УДК 004.056.55

НЕСИММЕТРИЧНЫЕ КРИПТОСИСТЕМЫ НА ОСНОВЕ АЛГЕБРАИЧЕСКОГО КОДИРОВАНИЯ: СОВРЕМЕННОЕ СОСТОЯНИЕ, СУЩЕСТВУЮЩИЕ ПРОТИВОРЕЧИЯ И ПЕРСПЕКТИВЫ ПРАКТИЧЕСКОГО ИСПОЛЬЗОВАНИЯ НА ПОСТ-КВАНТОВЫЙ ПЕРИОД

Александр Кузнецов[1], Андрей Пушкарев[2], Сергей Кавун[3], Вячеслав Калашников[4]

[1] Харьковский национальный университет имени В.Н. Каразина, площадь Свободы, 4, г. Харьков, 61022, Украина
kuznetsov@karazin.ua

[2] Департамент Государственной службы специальной связи и защиты информации Украины, г. Киев, Украина

[3] Харьковский учебно-научный институт ГВУЗ "Университет банковского дела", пр. Победы 55, г. Харьков, 61174, Украина
kavserg@gmail.com

[4] Технологический университет Монтеррея, Монтеррей, Мексика
пр. Еухеніо Гарса Сада 2501, 64849 Монтеррей, Нуево-Леон, Мексика
kalash@itesm.mx

Рецензент: Роман Олійников, д.т.н., проф., Харьковский национальный университет имени В.Н. Каразина, площадь Свободы, 4, г. Харьков, 61022, Украина
roliynykov@gmail.com

Поступила в августе 2016

Аннотация. *Рассматриваются несимметричные криптосистемы на основе алгебраического кодирования, исследуются современное состояние, существующие противоречия и перспективы практического использования на пост-квантовый период. Предлагается новая схема криптопреобразования, существенно повышающая относительную скорость передачи информации. Показано преимущество предлагаемой конструкции по сравнению с уже известными схемами Мак-Элиса и Нидеррайтера.*

Ключевые слова: *несимметричные криптосистемы, пост-квантовая криптография, криптография на основе кодов.*

1 Введение

Для предоставления базовых услуг безопасности в современных информационно-телекоммуникационных системах применяются различные криптографические механизмы, в частности, несимметричные (двухключевые) криптосистемы, в которых задача поиска секретного ключа по известному открытому ключу связана с решением известной и очень сложной математической задачи (факторизации, дискретного логарифмирования и пр.) [1-3]. Однако с появлением квантовых вычислений, основанных на принципах квантовой механики, в частности, на принципе суперпозиции и явлении квантовой запутанности, скорость решения определенных математических задач значительно возрастает [4]. Существует ряд квантовых алгоритмов, например, алгоритмы Дойча и Йожи, Саймона, Гровера, Шора и другие, выполнение которых занимает гораздо меньше времени, чем выполнение любого вероятностного классического алгоритма [5-11]. Алгоритм Шора позволяет найти за конечное (и приемлемое) время все простые множители больших чисел или решить задачу дискретного логарифмирования, и, как следствие, найти секретный ключ соответствующего асимметричного криптоалгоритма (RSA, ECC, или других) [10]. Следовательно, разработка и теоретическое обоснование новых криптографических алгоритмов, в которых сложность поиска секретного параметра по известному открытому ключу остается высокой даже с учетом возможного применением квантовых вычислений (т.е. для пост-квантового периода), является чрезвычайно важной научной задачей [12-14].

Среди возможных кандидатов для пост-квантовой криптографии *(Post-Quantum Cryptography)* особое место занимают алгоритмы, построение которых основано на использовании алгебраических кодов, замаскированных под код общего положения (случайный код, полный код) [15-18]. В русскоязычной литературе подобные алгоритмы получили название теоретико-кодовых схем [19,20], или крипто-кодовых преобразований [21,22]. Наряду с высокой скоростью криптографического преобразования и возможностью совмещать контроль ошибок с защитой от несанкционированного ознакомления [23] крипто-кодовые преобразования остаются стойкими даже в случае использования квантовых вычислений. Кроме того, на сегодняшний день уже известны различные криптографические примитивы (алгоритмы несимметричного [15,16,18,20] и симметричного [17] шифрования, генераторы псевдослучайных последовательностей и поточного шифрования [24 - 26], протоколы доказательства с нулевым разглашением *(Zero-knowledge proof)* [27,28], схемы электронной цифровой подписи [29,30], идентификации [31,32] и пр.), основанные на использовании алгебраических кодов, что делает это направление универсальным инструментом, позволяющим на едином математическом и программном обеспечении реализовать широкий спектр эффективных механизмов криптографической защиты информации. И хотя известны также и вычислительно эффективные атаки на отдельные варианты теоретико-кодовых схем [19, 33 - 36], базовая конструкция [15], предложенная около 40 лет назад, остается стойкой ко всем известным методам криптоанализа, что с исторической ретроспективы подтверждает надежность и перспективность крипто-кодовых преобразований, особенно в контексте построения эффективных пост-квантовых алгоритмов криптографической защиты [37].

В данной статье рассматриваются общетеоретические положения алгебраического кодирования и несимметричные криптосистемы на их основе. Исследуются современное состояние, существующие противоречия и перспективы практического использования несимметричных кодовых криптосистем на пост-квантовый период. Предлагается новая схема криптопреобразования, существенно повышающая относительную скорость передачи информации. Показано преимущество предлагаемой конструкции по сравнению с уже известными схемами Мак-Элиса и Нидеррайтера [15,16].

2 Общие положения алгебраической теории блоковых кодов, используемые для описания теоретико-кодовых схем

Введем основные термины и определения алгебраической теории кодирования [38-40], используемые в дальнейшем при рассмотрении несимметричных кодовых криптосистем (теоретико-кодовых схем), в том числе алгоритмов формирования и проверки ЭЦП.

Зафиксируем конечное поле $GF(q)$ и рассмотрим векторное пространство V_n, как множество n-последовательностей с элементами из $GF(q)$ с покомпонентным сложением и умножением на скаляр.

Линейный (n,k,d) *код* V_k над $GF(q)$ есть подпространство в V_n, т.е. непустое множество n-последовательностей (*кодовых слов*) с элементами из $GF(q)$, где k – *размерность* линейного подпространства, d - минимальный вес Хемминга (число ненулевых элементов) $w_h(c)$ произвольного ненулевого кодового слова c кода V_k:

$$d = \min_{\forall c \in V_k, c \neq 0} w_h(c).$$

В виду линейности подпространства V_k набор весов различных ненулевых кодовых слов совпадает с набором расстояний по Хеммингу между различными кодовыми словами, т.е. d называют также *минимальным кодовым расстоянием* по Хеммингу кода V_k. Величину $R = \frac{k}{n}$ называют *относительной скоростью кода*, а $\delta = \frac{d}{n}$ называют *относительным минимальным кодовым расстоянием*.

Основная проблема теории избыточного кодирования впервые сформулирована в работе К. Шеннона [41]: найти коды с большой относительной скоростью R и с большим минимальным кодовым расстоянием d. Она вытекает из следующей теоремы.

Теорема 1 [41]. Пусть $C(P_0)$ – пропускная способность дискретного симметричного канала с вероятностью ошибки P_0. Тогда для любого $\varepsilon > 0$, если $R < C(P_0)$ и n достаточно велико, существует (n, k, d) код с относительной скоростью $k / n \geq R$, вероятность ошибки которого $P_{ош} < \varepsilon$.

Теорема 1 доказана вероятностными методами и не дает механизма для построения кодов с высокими R и δ. Для линейных блоковых кодов справедлива следующая оценка (граница Варшамова-Гилберта).

