КРИПТОСИСТЕМА МАК-ЭЛИСА И ПРОБЛЕМЫ ЕЁ ВНЕДРЕНИЯ

MCELIECE'S CRYPTOSYSTEM AND PROBLEMS OF ITS IMPLEMENTATION

УДК 004.453.2

Баев Дмитрий Александрович, Студент кафедры кибербезопасности информационных систем, Донской государственный технический университет, г. Ростов-на-Дону

Baev D.A. fantom-sj@yandex.ru

Аннотация

В статье рассматривается криптографическая система Мак-Элиса, как кандидат на криптосистему, устойчивую к атакам, основанным на квантовых алгоритмах. Также рассматриваются проблемы, обуславливающие отсутствия широкого распространения данной криптосистемы в настоящее время. С целью их решения было проведено исследование данного алгоритма на криптоустойчивость, в рамках которого были определены рекомендуемые параметры кодов Гоппа, необходимых для построения достаточно безлопастного алгоритма шифрования Мак-Элиса. Также был рассмотрен и сам алгоритм шифрования Мак-Элиса, построена его общая схема и выявлены эффективные алгоритмы для декодирования кодов Гоппа лежащих в его основе.

Annotation

The article discusses the McEliece cryptographic system as a candidate for a cryptosystem resistant to attacks based on quantum algorithms. It also discusses the problems that cause the lack of widespread use of this cryptosystem at the present time. In order to solve them, a study of this algorithm for cryptographic stability was carried out, within the framework of which the recommended parameters of the Gopp codes

were determined, which are necessary to build a sufficiently bladeless McEliece encryption algorithm. The McEliece encryption algorithm itself was also considered, its general scheme was built and effective algorithms for decoding the Gopp codes underlying it were revealed.

Ключевые слова: криптография, постквантовая криптография, асимметричная криптография, криптосистема Мак-Элиса, коды Гоппа.

Keywords: cryptography, post-quantum cryptography, asymmetric cryptography, McEliece cryptosystem, Gopp codes.

Криптосистема Мак-Элиса – это криптосистема с открытым ключом, основанная на сложности декодирования полных линейных кодов, которая имеет очень высокий уровень безопасности [1]. Она была предложена в 1978 г. Р. Дж. Мак-Элисом [2].

Основная идея построения криптосистемы состоит в маскировке некоторого линейного кода, под код, не обладающий видимой алгебраической и комбинаторной структурой. Такие коды принято называть кодами общего положения [3]. Предполагается, что декодирование кода общего положения является трудной задачей. Не зная структуры кода, невозможно построить эффективный алгоритм декодирования такого кода. Именно эта идея и заложена в конструкции криптосистемы, предложенной Мак-Элисом [3].

Первоначальная криптосистема Мак-Элиса все еще не взломана, а именно атака, позволяющая взломать данный шифр полностью, до сих пор не найдена. Также не были найдены и атаки локального характера, что делает задачу взлома данной системы на практике практически неразрешимыми. Более того, эта система на два-три порядка быстрее, чем конкурирующие решения, такое как криптосистема RSA, которая в настоящее время является одним из самых популярных алгоритмов с использованием открытого ключа [1].

Несмотря на это, криптосистема Мак-Элиса редко используется на практике. Это связано с тем, что она имеет два основных недостатка [1]:

– большой размер открытого ключа;

– низкая скорость передачи (около 0,5).

В оригинальном варианте криптосистемы Мак-Элиса используются коды Гоппа длиной n = 1024, размерностью k = 524 и минимальным кодовым расстоянием d не менее 101, которые могут исправить t = 50 ошибок. Позднее было предпринято несколько попыток преодоления недостатков исходной системы для устранения недостатков данной системы, но принятие альтернативных семейств кодов оказалось невозможным без ущерба для её безопасности [2].

Одной из таких попыток стало использование кодов проверки на четность с низкой плотностью или LDPC кодов [4]. Использование данного семейства кодов решает проблему длины открытого ключа, позволяя значительно её сократить. Также с применением данных кодов можно значительно ускорить процесс декодирования, так как коды LDPC обладают хорошей масштабируемостью в аппаратных реализациях, так как у них намного проще алгоритмы декодирования. Тем не менее, использование кодов LDPC в криптосистеме Мак-Элиса не считается таким же безопасным, как использование кодов Гоппа в оригинальной реализации [4].

Процесс шифрования информации в криптосистеме Мак-Элиса начинается с выбора такого корректирующего кода, способного исправить заданное число ошибок, для которого известен эффективный алгоритм декодирования. Далее выбранный код замаскировывается под обычный линейный код, для которого невозможно подобрать эффективного алгоритма декодирования, способно справиться с поставленной задачей за обозримый промежуток времени [5].

