УДК 004.056.55

А. А. Красавин

Московский физико-технический институт (государственный университет)

Использование модифицированной (U|U+V)-конструкции в криптосистеме McEliece

Предложена модификация (U|U+V) конструкции для использования ее в криптосистеме McEliece, позволяющая уменьшить размер ключа без значительной потери криптостойкости системы.

Предложена модификация криптосистемы McEliece на основе предложенной кодовой конструкции.

Ключевые слова: (U|U+V) конструкция, криптосистема McEliece, криптография на кодах исправления ошибок.

A. A. Krasavin

Moscow Institute of Physics and Technology (State University)

(U|U+V) modification for McEliece cryptosystem

A modification of (U|U+V) construction for its use in the McEliece cryptosystem is proposed. The proposed algorithm allows us to reduce public key size without a significant loss of system cryptographic strength. A modification of the McEliece cryptosystem based on the proposed code structure is offered.

Key words: (U|U+V) scheme, McEliece cryptosystem, code-based cryptography.

1. Введение

Криптосистема McEliece, предложенная в 1978 году [1], является одним из возможных кандидатов для постквантовой криптографии [2].

Криптосистема McEliece может быть описана следующим образом [1].

Для генерации ключа выбирается линейный q-ичный (n, k, d)-код $\mathbb{C}$ с порождающей матрицей G, для которого известен эффективный алгоритм декодирования ошибок, весом не более t. Выбираются случайные невырожденные $k \times k$ матрица S и $n \times n$ перестановочная матрица P. В качестве открытого ключа системы публикуется матрица $\hat{G} = S \cdot G \cdot P$. Закрытым ключом системы являются алгоритм декодирования кода $\mathbb{C}$ и матрицы S, и P.

Криптосистема характеризуется следующими параметрами:

Размер открытого ключа: $PublicKeySize_0 = PKS_0 = n \cdot k$.

Скорость кода: $R_0 = \frac{k}{n}$.

Для шифрования сообщения m, представленного в виде q-ичного вектора длины k выбирается случайный q-ичный вектор e длины n веса не более t. Шифротекст рассчитывается по формуле:

$$c = m \cdot \hat{G} + e. \tag{1}$$

Для расшифрования сообщения c пользователь вычисляет $\hat{c} = c \cdot P^{-1}$, декодирует $\hat{c}$ алгоритмом декодирования для кода $\mathbb{C}$: $\hat{c} \to \{\hat{x}, e\} : \ \hat{c} = \hat{x} \cdot G + e$ и получает $m = \hat{x} \cdot S^{-1}$.

Оригинальная криптосистема McEliece была построена на кодах Гоппы [1]. Впоследствии для увеличения скорости работы или уменьшения размеров ключей криптосистемы было предложено использовать другие коды, например коды Рида–Маллера [4], обобщенные коды Рида–Соломона [3], LDPC-коды [5] и другие. Однако для многих из предложенных вариантов криптосистемы были найдены атаки на структуру матрицы $\hat{G}$, позволяющие вычислить закрытый ключ из открытого. Все эти атаки использовали различные особенности данных кодов. Так, атаки на криптосистему McEliece, построенную на кодах Рида–Маллера [6] или на полярных кодах [8], использовали поиск слов кода минимального веса.

Для защиты от подобных атак в нескольких работах [4] [7] было предложено использовать (U|U+V)-конструкцию и ее модификации. Данная конструкция позволяет создать код длины n из двух кодов U и V длины $\frac{n}{2}$ каждый:

$$(U|U+V) = \{(u|u+v) : \forall u \in U, \forall v \in V\}.$$

При использовании такого кода в криптосистеме McEliece сложность структурных атак на открытый ключ криптосистемы будет основана на задаче выделения порождающих матриц G_U и G_V кодов U и V из публичного ключа – матрицы $\hat{G} = S \cdot \begin{pmatrix} G_U & G_U \\ O & G_V \end{pmatrix} \cdot P$. Данная задача считается вычислительно сложной, а при использовании случайных кодов является NP-сложной [9]. В общем же случае ее принадлежность к классам P или NP не была доказана.