Теорема 2 [40]. Если выполняется равенство

$$q^{n-k} \geq \sum_{i=0}^{d-2} C_{n-1}^{i} (q-1)^{i}, \tag{1}$$

тогда существует линейный (n, k, d) код над $GF(q)$.

На практике чаще используют асимптотические границы, которые дают представление о предельных кодовых характеристиках при бесконечно большой длине кода. Прологарифмируем выражение (1), получим

$$n - k \geq \log_q \left(\sum_{i=0}^{d-2} C_{n-1}^{i} (q-1)^{i} \right).$$

Устремим $n \to \infty$, получим асимптотическую границу Варшамова-Гилберта:

$$R \leq 1 - H_q(\delta), \tag{2}$$

где $H_q(x)$ - q - ичная функция энтропии на отрезке $\left[0, \frac{q-1}{q} \right]$, причем

$$H_q(x) = x \log_q (q-1) - x \log_q (x) - (1-x) \log_q (1-x), \; 0 < x \leq \frac{q-1}{q}.$$

Таким образом, проблема помехоустойчивого кодирования состоит в поиске регулярных алгоритмов построения таких линейных блоковых (n, k, d) кодов, параметры которых удовлетворяют кодовой границе (1) и/или асимптотические кодовые границы которых удовлетворяют (2).

Линейный код V_k (как линейное подпространство в V_n) задается набором базисных (линейно независимых) векторов

$$(g_{0,0}, g_{0,1}, ..., g_{0,n-1}),$$
$$(g_{1,0}, g_{1,1}, ..., g_{1,n-1}),$$
$$\dots$$
$$(g_{k-1,0}, g_{k-1,1}, ..., g_{k-1,n-1}),$$

которые обычно представляются в матричном виде через порождающую матрицу

$$G = \begin{pmatrix} g_{0,0} & g_{0,1} & \dots & g_{0,n-1} \\ g_{1,0} & g_{1,1} & \dots & g_{1,n-1} \\ \dots & \dots & \dots & \dots \\ g_{k-1,0} & g_{k-1,1} & \dots & g_{k-1,n-1} \end{pmatrix}$$

ранга $rank(G) = k$ и размерности $k \times n$.

Произвольное кодовое слово $c = (c_0, c_1, ..., c_{n-1})$ кода V_k есть линейная комбинация строк из матрицы G. Кодирование заключается в сопоставлении каждого *информационного слова* $i = (i_0, i_1, ..., i_{k-1})$ с символами из $GF(q)$ некоторому кодовому слову $(c_0, c_1, ..., c_{n-1})$. Наиболее простой способ кодирования задается выражением

$$c = iG. \tag{3}$$

Посредством линейных операций над строками матрицу G удобно привести к каноническому виду

$$G^* = \begin{pmatrix} 1 & 0 & ... & 0 & g^*_{0,k} & g^*_{0,k+1} & ... & g^*_{0,n-1} \\ 0 & 1 & ... & 0 & g^*_{1,k} & g^*_{1,k+1} & ... & g^*_{1,n-1} \\ ... & ... & ... & ... & ... & ... & ... & ... \\ 0 & 0 & ... & 1 & g^*_{k-1,k} & g^*_{k-1,k+1} & ... & g^*_{k-1,n-1} \end{pmatrix} = I \| P, \tag{4}$$

где I - единичная подматрица размером $k \times k$, P - подматрица размером $k \times (n-k)$ в правой части матрицы G^*, $\|$ - символ конкатенации (объединения).

Тогда при использовании выражения $c = iG$ имеем *систематическое* правило кодирования $c = i \| iP$, т.е. информационный вектор $i = (i_0, i_1, ..., i_{k-1})$ будет в явном виде содержаться в кодовом слове $c = (c_0, c_1, ..., c_{n-1})$.

Линейное подпространство, отождествляющее код V_k, имеет ортогональное дополнение (обозначим его U_{n-k}). Базис подпространства U_{n-k} задается векторами

$$(h_{0,0}, h_{0,1}, ..., h_{0,n-1}),$$
$$(h_{1,0}, h_{1,1}, ..., h_{1,n-1}),$$
$$...$$
$$(h_{n-k-1,0}, h_{n-k-1,1}, ..., h_{n-k-1,n-1})$$

и обычно представляется в матричном виде через проверочную матрицу

$$H = \begin{pmatrix} h_{0,0} & h_{0,1} & ... & h_{0,n-1} \\ h_{1,0} & h_{1,1} & ... & h_{1,n-1} \\ ... & ... & ... & ... \\ h_{n-k-1,0} & h_{n-k-1,1} & ... & h_{n-k-1,n-1} \end{pmatrix}$$

ранга $rank(H) = n-k$ и размерности $(n-k) \times n$.

Условие ортогональности векторов из V_k и U_{n-k} в матричном виде записывается как

$$GH^T = 0, \tag{5}$$

где под нулем понимается нулевая матрица, размерности $k \times (n-k)$.

Линейное подпространство U_{n-k} называют *дуальным* (двойственным) к V_k кодом над $GF(q)$. *Определение кода* V_k *через ортогональное дополнение* U_{n-k} (через дуальный код) можно сформулировать следующим образом: произвольная n-последовательностей $c = (c_0, c_1, ..., c_{n-1})$ с элементами из $GF(q)$ является кодовым словом кода V_k тогда и только тогда, когда она ортогональна каждой строке проверочной матрицы H, т.е. при $cH^T = 0$.

Посредством линейных операций над строками приведем матрицу H к каноническому виду

$$H^* = \begin{pmatrix} h^*_{0,0} & h^*_{0,1} & ... & h^*_{0,k-1} & 1 & 0 & ... & 0 \\ h^*_{1,0} & h^*_{1,1} & ... & h^*_{1,k-1} & 0 & 1 & ... & g^*_{1,n-1} \\ ... & ... & ... & ... & ... & ... & ... & ... \\ h^*_{n-k-1,0} & h^*_{n-k-1,1} & ... & h^*_{n-k-1,k-1} & 0 & 0 & ... & 1 \end{pmatrix} = P^* \| I, \tag{6}$$

где P^*- подматрица размером $(n-k) \times k$ в левой части матрицы H^*, I - единичная подматрица размером $(n-k) \times (n-k)$, $\|$ - символ конкатенации (*объединения*).

Тогда из условия $G*H*^T = 0$ имеем $P* = -P^T$ (с операциями над $GF(q)$).

Единичные вектора-столбцы в матрицах $G*$ и $H*$ могут быть выбраны произвольно с соответствующим формированием единичных подматриц и систематическим размещением информационных символов в кодовом слове.

Основной целью избыточного кодирования информации является контроль *(обнаружение и исправление)* ошибок, произошедших при передаче сообщения по каналу с шумами [38-40]. Для контроля ошибок кодирующее устройство вносит избыточность *(проверочную часть длины* $r = n - k$*)* в передаваемые данные. На приемной стороне, анализируя свойства проверочной части и ее соответствие передаваемым данным, декодер уменьшает влияние ошибок, возникших при передаче.

Обозначим вектор ошибок, воздействующий на передаваемое кодовое слово c, как n-последовательность $e = (e_0, e_1, ..., e_{n-1})$ с элементами из $GF(q)$. Искаженное кодовое слово обозначим вектором $c* = c + e = (c_0 + e_0, c_1 + e_1, ..., c_{n-1} + e_{n-1})$.

