}} \sum_{i=0}^{\ell} \begin{bmatrix} \ell \\ i \end{bmatrix} \frac{1}{2^{i(\ell-i)}} = 2\frac{2^{\binom{\ell}{2}}}{2^{n/2-k}} \sum_{i=0}^{\ell} \frac{g_{2,\ell}}{g_{2,i} g_{2,\ell-i}}.$$

Итого, учитывая, что $g_{2,\ell} \leq g_{2,i}$ и $g_{2,\ell} \leq g_{2,\ell-i}$, получим

$$\frac{g_{2,\ell}}{g_{2,i} g_{2,\ell-i}} \leq \min\left(\frac{1}{g_{2,i}}, \frac{1}{g_{2,\ell-i}}\right) \leq \frac{1}{g_{2,\left\lfloor \frac{\ell}{2} \right\rfloor}}.$$

Отсюда

$$\left|\gamma_{n,k,l}\right| \leq 2\frac{2^{\binom{\ell}{2}}}{2^{\frac{n}{2}-k}} \frac{\ell+1}{g_{2,\left\lfloor \frac{\ell}{2} \right\rfloor}}.$$

□

Следствие 3. Существует константа K, что для всех $\ell,\ 0 \le \ell \le k$,

$$\frac{|\gamma_{n,k,\ell}|}{u_\ell} \le K \frac{k}{2^{\frac{n}{2}-k+\ell}}.$$

Доказательство. Из утверждения 10 и следствия 2 можно легко найти эту константу K. □

Используя полученные результаты легко найти размерность оболочки комбинированного кода.

Теорема 2. Пусть n – целое положительное число. Для всех $k \le n/2,\ \ell \le k$ число кодов, порождаемых матрицей $(G(r,m)|R)$, размерность оболочки которых равна ℓ, равно

$$A_{n,k,\ell} = \begin{bmatrix} n \\ k \end{bmatrix} \frac{1}{2^{\ell(\ell+1)/2}} \frac{g_{2,k}}{g_{4,\lfloor (k-\ell)/2 \rfloor} g_{2,\ell}} \left(1 + O\left(\frac{k}{2^{n/2-\ell}}\right)\right).$$

Доказательство. По определению

$$A_{n,k,k-\ell} 2^{\frac{k(k+1)}{2}} = \begin{bmatrix} n \\ k \end{bmatrix} \begin{bmatrix} k \\ \ell \end{bmatrix} a_{n,k,\ell},$$

где $a_{n,k,k-\ell} = u_{k-\ell} + \gamma_{n,k,k-\ell}$. По следствию 3 получаем

$$A_{n,k,\ell} q^{k(k+1)/2} = \begin{bmatrix} n \\ k \end{bmatrix} \begin{bmatrix} k \\ \ell \end{bmatrix} u_{k-\ell} \left(1 + O\left(\frac{k}{2^{n/2-\ell}}\right)\right).$$

Из следствия 2 и (1) получаем утверждение теоремы. □

Пусть задано вероятностное пространство, множество элементарных исходов Ω которого состоит из всех возможных размеров ℓ оболочки кода C, порождаемого матрицей $C = (G(r,m)|R)$.

По построению количество различных кодов C длины $2n$ и размерности k равно числу различных кодов длины n и размерности k. Количество двоичных линейных $[n,k]$– кодов равно $\begin{bmatrix} n \\ k \end{bmatrix}$. В вероятностном пространстве Ω можно ввести дискретную случайную величину $h = \dim(\mathcal{H}(C))$. Тогда, с учетом теоремы 1, её распределение вероятности определяется следующим образом:

$$p_\ell = \Pr\{dim(\mathcal{H}(C)) = \ell\} = \frac{A_{n,k,\ell}}{\begin{bmatrix} n \\ k \end{bmatrix}}$$