Недостатком использования (U|U+V)-конструкции в криптосистеме McEliece является увеличение размеров ключей системы по сравнению с размерами ключей оригинальной системы McEliece той же криптостойкости. Так, при использовании (U|U+V) конструкции получатся следующие параметры криптосистемы:

Размер открытого ключа: $PKS = 4 \cdot n \cdot k = 4 \cdot PKS_0$.

Скорость кода: $R = \frac{2k}{2n} = R_0$.

2. Модификация (U|U+V) конструкции

Для уменьшения размера порождающей матрицы (U|U+V) кода можно рассмотреть ряд модификаций.

Пусть $\mathbb{C}$ – (n, k, d) q-ичный код с порождающей матрицей G и эффективным алгоритмом декодирования ошибок веса менее t. Пусть K – случайная матрица размером $L \times n$. Можно задать K так, чтобы код $\hat{\mathbb{C}}$, заданный порождающей матрицей $\hat{G} = \begin{pmatrix} G \\ K \end{pmatrix}$ был $(n, k+L, d_0 \leq d)$-кодом.

Если B – порождающая матрица случайного (m, L, d_B) кода, то код $\mathbb{D}$, заданный порождающей матрицей

$$J = \begin{pmatrix} G & 0 \\ K & B \end{pmatrix}, \tag{2}$$

будет иметь минимальное расстояние:

$$d_J \geq \min(d, d_0 + d_B). \tag{3}$$

Чтобы $d_J \geq d$, необходимо потребовать, чтобы было выполнено условие $d_0 + d_B \geq d$, то есть чтобы выполнялось следующее условие:

$$d \geq d_B \geq d - d_0 \geq \frac{d}{2}.$$

Если $\mathbb{B}$ – код с максимальным кодовым расстоянием, например, код Рида–Соломона, то $m - L + 1 = d_B$. В этом случае $m = L + d_B - 1$, и потому

$$L + d - 1 \geq m \geq L + \frac{d}{2} - 1.$$

Пусть $\mathbb{B}$ – $(L+d-1, L, d_b \geq d)$ код. Код $\mathbb{D}$ с порождающей матрицей J, заданной согласно (2), может исправить не менее t ошибок по следующему алгоритму декодирования:

Пусть принято сообщение $c = xJ + e,\ wt(e) \leq t$.

1) c представляется в виде: $c = (z_1|z_2)$, длина z_1 равна n;

2) Перебираются все вектора e_i – q-ичные последовательности длины L;

3) Для каждого e_i вычисляется $v_i = z_1 - e_i K$;

4) v_i декодируется по алгоритму для кода $\mathbb{C}$: $v_i = u_i G + \xi_i;\ wt(\xi_i) \leq t$;

5) Если вес вектора $\xi = (\xi_i|z_2 - e_i\ B)$ меньше t, то $x = (u_i|e_i),\ e = \xi$, иначе повторяются пп. 2–5.

С учетом (3) и параметров выбранного кода $\mathbb{B}$, расстояние кода $\mathbb{D}$ больше d и поэтому в результате работы алгоритма декодирования всегда найдется единственное решение.

Предложенный алгоритм декодирования по сложности в q^L раз хуже алгоритма декодирования кода $\mathbb{C}$.

Способ построения кода $\mathbb{D}$ из $\mathbb{C}$ назван *КА-схемой*.

3. Использование модифицированной (U|U+V) конструкции в криптосистеме McEliece

Пусть A – случайная матрица размром $k \times L$.

Можно заметить, что подкод $\mathbb{D}_1$ кода $\mathbb{D}$, заданный порождающей матрицей $J_1 = \begin{pmatrix} G + AK & AB \end{pmatrix}$, имеет расстояние не меньше d. При этом код может исправить не менее t ошибок по следующему алгоритму декодирования:

Пусть принято сообщение $c = xJ_1 + e,\ wt(e) \leq t$.

1) с представляется в виде: $c = (z_1|z_2)$, длина z_1 равна n.

2) Перебираются все вектора e_i – q-ичные последовательности длины L.

3) Для каждого e_i вычисляется $v_i = z_1 - e_i K$.

4) v_i декодируется по алгоритму для кода C: $v_i = u_i G + \xi_i,\ wt(\xi_i) \leq t$.