Синдромом в теории кодирования называют вектор $s = (s_0, s_1, ..., s_{n-k-1})$ с элементами из $GF(q)$, который характеризует воздействие вектора ошибок на произвольное кодовое слово:

$$s = c*H^T = cH^T + eH^T = eH^T, \quad (7)$$

т.е. значение вектора s зависит только от вектора ошибок $e = (e_0, e_1, ..., e_{n-1})$ и не зависит от выбранного кодового слова $c = (c_0, c_1, ..., c_{n-1})$.

Таким образом, процесс декодирования состоит в анализе синдрома: при $s = 0$ принимается решение об отсутствии ошибок; при $s \neq 0$ принимается решение об искажении кодового слова ненулевым вектором ошибок. Дальнейшие действия зависят от принятой стратегии: в системах обнаружения ошибок с переспросом осуществляется запрос на повторную передачу кодового слова; в системах с прямым исправлением ошибок осуществляется поиск вектора $e = (e_0, e_1, ..., e_{n-1})$ по вычисленному значению $s \neq 0$.

Следует отметить, что при больших n и k задача поиска вектора e по ненулевому синдрому s для случайно выбранного в пространстве V_n линейного кода V_k является чрезвычайно сложной математической задачей. В общем случае эта задача относится к классу NP-сложных [37]. Однако для алгебраических кодов, со специфической структурой матриц G и H, декодирование (задача поиска вектора ошибок e и/или восстановление безошибочного кодового слова c) является полиномиально разрешимой задачей.

Алгебраическое кодирование основано на использовании специальных алгебраических уравнений, позволяющих однозначно представить информационные и кодовые слова, вектора ошибок и синдромов и свести задачу декодирования к решению систем линейных уравнений. Действительно, каждый вектор из V_n можно представить многочленом от формальной переменной x степени не выше $n-1$. При этом элементы вектора отождествляются с коэффициентами многочлена, а множество многочленов имеет структуру векторного пространства, идентичную структуре пространства V_n, а так же структуру кольца многочленов по модулю двучлена $x^n - 1$. Рассмотрим следующие многочлены:

- информационный многочлен $i(x) = i_0 x^0 + i_1 x^1 + ... + i_{k-1} x^{k-1}$ (соответствует вектору i);
- кодовый многочлен $c(x) = c_0 x^0 + c_1 x^1 + ... + c_{n-1} x^{n-1}$ (соответствует вектору c);
- многочлен ошибок $e(x) = e_0 x^0 + e_1 x^1 + ... + e_{n-1} x^{n-1}$ (соответствует вектору e);
- кодовый многочлен с ошибками $c*(x) = c*_0 x^0 + c*_1 x^1 + ... + c*_{n-1} x^{n-1}$ (соответствует вектору $c*$);
- синдромный многочлен $s(x) = s_0 x^0 + s_1 x^1 + ... + s_{n-k-1} x^{n-k-1}$ (соответствует вектору s).

Зададим, например, с помощью корней $X_0, X_1, ..., X_{n-k-1} \in GF(q^m)$, приведенный ненулевой многочлен $g(x) = (x - X_0)(x - X_1)...(x - X_{n-k-1})$ степени $r = n - k$ и правило кодирования

$$c(x) = i(x)g(x), \tag{8}$$

которое является полиномиальным аналогом выражения (3).

Многочлен $g(x)$ по аналогии с матрицей G называют *порождающим*, а соответствующая ему матрица G может быть получена циклически сдвинутой построчной записью коэффициентов многочлена $g(x)$ [38-40]. Линейные блоковые коды заданные таким образом называют *циклическими*, т.к. из принадлежности пространству V_k некоторой последовательности c (и соответствующего многочлена $c(x)$) следует также и принадлежность любой циклически сдвинутой последовательности (что в терминах многочленов трактуется как многочлен $x^i c(x)$, $i \in 0,1,...,n-k-1$ с операциями в кольце по модулю двучлена $x^n - 1$).

Многочлен $g(x)$ в общем случае определен над $GF(q^m)$ и тогда кодовые многочлены $c(x)$ также будут определены над расширенным полем $GF(q^m)$. Однако в случае, если все корни $X_0, X_1, ..., X_{n-k-1} \in GF(q^m)$ являются также всеми корнями некоторого набора минимальных многочленов элементов $GF(q^m)$, тогда порождающий многочлен $g(x)$ всегда будет иметь коэффициенты из подполя $GF(q)$, причем:

$$g(x) = H.O.K.\left(\prod_i f_i(x)\right),$$

где i пробегает по всем классам сопряженных элементов поля $GF(q^m)$, $f_i(x)$ - минимальный многочлен элемента $\alpha^i \in GF(q^m)$, α - примитивный элемент, $H.O.K.$ - наименьшее общее кратное.

Значение многочлена в его корне равно нулю, т.е. для всех $X_j \in \{X_0, X_1, ..., X_{n-k-1}\}$ выполняется равенство

$$c(X_j) = c_0 X_j^0 + c_1 X_j^1 + ... + c_{n-1} X_j^{n-1},$$

что в матричном виде соответствует записи:

$$(c_0, c_1, c_2, ..., c_{N-1}) \cdot \begin{vmatrix} 1 & X_0 & X_0^2 & ... & X_0^{n-1} \\ 1 & X_1 & X_1^2 & ... & X_1^{n-1} \\ 1 & X_2 & X_2^2 & ... & X_2^{n-1} \\ ... & ... & ... & ... & ... \\ 1 & X_{r-1} & X_{r-1}^2 & ... & X_{r-1}^{n-1} \end{vmatrix}^T = 0.$$

Полученное выражение соответствует условию взаимной ортогональности произвольного кодового слова $c = (c_0, c_1, c_2, ..., c_{n-1})$ и матрицы в правой части произведения. Следовательно, положим

$$cH^T = 0,\ H = \begin{pmatrix} X_0^0 & X_0^1 & ... & X_0^{n-1} \\ X_1^0 & X_1^1 & ... & X_1^{n-1} \\ ... & ... & ... & ... \\ X_{n-k-1}^0 & X_{n-k-1}^1 & ... & X_{n-k-1}^{n-1} \end{pmatrix}, \tag{9}$$

где H – проверочная матрица кода, заданная корнями порождающего многочлена.

Для построения матрицы H с элементами из подполя $GF(q)$ следует заменить каждый элемент $X_j^i \in GF(q^m)$ в (9) вектором-столбцом из m элементов поля $GF(q)$.

Если выбрать в качестве корней многочлена $g(x)$ $2t$ подряд следующих элементов $X_0 = \alpha^j, X_1 = \alpha^{j+1}, ..., X_{n-k-1} = \alpha^{j+2t-1} \in GF(q^m)$, тогда, по теореме Боуза-Чоудхури-Хоквингема [38-40], полученный код (БЧХ) будет иметь минимальное кодовое расстояние равное $d = 2t+1$. Для кода БЧХ над $GF(q^m)$ проверочная матрица примет вид

$$H = \begin{pmatrix} \alpha^0 & \alpha^j & \dots & \alpha^{j(n-1)} \\ \alpha^0 & \alpha^{j+1} & \dots & \alpha^{(j+1)(n-1)} \\ \dots & \dots & \dots & \dots \\ \alpha^0 & \alpha^{j+2t-1} & \dots & \alpha^{(j+2t-1)(n-1)} \end{pmatrix}. \quad (10)$$

Заданные таким образом коды называют *кодами Рида-Соломона*, их (n,k,d) параметры над $GF(q^m)$ связаны соотношением $d = n-k+1$ (верхняя граница Синлгтона), т.е. они обладают *максимально-достижимым кодовым расстоянием* (МДР) [38-40]. Ограничением на подполе $GF(q)$ с заменой всех $\alpha^i \in GF(q^m)$ в (10) соответствующими векторами-столбцами из m элементов поля $GF(q)$ получают коды над $GF(q)$, (n,k,d) параметры которых удовлетворяют ограничению (нижняя граница БЧХ) $d \ge n-km+1$.