Тогда её математическое ожидание равно

$$M\mathrm{h} = \sum_{\ell=1}^{k} \frac{\ell}{2^{\ell(\ell+1)/2}} \frac{g_{2,k}}{g_{4,\lfloor (k-\ell)/2 \rfloor} g_{2,\ell}} \left(1 + O\left(\frac{k}{2^{n/2-\ell}}\right)\right) \quad (2)$$

V. РЕЗУЛЬТАТЫ ПРОГРАММНОГО МОДЕЛИРОВАНИЯ

Атака была реализована программно на языке C++. Исследования проводились на кодах с параметрами $(k \times 2n) = \{(22 \times 128), (93 \times 512), (130 \times 1024)\}$.

С использованием формулы (2) было вычислено, что с наибольшей вероятностью размер оболочки для исследуемых кодов равен 1, а вероятность получить размерность больше трех близка к нулю.

Ввиду небольшого размера оболочки рассматриваемого комбинированного кода можно надеяться на эффективность предложенной атаки.

Для рассматриваемых примеров размерность q произведений Адамара, строящегося на шаге пять, соответственно равна $\{128(q=3),\ 512(q=2), 1024(q=3)\}$. И как видно, она относительно невилика.

Ниже в таблице представлены результаты проведенных экспериментов. Для минимальных значений размера порождающей матрицы кода результаты получены на 1000 матрицах, для больших размеров – на 10 матрицах.

Размер K_{pub} $(k \times 2n)$ (Кбайт)	S_1	S_2	Размер оболочки q произведений Адамара	$\|U\|$	Сложность полного перебора	Сложность атаки
(22×128) (5.6 Кб)	1	1	1	7	2^{81}	2^{33}
(93× 512) (95.2 Кб)	1	1	9	10	2^{350}	2^{49}
(130×1024) (266.2 Кб)	1	1	1	21	2^{560}	2^{96}

Построенная атака была применена и к оригинальной криптосистеме Кабатянского–Тавернье [7], в которой код Рида-Маллера комбинируется с кодом Гоппы.

Результаты экспериментов приведены в таблице ниже. Данные получены на следующих значениях параметров (r, m_1) кода Рида-Маллера и $(m_2, t, |L|)$ – параметров кода Гоппы:

$[\{(2,6),(6,7,64)\},\{(3,8)(7,5,128)\},\{(3,9)(8,10,209)\}]$.

Аналогично для небольших размеров параметров данные получены на 100 матриц, для больших – на 5.

Размер K_{pub} $(k \times 2n)$ (Кбайт)	S_1	S_2	Размер оболочки q произведений Адамара	$\|U\|$	Сложность полного перебора	Сложность атаки
(22×128) (5.6 Кб)	1	1	1	10	2^{81}	2^{43}
(93×384) (35.7 Кб)	1	1	9	23	2^{350}	2^{97}
(130×721) (93.7 Кб)	1	1	1	38	2^{560}	2^{176}

VI. ЗАКЛЮЧЕНИЕ

В работе рассматривается криптосистема типа Мак-Элиса–Сидельникова, построенная на комбинировании кодов Рида–Маллера, случайных линейных двоичных кодов или кодов Гоппы. На такую криптосистему была построена структурная атака, восстанавливающая секретный ключ в модели, когда противнику известна порождающая матрица случайного кода. Атака показывает эффективность на достаточно больших открытых ключах, поэтому на практике размеры ключей должны быть значительно увеличены.

В целом авторы полагают, что модификация Кабатянского–Тавернье не только не увеличивает стойкости криптосистемы Мак-Элиса, но и понижает её в модели, в которой фиксированы используемые коды. Так атаки на основе алгоритма разделения носителя не применимы к кодам Рида–Маллера или к кодам Гоппы, но комбинация этих кодов вносит полезные свойства криптосистемы, которые противник может использовать для построения атак.

Один из выходов для защиты от такого рода атак – оставлять в секрете параметры кодов Гоппы при использовании криптосистемы на практике.