5) Если вес вектора $\xi = (\xi_i|z_2 - u_i AB)$ меньше t, то $x = u_1$,$e = \xi$, иначе повторяются пп. 2–5.

При использовании в криптосистеме McEliece кода $\mathbb{D}_1$, построенного при помощи предложенной КА-схемы, открытым ключом будет являться матрица

$$\hat{J} = S \cdot J_1 \cdot P = S \cdot \begin{pmatrix} G + AK & AB \end{pmatrix} \cdot P. \tag{4}$$

При этом для осуществления структурных атак на систему необходимо отделить столбцы матриц $G + AK$ от AB и слова кода, заданные матрицей G от слов, заданных матрицей AK.

Без знания P вероятность случайного выбора одного столбца матрицы AB – $\frac{1}{\binom{n+L+d-1}{L+d-1}}$. Значит, при случайном выборе столбцов в среднем потребуется полиномиальное (от n) число попыток для нахождения одного столбца матрицы AB.

Решение второй задачи – отделения слов кода, заданных матрицей G от слов, заданных матрицей AK, – сводится к выделению ядра отображения φ_A, заданного матрицей A, так как:

$$\forall a \in Ker_{\varphi_A} :\ a \cdot (G + AK) = a \cdot G \in \mathbb{C}. \tag{5}$$

При случайном выборе матриц A, K и B, это возможно сделать только перебором всех векторов пространства F_{q^k}, что требует порядка $\binom{q^k}{q^L}$ операций и, потому, является вычислительно сложной задачей.

Значит, структурные атаки на открытый ключ системы будут вычислительно сложными, при этом схема будет обладать следующими параметрами:

$$\begin{gathered} PKS = (n+L+d-1)\cdot k = n\cdot k\cdot(1+\frac{d}{n}+\frac{L-1}{n}) = PKS_0\cdot(2-R_0+\frac{L-1}{n}); \\ R = \frac{k}{n+L+d-1} = R_0\cdot\frac{1}{2-R_0+\frac{L-1}{n}}. \end{gathered} \tag{6}$$

4. Модификация криптосистемы McEliece

С учетом предложенных идей можно предложить и следующую модификацию криптосистемы McEliece:

Пусть $\mathbb{C}$ – (n,k,d) q-ичный код с порождающей матрицей G и эффективным алгоритмом декодирования ошибок веса менее t. Для генерации ключей выбираются случайные матрицы A и K размером $k\times L$ и $L\times n$ соответственно, случайные невырожденные матрицы S и P, размером $k\times k$ и $n\times n$ соответственно. Пусть $h(\cdot)$ – криптографически стойкая хэш-функция.

В качестве открытого ключа публикуется матрица $\hat{G} = S(G+AK)P$ и хэш-функция $h(\cdot)$.

Для шифрования сообщения m, представленного в виде q-ичного вектора длины k выбирается случайный q-ичный вектор e длины n веса не более t. Шифротекст рассчитывается по формуле:

$$c = (m\mathbb{G} + e \| h(m\|e)).$$

Для расшифрования сообщения c:

1) Вычисляется $\hat{c} = cP^{-1} = (c_1|c_2)$, длина c_1 равна n.

2) Перебираются все вектора e_i – q-ичные последовательности длины L.

3) Для каждого e_i вычисляется $d_i = c_1 - e_iK$.

4) d_i декодируется по алгоритму для кода $\mathbb{C}$: $d_i = x_iG + \xi_i,\ wt(\xi_i) \le t$.

5) Если $c_2 \equiv h(x_i, \xi_i)$, то вычисляется $m = x_iS^{-1}$, иначе повторяются пп. 2–5.

5. Заключение

Предложена модификация (U|U+V) конструкции, позволяющая использовать любые коды в криптосистеме McEliece без существенной потери производительности и увеличения размеров ключей криптосистемы по сравнению с криптосистемами McEliece на тех же кодах, не использующих модификаций. При этом увеличивается стойкость криптосистемы к структурным атакам на открытый ключ.

По сравнению с криптосистемой McEliece, использующей (U|U+V) конструкцию, предложенная модификация позволяет существенно уменьшить размеры ключей криптосистемы.