Предположим, что кодовое слово c исказилось при его передаче, а число ошибок на блоке из N символов не превышает исправляющей способности $t = \left\lfloor \frac{d-1}{2} \right\rfloor$ алгебраического (n,k,d) кода. Другими словами, многочлен $e(x)$ содержит не более t ненулевых коэффициентов. Из (7) и (9) следует равенство

$$(s_0, s_1, \dots, s_{n-k-1}) = (e_0, e_1, e_2, \dots, e_{N-1}) \cdot \begin{vmatrix} 1 & X_0 & X_0^2 & \dots & X_0^{N-1} \\ 1 & X_1 & X_1^2 & \dots & X_1^{N-1} \\ 1 & X_2 & X_2^2 & \dots & X_2^{N-1} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & X_{n-k-1} & X_{n-k-1}^2 & \dots & X_{n-k-1}^{N-1} \end{vmatrix}^T,$$

что эквивалентно следующей системе уравнений:

$$\begin{aligned} s_0 &= e_0 + e_1 X_0 + e_2 X_0^2 + \dots + e_{n-1} X_0^{n-1} = \sum_{i=0}^{n-1} e_i X_0^i, \\ s_1 &= e_0 + e_1 X_1 + e_2 X_1^2 + \dots + e_{n-1} X_1^{n-1} = \sum_{i=0}^{n-1} e_i X_1^i, \\ &\dots \\ s_{n-k-1} &= e_0 + e_1 X_{n-k-1} + e_2 X_{n-k-1}^2 + \dots + e_{n-1} X_{n-k-1}^{n-1} = \sum_{i=0}^{n-1} e_i X_{n-k-1}^i. \end{aligned} \quad (11)$$

Задача декодирования вектора c^* состоит в нахождении всех e_i, $i = 0,\dots, n-1$ по известным элементам вектора $s = (s_0, s_1, \dots, s_{n-k-1})$. Система уравнений (11) содержит $n-k$ нелинейных уравнения от n неизвестных, прямых методов ее решения такой системы не известно. В алгебраической теории кодирования [38-40] для нахождения элементов вектора $e = (e_0, e_1, \dots, e_{n-1})$ используют искусственный прием, состоящий в рассмотрении многочлена локаторов ошибок $\Lambda(x)$, корнями которого являются ненулевые элементы вектора e, т.е.

$$\Lambda(x) = \prod_j (x + X_j), \quad (12)$$

где j – индекс ненулевых элементов вектора e, X_j - т.н. локатор ошибки, произошедшей в j-ом символе кодового слова.

Раскроем скобки в выражении (12), получим

$$\Lambda(x) = x^u + \lambda_{u-1} x^{u-1} + \dots + \lambda_1 x + \lambda_0, \quad (13)$$

где степень u многочлена $\Lambda(x)$ задает число произошедших ошибок на блоке из n символов, $u \le t$, т.е. число ненулевых элементов вектора e.

Набор коэффициентов $(\lambda_0, \lambda_1, ..., \lambda_{u-1})$ многочлена (13) однозначно задает его корни, которые однозначно указывают (локализуют) расположение произошедших ошибок.

Умножим многочлен (13) на $e_i X^i$ и вычислим его значение в X_j, получим:

$$e_i X_j^{u+i} + e_i \lambda_{u-1} X_j^{u+i-1} + ... + e_i \lambda_1 X_j^{i+1} + e_i \lambda_0 X_j^i = 0,$$

где $X_j \in GF(q^m)$, т.е. $X_j = \alpha^{Jj}$ (где α - примитивный элемент поля $GF(q^m)$) для некоторого J_j.

Следовательно, $X_j^{a+b} = \alpha^{a+b+Jj} = X^b_{j+a}$, т.е. справедливо выражение

$$e_i X^i_{j+u} + e_i \lambda_{u-1} X^i_{j+u-1} + ... + e_i \lambda_1 X^i_{j+1} + e_i \lambda_0 X^i_j = 0$$

Последнее равенство выполняется для любых j и i. Просуммируем по всем $i = 0...n-1$:

$$\sum_{i=0}^{N-1} \left(e_i X^i_{j+u} + e_i \lambda_{u-1} X^i_{j+u-1} + ... + e_i \lambda_1 X^i_{j+1} + e_i \lambda_0 X^i_j \right) = 0.$$

Изменим порядок суммирования, вынесем коэффициенты многочлена $\Lambda(x)$ за знак суммирования, получим:

$$\sum_{i=0}^{N-1} e_i X^i_{j+u} + \lambda_{u-1} \cdot \sum_{i=0}^{N-1} e_i X^i_{j+u-1} + ... + \lambda_1 \cdot \sum_{i=0}^{N-1} e_i X^i_{j+1} + \lambda_0 \cdot \sum_{i=0}^{N-1} e_i X^i_j = 0.$$

Значение каждого слагаемого в последнем выражении соответствует произведению коэффициентов многочлена $\Lambda(x)$ на соответствующие синдромы в выражении (11), так что запишем

$$s_{j+u} + \lambda_{u-1} \cdot s_{j+u-1} + ... + \lambda_1 \cdot s_{j+1} + \lambda_0 \cdot s_j = 0.$$

Перепишем выражение для каждого $j = 0...u$, получим систему линейных уравнений:

$$\begin{aligned} s_u + \lambda_{u-1} \cdot s_{u-1} + ... + \lambda_1 \cdot s_1 + \lambda_0 \cdot s_0 &= 0, \\ s_{u+1} + \lambda_{u-1} \cdot s_u + ... + \lambda_1 \cdot s_2 + \lambda_0 \cdot s_1 &= 0, \\ &... \\ s_{2 \cdot u} + \lambda_{u-1} \cdot s_{2 \cdot u-1} + ... + \lambda_1 \cdot s_{u+1} + \lambda_0 \cdot s_u &= 0. \end{aligned} \quad (14)$$

Система из u линейных уравнений (14) с u неизвестными разрешима, сложность ее решения растет полиномиально от числа неизвестных [38-40]. Так, например, для решения системы (14) методом Гаусса необходимо выполнить u^3 арифметических операций (сложений и умножений над элементами поля $GF(q^m)$).

Решение системы (14) дает значения коэффициентов многочлена локаторов ошибок (13). Корнями многочлена (13) являются локаторы – такие элементы поля $GF(q^m)$, которые однозначно указывают расположение ненулевых элементов вектора ошибок e. Следовательно, для локализации ошибок необходимо найти корни уравнения (13).

Наиболее простая процедура поиска корней $\Lambda(x)$ состоит в подстановке в многочлен всех n элементов $X_j \in GF(q^m)$ и выборе таких элементов, которые обращают в нуль $\Lambda(x)$. В литературе такой прием получил название процедура Ченя [38-40]. Используя схему Горнера перепишем многочлен $\Lambda(x)$ в виде

$$\Lambda(x) = x^u + \lambda_{u-1} x^{u-1} + \lambda_{u-2} x^{u-1} ... + \lambda_1 x + \lambda_0 = (...((x + \lambda_{u-1})x + \lambda_{u-2})x + ... + \lambda_1)x + \lambda_0.$$

Для вычисления значения многочлена $\Lambda(x)$ в такой форме потребуется не более $u-1$ арифметических операций (сложений и умножений над элементами поля $GF(q^m)$), т.е. сложность этого этапа декодирования не превысит $n(u-1)$ арифметических операций.