При этом, как отправитель сообщения, так и его получатель заранее знают параметры того кода, который будет использован в процессе шифрования. К таким параметрам относятся:

– k – длина исходного информационного вектора;

– n – длина кодового слова;

– t – количество ошибок, которые способен исправить данный код.

В свою очередь при выборе данных параметров стоит учитывать ограничения, которые накладываются на них для определения кода Гоппа, а именно зная степень t порождающего многочлена получим следующие зависимости для параметров n и k, определённые согласно формулам [6]:

$$n \leq 2^m \qquad (4)$$

$$k \geq n - mt \, (5)$$

$$d \geq 2t + 1 \; (6)$$

где d – это минимальное кодовое расстояние.

Также стоит учитывать, что выбранные параметры должны обеспечивать достаточную криптостойкость для известных типов атак на систему Мак-Элиса. Один из способов атаки на данный алгоритм был описан Анн Канто и Флоран Шабо, который заключается в поиске слов минимального веса в больших линейных кодах, что позволяет реализовать атаку по открытому ключу в частности и на систему Мак-Элиса [7].

Таким образом, параметры безопасности стоит выбирать в соответствии с алгоритмом Канто-Шабо, оценка надежности которых показана на рисунке 1.

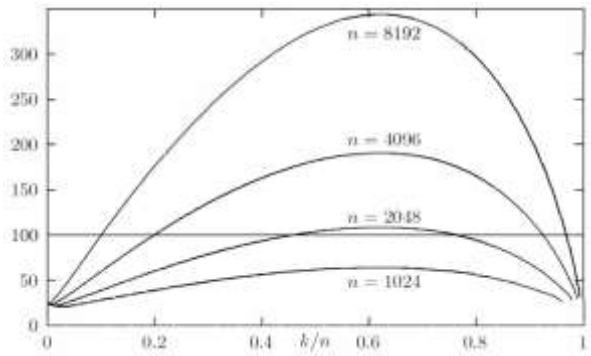

Рисунок 1 – Сложность работы алгоритма Канто-Шабо для кодов Гоппа

Как видно из данной зависимости показанной на алгоритмической шкале сложность алгоритма значительно возрастает при увеличении длины закодированных слов. Тем не менее Канто-Шабо в своей статье приводят следующие минимальные параметры для кода Гоппа чтобы компенсировать данную атаку на алгоритм Мак-Элиса: $n = 1024, k = 614, t = 41$ [7].

После того как были определены необходимые параметры, по которым будет строиться используемый в процессе шифрования код Гоппа, необходимо

определить порождающую матрицу G двоичного алгебраического кода, для которого известен эффективный алгоритм декодирования и способного исправить заданное число ошибок t. Размер данной матрицы определяется заданными параметрами n и k [6]. Для её получения необходимо воспользоваться следующим равенством:

$$GH^T = 0$$

Таким образом для получения матрицы G необходимо разложить проверочную матрицу над полем $\mathbb{F}(2^m)$ и привести к каноническому виду [5].

Данная матрица является частью закрытого ключа системы шифрования Мак-Элиса. При этом ключ также включает в себя по мимо порождающей матрицы, две матрицы, позволяющие скрыть данный код, а именно:

– двоичная подстановочная матрица P, размера n на n;
– двоичная невырожденная матрица S, размера k на k;

И тогда закрытым ключом будет являться тройка (S, G, P), при этом параметры закрытого ключа известны только абоненту, который должен принять зашифрованное сообщение и расшифровать его с помощью данных параметров.

Для получения открытого ключа необходимо получить матрицу G_p, которая необходима для шифрования очередного блока сообщения.

Для её получения необходимо воспользоваться следующей формулой [6]:

$$G_p = SGP$$

Тогда открытым ключом будет считаться пара (G_p, t), где t – это количество исправляемых кодом ошибок.

Знание параметра t для абонента, кодирующего сообщение является обязательным условием правильного выполнения процесса шифрования. Для его выполнения абонент должен представить сообщение в виде множества двоичных векторов u_i длины k.

Также необходимо выбрать случайный двоичный вектор ошибок также длиной k, при этом для каждого вектора u_i должен быть выбран свой вектор ошибок e_i [6]. При этом каждый вектор e_i должен указывать не более чем на t ошибок.

После чего каждый вектор u_i должен быть подвергнут процессу шифрования согласно формуле [6].

$$y_i = u_i G_p \oplus e_i$$

Полученные вектора y_i передаются по открытому каналу связи другому абоненту. Который в свою очередь получив их приступает к процессу расшифровывания.

Процесс расшифровывания состоит на первом этапе в нахождении для каждого зашифрованного вектора, вектора y_i', для чего необходимо воспользоваться следующей формулой [8]:

$$y_i' = y_i P^{-1}$$

где P^{-1} – это обратная матрица к секретной матрице P.