Литература

1. *McEliece R.J.* A public-key cryptosystem based on algebraic coding theory // DSN Progress Report, Jet Propulsion Laboratory, Pasadena. 1978. P. 114–116.

2. *Dinh H., Moore C., Russell A.* McEliece and Niederreiter Cryptosystems That Resist Quantum Fourier Sampling Attacks // Advances in Cryptology — CRYPTO 2011: 31st Annual Cryptology Conference, Santa Barbara, CA, USA, August 14–18, 2011, Proceedings. 2011. P. 761–779.

3. *Niederreiter H.* Knapsack-Type Cryptosystems and Algebraic Coding Theory // Problems of Control and Information Theory. 1986. V. 15, N 2. P. 159–166.

4. *Sidelnikov V.M.* A public-key cryptosytem based on ReedMuller codes // Discrete Math. Appl. 1994. V. 4, N 3. P. 191–207.

5. *Baldi M., Chiaraluce F.* Cryptanalysis of a new instance of McEliece cryptosystem based on QC-LDPC codes // Proc. IEEE Int. Symposium Inf. Theory – ISIT. 2007. P. 2591–2595.

6. *Minder L., Shokrollahi A.* Cryptanalysis of the Sidelnikov cryptosystem // Advances in Cryptology – EUROCRYPT 2007. 2007. V. 4515 of Lecture Notes in Comput. Sci. P. 347–360.

7. *Corbella I.M., Tillich J.-P.* Attaining capacity with iterated (U|U + V) codes based on AG codes and Koetter-Vardy soft decoding // ISIT 2017. 2017. P. 6–10.

8. *Bardet M., Chaulet J., Dragoi V., Otmani A., Tillich J.-P.* Cryptanalysis of the McEliece Public Key Cryptosystem Based on Polar Codes // Post-Quantum Cryptography: 7th International Workshop, PQCrypto 2016, Fukuoka, Japan. 2016. P. 118–143.

9. *Debris-Alazard T., Sendrier N., Tillich J.-P.* SURF: A new code-based signature scheme arXiv preprint arXiv:1706.08065 2017.

References

1. *McEliece R.J.* A public-key cryptosystem based on algebraic coding theory. DSN Progress Report, Jet Propulsion Laboratory, Pasadena. 1978. P. 114–116.

2. *Dinh H., Moore C., Russell A.* McEliece and Niederreiter Cryptosystems That Resist Quantum Fourier Sampling Attacks. Advances in Cryptology — CRYPTO 2011: 31st Annual Cryptology Conference, Santa Barbara, CA, USA, August 14–18, 2011, Proceedings. 2011. P. 761–779.

3. *Niederreiter H.* Knapsack-Type Cryptosystems and Algebraic Coding Theory. Problems of Control and Information Theory. 1986. V. 15, N 2. P. 159–166.

4. *Sidelnikov V.M.* A public-key cryptosytem based on ReedMuller codes. Discrete Math. Appl. 1994. V. 4, N 3. P. 191–207.

5. *Baldi M., Chiaraluce F.* Cryptanalysis of a new instance of McEliece cryptosystem based on QC-LDPC codes. Proc. IEEE Int. Symposium Inf. Theory – ISIT. 2007. P. 2591–2595.

6. *Minder L., Shokrollahi A.* Cryptanalysis of the Sidelnikov cryptosystem. Advances in Cryptology – EUROCRYPT 2007. 2007. V. 4515 of Lecture Notes in Comput. Sci. P. 347–360.

7. *Corbella I.M., Tillich J.-P.* Attaining capacity with iterated (U|U + V) codes based on AG codes and Koetter-Vardy soft decoding. ISIT 2017. 2017. P. 6–10.

8. *Bardet M., Chaulet J., Dragoi V., Otmani A., Tillich J.-P.* Cryptanalysis of the McEliece Public Key Cryptosystem Based on Polar Codes. Post-Quantum Cryptography: 7th International Workshop, PQCrypto 2016, Fukuoka, Japan. 2016. P. 118–143.

9. *Debris-Alazard T., Sendrier N., Tillich J.-P.* SURF: A new code-based signature scheme. arXiv preprint arXiv:1706.08065 2017.

Поступила в редакцию 01.03.2018