После локализации ошибок - нахождении локаторов ошибок X_j, необходимо вычислить значения ошибок в j-ом символе, т.е. вычислить элементы вектора e и восстановить кодовое слово: $c = c^* - e$. Для нахождения значений ошибок воспользуемся выражением (11). Подставим значения найденных локаторов X_j и неизвестные значения e_j в систему уравнений. Остальные e_i при $i \neq j$ равны нулю. Следовательно, система уравнений (11) запишется в виде:

$$s_0 = \sum_{i \in J} e_i X_0^i,$$

$$s_1 = \sum_{i \in J} e_i X_1^i,$$

$$\dots$$

$$s_{n-k-1} = \sum_{i \in J} e_i X_{n-k-1}^i,$$

где J – множество индексов ненулевых элементов вектора ошибок, т.е. набор номеров локаторов ошибок, причем $|J| = u \le t$. Полученная система из $n-k$ линейных уравнений содержит $|J| = u \le t$ неизвестных значений ошибок e_i, причем $t < n - k$. Следовательно, система разрешима, ее решение дает неизвестные ненулевые значения ошибок вектора e. Для решения системы уравнений от u неизвестных методом Гаусса, необходимо выполнить u^3 арифметических операций (сложений и умножений над элементами поля $GF(q^m)$). Для восстановления кодового слова длины n кодовых символов достаточно снять действие найденного вектора ошибок: $c = c^* - e$, т.е. выполнить u арифметических операций.

Таким образом, задача декодирования алгебраического блокового (n,k,d) кода (нахождения вектора ошибок $e = (e_0, e_1, ..., e_{n-1})$) сводится к решению двух систем линейных уравнений от $u \le t = \left\lfloor \frac{d-1}{2} \right\rfloor$ неизвестных и вычислению n значений многочлена $\Lambda(x)$ степени u. Для обращения матриц и решения систем линейных уравнений потребуется порядка u^3 арифметических операций, что при больших u может потребовать существенных вычислительных затрат. На практике для декодирования алгебраических кодов используют алгоритм Берлекэмпа-Мэсси, суть которого состоит в итеративном построении минимального регистра сдвига с обратной связью, генерирующего известную последовательность синдромов $s = (s_0, s_1, ..., s_{n-k-1})$. Сложность такого алгоритма составляет u^2 арифметических операций. Еще более эффективным, с точки зрения вычислительной сложности, является рекуррентный алгоритм Берлекэмпа-Месси [38-40]. Асимптотическая сложность декодирования кодов Рида-Соломона в этом случае не превосходит величины $O(n\log^2 n)$, причем очень близка к величине $O(n\log n)$.

Коды Рида-Соломона имеют небольшою длину – число корней $X_0, X_1, ..., X_{n-k-1} \in GF(q^m)$ не может быть больше числа элементов поля $GF(q^m)$ и, следовательно, $n \le q^m - 1$ (процедурами модификации кодов можно увеличить длину кода еще на 2 символа). Ограничением на подполе $GF(q)$ можно получить большую длину n кода при фиксированном q, однако кодовые (n,k,d) параметры лежат значительно ниже кодовых границ (1), (2) и с увеличением длины n это тенденция усиливается. Тем не менее существуют классы алгебраических кодов, которые лежат выше границ (1) и (2).

Определение 1 [38-40]. Пусть $X = (X_0, X_1, ..., X_{n-1})$ вектор над $GF(q^m)$, причем все X_i – различные элементы $GF(q^m)$. Пусть также $B = (B_0, B_1, ..., B_{n-1})$ – вектор над $GF(q^m)$ с необязательно различными B_i элементами $GF(q^m)$. Тогда (n, k, d) обобщенный код Рида-Соломона $ОРС_k(X, h)$ состоит из всех векторов вида

$$(B_0 \cdot F(X_0), B_1 \cdot F(X_1), ..., B_{n-1} \cdot F(X_{n-1})),$$

где $F(x)$ – любой многочлен с коэффициентами из $GF(q^m)$, степень которого не превосходит k.

Код ОРС является МДР кодом, его проверочная матрица $ОРС_k(X, h)$ равна:

$$H=\begin{pmatrix} Y_0 & Y_1 & \dots & Y_{n-1} \\ X_1\cdot Y_0 & X_2\cdot Y_1 & \dots & X_{n-1}\cdot Y_{n-1} \\ X_1^2\cdot Y_0 & X_2^2\cdot Y_1 & \dots & X_{n-1}^2\cdot Y_{n-1} \\ \dots & \dots & \dots & \dots \\ X_1^{n-k-1}\cdot Y_0 & X_2^{n-k-1}\cdot Y_1 & \dots & X_{n-1}^{n-k-1}\cdot Y_{n-1} \end{pmatrix}= \\ =\begin{pmatrix} 1 & 1 & \dots & 1 \\ X_1 & X_2 & \dots & X_{n-1} \\ X_1^2 & X_2^2 & \dots & X_{n-1}^2 \\ \dots & \dots & \dots & \dots \\ X_1^{n-k-1} & X_2^{n-k-1} & \dots & X_{n-1}^{n-k-1} \end{pmatrix}\cdot\begin{pmatrix} Y_0 & 0 & \dots & 0 \\ 0 & Y_1 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & Y_{n-1} \end{pmatrix}, \tag{15}$$

где вектор $Y = (Y_0, Y_1, \dots, Y_{n-1})$ такой, что $\forall Y_i \in GF(q^m)$, $Y_i \neq 0$ и дуальным к $ОРС_k(X, B)$ является $ОРС_{n-k}(X, Y)$.

Через определение ОРС вводится обширный класс т.н. альтернантных кодов [38-40].

Определение 2 [38-40]. Альтернантный (n, k, d) код $A(X, B)$ состоит из всех слов кода $ОРС_k(X, B)$ таких, что их компоненты лежат в поле $GF(q)$. Другими словами, $A(X, B)$ равен ограничению кода $ОРС_k(X, B)$ на подполе $GF(q)$, т.е. он состоит из всех векторов c над $GF(q)$, для которых выполняется равенство $cH^T = 0$, где H – проверочная матрица $ОРС_k(X, B)$, задаваемая выражением (15). Порождающая матрица $A(X, B)$ может быть получена заменой каждого элемента матрицы H в (15) соответствующим вектором-столбцом длины m над $GF(q)$.

Параметры кода $A(X, B)$ связаны соотношением: $n - mr \le k \le n - r$; $d \ge r + 1$, причем доказано [38-40], что среди большого числа всех возможных альтернантных кодов при фиксированном n и k найдутся такие коды, параметры которых лежат выше кодовых границ (1) и (2). Одним из частных случаев $A(X, B)$ являются коды Гоппы [42,43].

Определение 3 [40]. Альтернантный (n, k, d) код Гоппы $\Gamma(L, G)$ над $GF(q)$ состоит из всех векторов $c = (c_1, c_2, \dots, c_n)$ таких, что

$$R_c(x) \equiv 0 \bmod G(x), \tag{16}$$

где

$$R_c(x) = \sum_{i=1}^{n} \frac{c_i}{x - \alpha_i},$$

$G(x)$ – многочлен с коэффициентами из $GF(q^m)$ (многочлен Гоппы), $L = (\alpha_1, \alpha_2, \dots, \alpha_n)$ – подмножество элементов из $GF(q^m)$ таких, что $G(\alpha_i) \neq 0 \ \ \forall \alpha_i \in L$.