Проведя данные вычисления получим следующее выражение:

$$y_i P^{-1} = u_i SG + e'$$

где e' – это изменённый векторы ошибок.

Ввиду того что у абонента отсутствует информация как о положении ошибок в полученном векторе, так и вовсе о их наличии, необходимо применить известный получателю эффективный алгоритм декодирования кода Гоппа, что позволит получить значения вектора $u_i SG$ и в дальнейшем используя разложение $[G^T | (u_i S)^T]$ получить значение вектора $u_i S$, согласно следующим вычислениям [8]:

Обозначим очередной вектор $u_i S$ как вектор m, тогда получим:

$$u_i S = (m_1, \dots, m_k)$$

Зная, что после исправления возможных ошибок в очередном кодовом слове получаем следующее выражение:

$$(m_1, \dots, m_k) \cdot G = (c_1, \dots, c_n)$$

что будет эквивалентно

$$G^T \cdot \begin{pmatrix} m_1 \\ \vdots \\ m_k \end{pmatrix} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$$

Тогда для нахождения вектора m необходимо с помощью простых алгебраических операций провести следующее разложение [8]:

$$\left(G^T \middle| \begin{matrix} c_1 \\ \vdots \\ c_n \end{matrix}\right) \sim \cdots \sim \left(\begin{array}{c|c} I_k & \begin{matrix} m_1 \\ \vdots \\ m_k \end{matrix} \\ \hline P & \end{array}\right)$$

где I_k – это единичная матрица размера $k \times k$, а матрица P имеет размер $(n-k) \times (k+1)$ [8].

После нахождения значения вектора $u_i S$ для окончания процесса расшифровки и получения значения информационного вектор u_i необходимо воспользоваться следующей формулой [8]:

$$u_i = u_i S \cdot S^{-1}$$

где S^{-1} – это обратная матрица к матрице S.

Таким образом обобщая все выше описанный вычисления можно представить процесс генерации закрытого и открытого ключей, шифрования, передачи информации по открытому каналу и процесс расшифровывания полученных данных следующим образом как показано на рисунке 2.

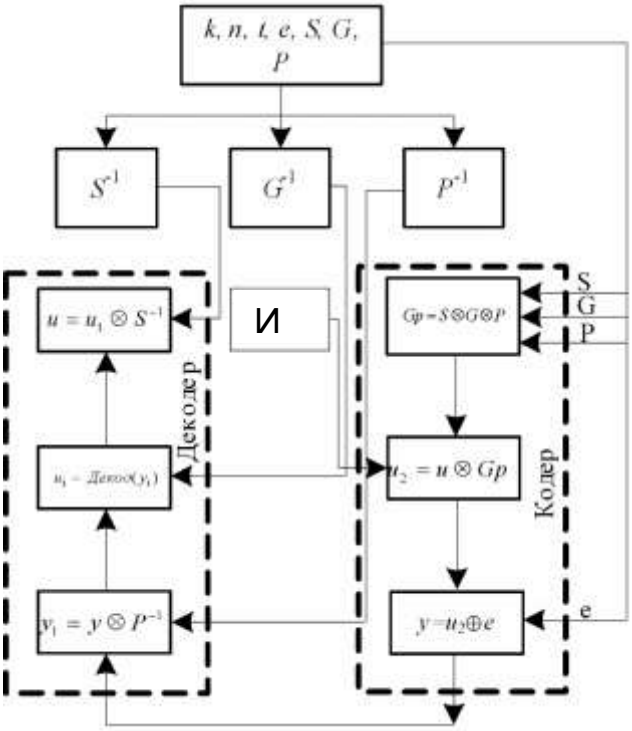

Рисунок 2 – Алгоритм шифрования Мак-Элиса

На данном рисунке показан полный цикл работы криптосистемы Мак-Элиса. Исходные данные, подставляемые для шифрования в систему обозначены как «ИД».

Стоит отметить, что в процессе описания расшифровывания информации в декодере Мак-Элиса для декодирования используемого алгебраического кода в момент шифрования информации используется как уже описывалось выше эффективный алгоритм декодирования. Реализация конкретного алгоритма зависит от того какой именно линейный код применяется в системе шифрования [9]. Его реализация зависит от конкретной реализации криптосистемы Мак-Элиса.

Для кодов Гоппа существует несколько возможных реализаций данного алгоритма, к ним относятся [9]:

– алгоритм декодирования Паттерсона;
– декодирования кода Гоппа на основе алгоритма Гоа;
– декодирования кода Гоппа на основе алгоритма Сугиямы;
– декодирования кода Гоппа на основе алгоритма Берлекэмпа-Месси.

Каждый из предложенных алгоритмов за исключением алгоритма Паттерсона подходит для декодирования кодов Гоппа над произвольным конечным полем. При этом алгоритм Паттерсона предназначен для декодирования исключительно двоичных кодов Гоппа, что обуславливается тем, что он построен на основе свойств полей вида $\mathbb{F}(2^m)$ [9].