Однако авторы считают, что описанная техника может

быть усовершенствована для построения эффективных атак в модели с неизвестными параметрами кодов, на основе которых строится криптосистема.

Библиография

[1] R. J. Mceliece, «A Public-Key Cryptosystem Based on Algebraic Coding Theory», *JPL DSN Prog. Rep.*, т. 44, сс. 123–125, 1978.

[2] P. W. Shor, «Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer», *SIAM J. Comput.*, т. 26, вып. 5, сс. 1484–1509, окт. 1997, doi: 10.1137/S0097539795293172.

[3] N. T. Courtois, M. Finiasz, и N. Sendrier, «How to Achieve a McEliece-Based Digital Signature Scheme», в *Advances in Cryptology — ASIACRYPT 2001*, Berlin, Heidelberg, 2001, сс. 157–174, doi: 10.1007/3-540-45682-1_10.

[4] «В. М. Сидельников, “Открытое шифрование на основе двоичных кодов Рида–Маллера”, Дискрет. матем., 6:2 (1994), 3–20; Discrete Math. Appl., 4:3 (1994), 191–207», Просмотрено: апр. 25, 2020. [Онлайн]. Доступно на: http://www.mathnet.ru/php/archive.phtml?wshow=paper&jrnid=dm&paperid=637&option_lang=rus.

[5] L. Minder и A. Shokrollahi, «Cryptanalysis of the Sidelnikov Cryptosystem», в *Advances in Cryptology - EUROCRYPT 2007*, т. 4515, M. Naor, Ред. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, сс. 347–360.

[6] М. А. Бородин, M. A. Borodin, И. В. Чижов, и I. V. Chizhov, «Эффективная атака на криптосистему Мак-Элиса, построенную на основе кодов Рида - Маллера», *Дискретная Математика*, т. 26, вып. 1, сс. 10–20, 2014, doi: 10.4213/dm1264.

[7] E. Egorova, G. Kabatiansky, E. Krouk, и C. Tavernier, «A new code-based public-key cryptosystem resistant to quantum computer attacks», *J. Phys. Conf. Ser.*, т. 1163, с. 012061, фев. 2019, doi: 10.1088/1742-6596/1163/1/012061.

[8] Ф. Дж. Мак-Вильямс и Н. Дж. А. Слоэн, *Теория кодов, исправляющих ошибки*. Москва: Связь, 1979.

[9] E. Berlekamp, R. J. McEliece, и H. C. van Tilborg, «On the Inherent Intractability of Certain Coding Problems», *Inf. Theory IEEE Trans. On*, т. 24, вып. 3, сс. 384–386, 1978, doi: 10.1109/TIT.1978.1055873.

[10] N. Sendrier, «Finding the permutation between equivalent linear codes: the support splitting algorithm», *IEEE Trans. Inf. Theory*, т. 46, вып. 4, сс. 1193–1203, июл. 2000, doi: 10.1109/18.850662.

[11] D. Mirandola и G. Zémor, «Schur products of linear codes: a study of parameters», *Diss Master Thesis Superv. G Zémor Univ Bordx.*, 2012.

[12] N. Sendrier, «On the Dimension of the Hull», *SIAM J. Discrete Math.*, т. 10, вып. 2, сс. 282–293, май 1997, doi: 10.1137/S0895480195294027.

Structural attack on McEliece-Sidelnikov type public-key cryptosystem based on a combination of random codes with Reed-Muller codes