Используя выражение (16) проверочную матрицу кода Гоппы можно задать следующим образом. Многочлен $x - \alpha_i$ в кольце многочленов по модулю $G(x)$ имеет обратный многочлен:

$$(x - \alpha_i)^{-1} = -\frac{G(x) - G(\alpha_i)}{x - \alpha_i} G^{-1}(\alpha_i).$$

Следовательно, вектор $c = (c_1, c_2, \dots, c_n)$ принадлежит коду Гоппы $\Gamma(L, G)$ тогда и только тогда, когда

$$\sum_{i=1}^{n} c_i \frac{G(x) - G(\alpha_i)}{x - \alpha_i} G^{-1}(\alpha_i) = 0. \tag{17}$$

Если $G(x) = \sum_{i=0}^{r} g_i x^i$, где $g_i \in GF(q^m)$ и $g_r \neq 0$, то

$$\frac{G(x) - G(\alpha_i)}{x - \alpha_i} = g_r(x^{r-1} + x^{r-2}\alpha_i + \dots + \alpha_i^{r-1}) + g_{r-1}(x^{r-2}\alpha_i + \dots + \alpha_i^{r-1}) + \dots + g_2(x + \alpha_i) + g_1.$$

Приравнивая согласно (17) нулю все коэффициенты при x^{r-1}, x^{r-2}, …, *1*, получим, что условие $cH^T = 0$ выполнится только если

$$H = \begin{pmatrix} g_r G^{-1}(\alpha_1) & g_r G^{-1}(\alpha_2) & \dots \\ (g_{r-1} + \alpha_1 g_r) G^{-1}(\alpha_1) & (g_{r-1} + \alpha_2 g_r) G^{-1}(\alpha_2) & \dots \\ \dots & \dots & \dots \\ (g_1 + \alpha_1 g_2 + \dots + \alpha_1^{r-1} g_r) G^{-1}(\alpha_1) & (g_1 + \alpha_2 g_2 + \dots + \alpha_2^{r-1} g_r) G^{-1}(\alpha_2) & \dots \end{pmatrix}$$

$$\begin{pmatrix} \dots & g_r G^{-1}(\alpha_n) \\ \dots & (g_{r-1} + \alpha_n g_r) G^{-1}(\alpha_n) \\ \dots & \dots \\ \dots & (g_1 + \alpha_n g_2 + \dots + \alpha_n^{r-1} g_r) G^{-1}(\alpha_n) \end{pmatrix} =$$

$$= \begin{pmatrix} g_r & 0 & \dots & 0 \\ g_{r-1} & g_r & \dots & 0 \\ \dots & \dots & \dots & \dots \\ g_1 & g_2 & \dots & g_r \end{pmatrix} \begin{pmatrix} 1 & 1 & \dots & 1 \\ \alpha_1 & \alpha_2 & \dots & \alpha_n \\ \dots & \dots & \dots & \dots \\ \alpha_1^{r-1} & \alpha_2^{r-1} & \dots & \alpha_n^{r-1} \end{pmatrix} \begin{pmatrix} G^{-1}(\alpha_1) & 0 & \dots & 0 \\ 0 & G^{-1}(\alpha_2) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & G^{-1}(\alpha_n) \end{pmatrix}.$$

Матрица

$$\begin{pmatrix} g_r & 0 & \dots & 0 \\ g_{r-1} & g_r & \dots & 0 \\ \dots & \dots & \dots & \dots \\ g_1 & g_2 & \dots & g_r \end{pmatrix}$$

- обратима. Следовательно, проверочная матрица

$$H = \begin{pmatrix} 1 & 1 & \dots & 1 \\ \alpha_1 & \alpha_2 & \dots & \alpha_n \\ \dots & \dots & \dots & \dots \\ \alpha_1^{r-1} & \alpha_2^{r-1} & \dots & \alpha_n^{r-1} \end{pmatrix} \begin{pmatrix} G^{-1}(\alpha_1) & 0 & \dots & 0 \\ 0 & G^{-1}(\alpha_2) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & G^{-1}(\alpha_n) \end{pmatrix} =$$

$$= \begin{pmatrix} G^{-1}(\alpha_1) & G^{-1}(\alpha_2) & \dots & G^{-1}(\alpha_n) \\ \alpha_1 G^{-1}(\alpha_1) & \alpha_2 G^{-1}(\alpha_2) & \dots & \alpha_n G^{-1}(\alpha_n) \\ \dots & \dots & \dots & \dots \\ \alpha_1^{r-1} G^{-1}(\alpha_1) & \alpha_2^{r-1} G^{-1}(\alpha_2) & \dots & \alpha_n^{r-1} G^{-1}(\alpha_n) \end{pmatrix}$$

также задает *(n, k, d)* код Гоппы *Γ(L, G)* над *GF(q)*.

Последнее выражение при $Y = (Y_1, Y_2, \dots, Y_n)$, $Y_1 = G^{-1}(\alpha_1)$, $Y_2 = G^{-1}(\alpha_2)$, …, $Y_n = G^{-1}(\alpha_n)$ эквивалентно выражению (15). Проверочную матрицу *Γ(L, G)* над *GF(q)* с элементами из *GF(q)* можно получить путем представления каждого элемента из $GF(q^m)$ вектором-столбцом длины *m* символов из *GF(q)*. Справедлива следующая оценка.

Теорема 3 [40,42,43]. Параметры *(n, k, d)* кода Гоппы *Γ(L, G)* связаны соотношениями: $n = |L|$, $k \geq n - mr$, $r = deg\ G(x)$, $d \geq r + 1$.

Для сепарабельных (когда многочлен *G(x)* не имеет кратных корней ни в одном расширении поля) двоичных кодов Гоппы минимальное кодовое расстояние $d \geq 2r + 1$. Причем, если *G(x)* - неприводимый многочлен степени *r* над $GF(q^m)$ и $L = GF(q^m)$, тогда существует код Гоппы над $GF(q^m)$ лежащий на границе Варшамова-Гилберта [40,42,43].

Теорема 3 гарантирует существование альтернантных кодов, построенных через многочлен Гоппы, с кодовыми характеристиками, удовлетворяющими (1), (2). При соответствую-

щем выборе вектора-шаблона $Y = (Y_1, Y_2, \ldots, Y_n)$ удается построить блоковые коды лежащие выше границы Варшамово-Гилберта [40,42,43]. Это свойство рассмотренных кодовых конструкций указывает на перспективность применения альтернантных кодов, в том числе и кодов Гоппы, для решения различных инженерных задач как в области повышения помехоустойчивости передачи данных, так и для криптографической защиты информационных ресурсов. В частности, использование рассмотренных положений алгебраической теории блоковых кодов в криптографических целях позволяет реализовать несимметричные криптосистемы доказуемой стойкости *(provable security)*, которые, помимо высокой скорости двухключевого криптографического преобразования и возможности совмещать контроль ошибок с защитой от несанкционированного ознакомления [15-22], остаются стойкими даже в случае использования квантовых вычислений [28].

3 Несимметричные криптосистемы на основе алгебраических блоковых кодов (теоретико-кодовые схемы)

Рассмотрим схемы несимметричного криптографического преобразования, построение которых основано на использовании алгебраических кодов, замаскированных под код общего положения (случайный код, полный код) [15-22]. Рассмотрим современное состояние, существующие противоречия и перспективы их практического использования в том числе на постквантовый период.

Криптосистема Мак-Элиса. Первой и наиболее изученной схемой несимметричного шифрования, основанной на использовании алгебраических блоковых кодов, является предложенная в 1978 году криптосистема Мак-Элиса (McEliece) [15]. Она обладает неоспоримыми преимуществами: высокой скоростью криптографического преобразования, а также возможностью совмещать контроль ошибок с защитой от несанкционированного ознакомления [15-22]. Подобные (крипто-кодовые) преобразования остаются стойкими и при использовании квантовых вычислений [28]. Открытым ключом в схеме Мак-Элиса является матрица

$$G_X = XGPD, \tag{18}$$

где G – порождающая матрица алгебраического (n,k,d) кода над $GF(q)$ (в оригинальной статье [15] предлагалось использовать рассмотренный выше двоичный код Гоппы), X - невырожденная $k \times k$ матрица с элементами из $GF(q)$, P и D – перестановочная и диагональная $n \times n$ матрицы (для двоичных кодов используется только матрица P).