В результате проведённого исследования криптосистемы Мак-Элиса было описано по каким алгоритмам выстраивается процесс шифрования в данной системе. В ходе работы были рассмотрены определения, которые необходимы для понимания принципов работы данной криптосистемы, а также особо подчёркнута значимость применения системы шифрования Мак-Элиса в условиях постквантовой криптографии.

Также был проведён выбор оптимальных параметров алгебраического кода, используемого в данной системе кодирования для осуществления безопасной передачи зашифрованной информации по открытому каналу связи.

Таким образом, можно сделать вывод, что развитие данной криптографической системы является необходимым этапом, который позволит преодолеть проблему применения квантовых алгоритмов для взлома существующих систем шифрования. Это позволяет считать данную систему актуальной и открывает возможности для её дальнейшего развития, с целью решения тех проблем, которые не позволяют применять её в настоящее время.

Литература:

1. Марко Бальди, Марко Бодрато, Франко Чиаралук. Новый анализ криптосистемы Мак-Элиса на основе кодов QC-LDPC, SCN 2008, LNCS 5229, стр. 246–262;
2. Бхаскар Бисвас, Николас Сендриер. Внедрение криптосистемы Мак-Элиса: теория и практика. Исследовательский центр INRIA, Париж, Франция, PQCrypto 2008, LNCS 5299, pp. 47-62;
3. Гордов Н.А. Криптосистемы Мак-Элиса и Нидеррайтера в атаках декодирования классической информации // Современные научные

исследования и инновации. 2020. № 6 [Электронный ресурс]. URL: http://web.snauka.ru/issues/ 2020/06/92527 (дата обращения: 10.01.2021);

4. Педро Бранко, Пауло Матеус, Карлос Салема, Андре Соуто. Информационные науки, том 510, 2020 г., страницы 243-255;
5. Фам Суан Нгиа. Анализ применения алгоритма Мак-Элис для электронной цифровой подписи Вестник РГРТА, Рязань, 2007 г.;
6. Гоппа В. Д. Новый класс линейных корректирующих кодов / В.Д. Гоппа // Проблемы передачи информации. - 1970. - Т. 6, вып. 3. - С. 24-30;
7. Анн Канто и Флоран Шабо. Новый алгоритм поиска слов с минимальным весом в линейном коде: приложение к криптосистеме Мак-Элиса на кодах BCH., Транзакции IEEE по теории информации, т. 44, 1998 г. с. 367-378;
8. Эллен Йохемс., Коды Гоппы и криптосистема Мак-Элиса., // Buluitreiking op woensdag 2002 г., 59 с.;
9. С. М. Рацеев, Об алгоритмах декодирования кодов Гоппы, Челяб. физ.-матем. журн., 2020, том 5, выпуск 3, с. 327–341

Literature

1. Marco Baldi, Marco Bodrato, Franco Chiaraluc. A New Analysis of the McEliece Cryptosystem Based on QC-LDPC Codes, SCN 2008, LNCS 5229, pp. 246-262, 2008;
2. Bhaskar Biswas, Nicolas Sendrier. McEliece Cryptosystem Implementation: Theory and Practice., Center de recherche INRIA Paris, France, PQCrypto 2008, LNCS 5299, pp. 47-62, 2008;
3. Gordov N.A. Cryptosystems of McEliece and Niederreiter in attacks of decoding classical information // Modern scientific research and innovations. 2020. No. 6 [Electronic resource]. URL: http://web.snauka.ru/issues/ 2020/06/92527 (date accessed: 01/10/2021);
4. Pedro Branco, Paulo Mateus, Carlos Salema, Andre Souto., Information Sciences, Volume 510, February 2020, Pages 243-255;

5. Pham Xuan Nghia. Analysis of the application of the McEliece algorithm for electronic digital signature Vestnik RGRTA. Issue 20. Ryazan, 2007;
6. Goppa VD A new class of linear correcting codes. Goppa // Problems of information transmission. - 1970 .-- T. 6, no. 3. - S. 24-30;
7. Anne Canteaut and Florent Chabaud A New Algorithm for Finding Minimum-Weight Words in a Linear Code: Application to McEliece's Cryptosystem and to Narrow-Sense BCH Codes of Length., IEEE Transactions on information theory, VOL. 44, NO. 1, January 1998 Pages 367-378;
8. Ellen Jochemsz., Goppa Codes & the McEliece Cryptosystem., // Buluitreiking op woensdag 28 augustus 2002, 59 p.;
9. SM Ratseev, On algorithms for decoding Goppa codes, Chelyab. phys.-math. zh., 2020, volume 5, issue 3, pp. 327–341