Ivan Chizhov, Elizaveta Popova

***Abstract* — This paper represents the investigation of McEliece-Sidelnikov cryptosystem, based on combination of random codes with Reed-Muller codes. Different modifications of classical McEliece cryptosystem has been studied. Sidelnikov's work introduced using the several samples of Reed-Muller code, and Kabatiansky and Tavernier's work proposed to use the concatenation of Goppa and Reed-Muller codes. The popularity of this cryptosystem explains with the fact that it's strength is based on the hardness of the decoding general linear code problem, so it will remain unbreakable in postquantum era. This paper investigates one of the modifications of McElice cryptosystem in a model when the attacker knows the public-key matrix and the generator matrix of random linear code. The goal is to reconstruct the permutation matrix from the secret-key. During the investigation of the hull of the code built by combining random codes with Reed-Muller codes the theorem about the number of such codes with fixed size of the hull has been proved. An attack based on signature method has been produced and programmed with the use of C++ programming language. All the results of this program's work are represented in this paper. With the use of proved theorem the hardness of provided attack has been calculated.**

Key words — postquantum cryptography, code-based cryptosystems, McEliece public key cryptosystem, Sidelnikov public key cryptosystem, combined linear codes, Reed-Muller codes.

REFERENCES

[1] R. J. Mceliece, «A Public-Key Cryptosystem Based on Algebraic Coding Theory», JPL DSN Prog. Rep., t. 44, ss. 123–125, 1978.
[2] P. W. Shor, «Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer», SIAM J. Comput., t. 26, vyp. 5, ss. 1484–1509, okt. 1997, doi: 10.1137/S0097539795293172.
[3] N. T. Courtois, M. Finiasz, i N. Sendrier, «How to Achieve a McEliece-Based Digital Signature Scheme», v Advances in Cryptology — ASIACRYPT 2001, Berlin, Heidelberg, 2001, ss. 157–174, doi: 10.1007/3-540-45682-1_10.
[4] «V. M. Sidel'nikov, "Otkrytoe shifrovanie na osnove dvoichnyh kodov Rida–Mallera", Diskret. matem., 6:2 (1994), 3–20; Discrete Math. Appl., 4:3 (1994), 191–207», Prosmotreno: apr. 25, 2020. [Onlajn]. Dostupno na: http://www.mathnet.ru/php/archive.phtml?wshow=paper&jrnid=dm&paperid=637&option_lang=rus.
[5] L. Minder i A. Shokrollahi, «Cryptanalysis of the Sidelnikov Cryptosystem», v Advances in Cryptology - EUROCRYPT 2007, t. 4515, M. Naor, Red. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, ss. 347–360.
[6] M. A. Borodin, M. A. Borodin, I. V. Chizhov, i I. V. Chizhov, «Jeffektivnaja ataka na kriptosistemu Mak-Jelisa, postroennuju na osnove kodov Rida - Mallera», Diskretnaja Matematika, t. 26, vyp. 1, ss. 10–20, 2014, doi: 10.4213/dm1264.
[7] E. Egorova, G. Kabatiansky, E. Krouk, i C. Tavernier, «A new code-based public-key cryptosystem resistant to quantum computer attacks», J. Phys. Conf. Ser., t. 1163, s. 012061, fev. 2019, doi: 10.1088/1742-6596/1163/1/012061.
[8] F. Dzh. Mak-Vil'jams i N. Dzh. A. Slojen, Teorija kodov, ispravljajushhih oshibki. Moskva: Svjaz', 1979.
[9] E. Berlekamp, R. J. McEliece, i H. C. van Tilborg, «On the Inherent Intractability of Certain Coding Problems», Inf. Theory IEEE Trans. On, t. 24, vyp. 3, ss. 384–386, 1978, doi: 10.1109/TIT.1978.1055873.
[10] N. Sendrier, «Finding the permutation between equivalent linear codes: the support splitting algorithm», IEEE Trans. Inf. Theory, t. 46, vyp. 4, ss. 1193–1203, ijul. 2000, doi: 10.1109/18.850662.
[11] D. Mirandola i G. Zémor, «Schur products of linear codes: a study of parameters», Diss Master Thesis Superv. G Zémor Univ Bordx., 2012.
[12] N. Sendrier, «On the Dimension of the Hull», SIAM J. Discrete Math., t. 10, vyp. 2, ss. 282–293, maj 1997, doi: 10.1137/S0895480195294027.