Матрицы X, P и D являются секретным ключом, который маскирует используемый алгебраический блоковый код под случайный код (код общего положения), т.е. открытый ключ G_X представляется злоумышленнику как случайно сформированная порождающая матрица некоторого линейного кода, для которого неизвестен алгоритм быстрого декодирования. Напротив, уполномоченный пользователь, знающий секретный ключ (матрицы X, P и D), может снять действие маскирующих матриц и воспользоваться быстрым алгоритмом декодирования алгебраического кода с порождающей матрицей G.

Криптограмма представляет собой вектор длины n, который вычисляется по правилу

$$c_X^* = IG_X + e, \tag{19}$$

где вектор

$$c_X = IG_X$$

является кодовым словом замаскированного кода, т.е. c_X принадлежит (n,k,d) коду с порождающей матрицей G_X, I – k-разрядный информационный вектор над $GF(q)$, вектор e – секретный вектор ошибок веса

$$w_h(e) \le t = \left\lfloor \frac{d-1}{2} \right\rfloor.$$

Вектор e следует рассматривать как одноразовый сеансовый секретный ключ, его вес определяет сложность декодирования искаженного кодового слова (криптограммы) c_X^*. Злоумышленнику необходимо декодировать кодограмму c_X^* используя известную ему порождающую матрицу G_X. Однако декодирование случайного кода (при соответствующих параметрах n, k, q и $w_h(e)$) вычислительно недостижимо. Не зная матрицы X, P и D злоумышленник не может восстановить матрицу G и воспользоваться алгоритмом декодирования полиномиальной сложности. Из этих соображений величину $w_h(e)$ следует максимизировать, например, при $w_h(e) = t$ сложность декодирования будет максимальной, что обеспечит наивысший уровень стойкости кодовой криптосистемы для заданных параметров n, k, q.

Для уполномоченного пользователя (знающего секретный ключ) декодирование – полиномиально разрешимая задача. Действительно, легитимный пользователь, получив вектор c_X^*, строит вектор $\overline{c}^* = c_X^* \cdot D^{-1} \cdot P^{-1}$. Матрица $\Lambda = PD$ сохраняет вес и расстояние по Хеммингу, т.е. для любых кодовых слов c и c' выполняются равенства:

$$w_h(c) = w_h(c\Lambda),\ w_h(c, c') = w_h(c\Lambda, c'\Lambda).$$

Это означает, что вектор $\overline{c}^*$ является искаженным не более чем в t разрядах кодовым словом алгебраического кода с порождающей матрицей G и его можно декодировать быстрым алгоритмом полиномиальной сложности [19].

Уполномоченный пользователь, используя алгоритмом полиномиальной сложности, декодирует вектор $\overline{c}^* = I' \cdot G + e'$, т.е. находит I'. Затем он вычисляет k-разрядный информационный вектор $I = I' X^{-1}$.

Таким образом, в криптосистеме Мак-Элиса основным средством маскировки линейного блокового *(n, k, d)* кода под линейный случайный код (код общего положения) являются матрицы *X, P, D*. Дополнительным секретным параметром, который можно использовать в случае кодов Гоппы, является многочлен Гоппы *G(x)*, или, в более широком смысле, вектор $Y = (Y_0, Y_1, \ldots, Y_{n-1})$ в случае альтернантных кодов (см. выражения (15)-(17)). Изменение шаблона не снижает конструктивных кодовых характеристик, т.е. с точки зрения криптографического преобразования не приведет к снижению безопасности. Однако знание вектора-шаблона $Y = (Y_0, Y_1, \ldots, Y_{n-1})$ (или многочлена *G(x)*) является необходимым для правильного декодирования информационного сообщения, т.е. для корректного расшифрования на приемной стороне.

Опубликовано большое число различных атак на крипто-кодовые схемы защиты информации [19,33-36], некоторые из которых оказались достаточно эффективными относительно отдельных вариантов кодовых криптосистем. Однако базовая конструкция [15], предложенная около 40 лет назад, остается стойкой ко всем известным, на сегодняшний день, методам криптоанализа, в том числе и в случае использования квантовых вычислительных систем.

Наиболее естественным направлением в развитии методов криптоанализа кодовой схемы Мак-Элиса является использование неалгебраических методов декодирования. Действительно, если существует вычислительно эффективный способ декодирования кодового слова (19) только по известной порождающей матрице (18), тогда информационное сообщение I может быть эффективно восстановлено и без знания секретного ключа (матриц X, P и D).

Среди универсальных методов декодирования линейных блоковых кодов, заданных произвольной порождающей матрицей, особое место занимают перестановочные алгоритмы [38-40]. Основная идея такого декодирования состоит в использовании различных наборов информационных множеств. Представим порождающую матрицу (19) в каноническом виде (4). Единичные вектора-столбцы в (4) могут быть выбраны произвольно с соответствующим формированием единичных подматриц и систематическим размещением k символов информационного множества. Оставшиеся $(n-k)$ символов однозначно вычисляются по элементам информационного множества. Позиции этих $(n-k)$ символов задают размещение

единичных вектор-столбцов соответствующей проверочной матрицы H^* (6). Если выбрать размещение $(n-k)$ единичных вектор-столбцов в (6) таким образом, чтобы они покрыли все t позиций ненулевых элементов вектора ошибок e, тогда кодовое слово, вычисленное по k символам информационного множества, не будет содержать ошибок, т.е. слово (19) можно декодировать даже без знания специальной алгебраической структуры порождающей (проверочной) матрицы используемого алгебраического кода.

Таким образом, при реализации переставного декодирования конкретная комбинация ошибок будет исправлена, только если удастся найти такое информационное множество, которое целиком содержит эту комбинацию. Такое множество, являющееся кровельной комбинацией ошибок, и набор проверочных множеств, которые покрывают все наборы ошибок данного типа, называют покрытием [38]. Задача декодера состоит в том, чтобы найти проверочное множество, которое покрывает неизвестную комбинацию ошибок.

Рассмотрим границы для количества кровельных множеств. Предположим, что с помощью *(n, k, d)* кода исправляются все комбинации из *t* или меньшего количества ошибок. Рассмотрим комбинацию только из *t* кратных ошибок, так как все ошибки меньшей кратности будут покрыты. Общее количество комбинаций ошибок во всех *n* позициях равно $C_n^t = \frac{n!}{t!(n-t)!}$. Поскольку объем кровельного множества равен *n – k*, максимальное количество комбинаций ошибок, которые могут быть покрытые данным множеством равно $C_{n-k}^t = \frac{(n-k)!}{t!(n-k-t)!}$. Наименьшее количество множеств, которые могут исправить все комбинации из *t* ошибок, ограничивается выражением [38]:

$$N \geq \frac{C_n^t}{C_{n-k}^t} = \frac{\frac{n!}{t!(n-t)!}}{\frac{(n-k)!}{t!(n-k-t)!}} = \frac{n!(n-k-t)!}{(n-t)!(n-k)!}. \qquad (20)$$

На рис. 1 приведены зависимости наименьшего числа кровельных множеств, которые потребуется для исправления всех комбинации из *t* ошибок произвольного линейного блокового кода. Оценки *N* приведены в логарифмическом масштабе в зависимости от относительной скорости кодирования *R = k/n* и рассчитаны для параметров двоичных сепарабельных кодов Гоппы: $n = 2^m$, $k \geq n - mr$, $r = deg\ G(x)$, $d \geq 2r + 1$.

Зависимости, приведенные на рис. 1, можно интерпретировать как оценки стойкости крипто-кодовых преобразований, выраженные в наименьшем числе покрывающих множеств, которые потребуется перебрать для декодирования любой конфигурации вектора ошибок e. Эти зависимости не учитывают вычислительную сложность формирования слов-кандидатов, вычисляемых по выбранной конфигурации информационного множества *(реальная стойкость будет еще выше)*.

Как следует зависимостей, представленных на рис.1, наибольшую стойкость схема Мак-Элиса обеспечивает при использовании кодов с относительной скоростью $R \approx 2/3$, что согласуется с выводами большинства исследований [28].

В таблице 1 приведены параметры некоторых схем с кодами Гоппы и $R \approx 2/3$, оценки стойкости к атаке перестановочного декодирования, оценки вычислительной сложности кодирования (зашифрования) и декодирования (расшифрования), а также аналогичные оценки для несимметричного шифрования RSA и блочного симметричного шифрования AES (FIPS-197) / Kalyna (DSTU 7624:2014) [44, 45].

Для схемы Мак-Элиса значения в таблице 1 оценивались следующим образом:

- размер открытого ключа оценивался как число двоичных элементов матрицы $G_X - kn$ бит;
- размер закрытого ключа оценивался как число элементов матрицы X (k^2 бит) плюс число элементов, необходимых для хранения правила перестановки ($n\log_2 n$ бит).

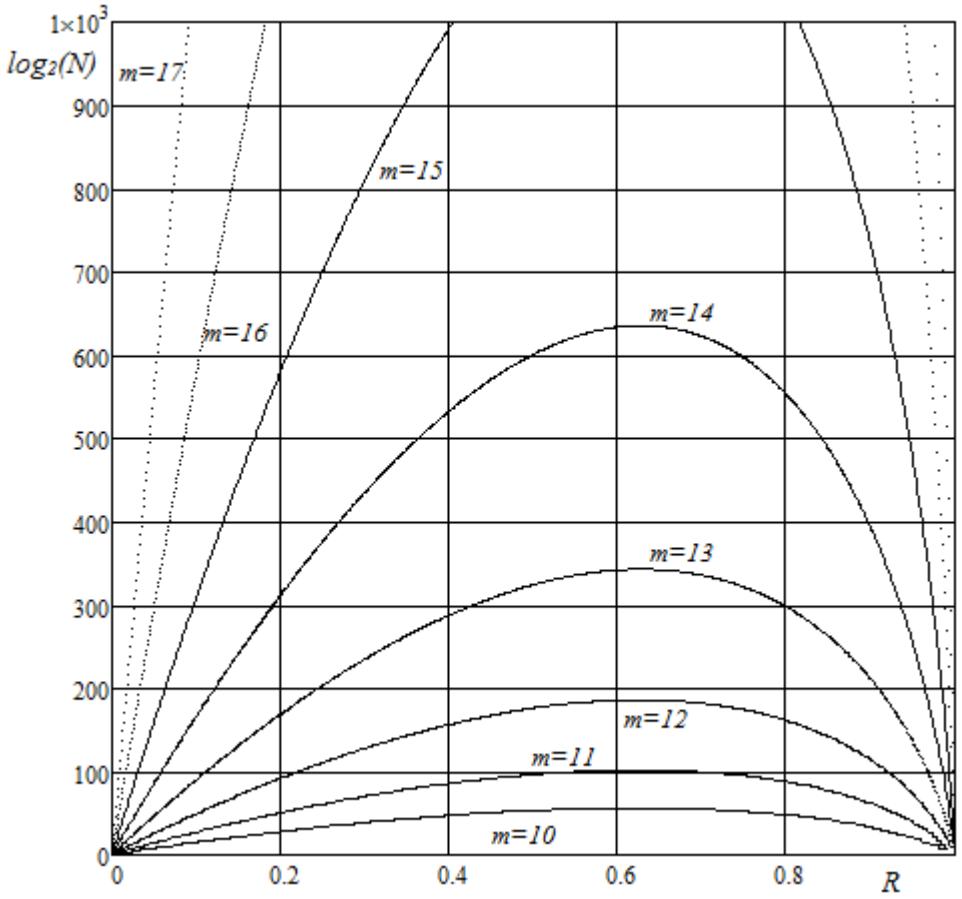

Рис. 1 – Оценка стойкости схемы Мак-Элиса на сепарабельных двоичных кодах Гоппы к атаке перестановочным декодированием

Таблица 1 – Оценка характеристик криптосистемы Мак-Элиса

Оцениваемые параметры	**Уровень стойкости** *(без учета квантового криптоанализа)*			
	Достаточный (2^{80} … 2^{128})	**Высокий** (2^{192} … 2^{256})		**Сверхвысокий** ($> 2^{512}$)
Криптосистема Мак-Элиса				
Параметры (*n, k, d*)	(2 048, 1 300, 137)	(4 096, 2 584, 253)		(16 384, 10 322, 867)
Размер секретного ключа, бит	1 712 528	6 726 208		106 773 060
Размер открытого ключа, бит	2 662 400	10 584 064		169 115 648
Сложность шифрования, операций XOR	1 300	2 584		10 322
Сложность расшифрования, операций над $GF(2^m)$	4 624	15 876		187 489
Оценка стойкости (эквивалентная длина ключа симметричного шифра), $\log_2 N$	102	186		636
Оценка стойкости к квантовому криптоанализу, бит	49	91		310
Криптосистема RSA				
Размер модуля и открытого (закрытого) ключа, бит	2 048	7 680	15360	
Сложность шифрования (расшифрования), битовых операций	$3{,}2 \cdot 10^9$	$1{,}7 \cdot 10^{11}$	$1{,}4 \cdot 10^{12}$	
Оценка стойкости (эквивалентная длина ключа симметричного шифра), бит	112	192	256	
Оценка стойкости к квантовому криптоанализу, бит	40	41	44	
Блочный симметричный шифр AES (FIPS-197) / Kalyna (DSTU 7624:2014)				
Размер секретного ключа (оценка стойкости), бит	128	196	256	512
Сложность шифрования (расшифрования), операций на слово	40	48	56	72
Оценка стойкости к квантовому криптоанализу, бит	64	98	128	256

Сложность шифрования оценивалась как максимальное число операций, которые необходимо выполнить для формирования кодового слова посредством матричного вычисления выражения (19). Для двоичного *(n, k, d)* кода это соответствует *k* операциям XOR над *n* битными словами. Если вычисления реализуются под управлением 32(64)-битной операционной системы, тогда каждое *n* битное кодовое слово представляется как набор из *n/32* (*n/64*) машинных слов и для вычисления каждого из них потребуется выполнить не более *k* операций XOR.

Сложность расшифрования оценивалась как максимальное число арифметических операций над конечным полем $GF(2^m)$, которые необходимо выполнить для декодирования кодового слова с ошибками. При этом сложность декодирования оценивалась как t^2. В практических приложениях операции сложения элементов поля $GF(2^m)$ реализуются операцией XOR, а операции умножения – табличным способом, т.е. через обращение к ячейке памяти с заданным входными аргументами адресом, так что оценка t^2 выглядит вполне правдоподобной.

Оценка стойкости кодовой криптосистемы Мак-Элиса к квантовому криптоанализу проведена в работах [46, 47]. В частности, в [47] приводится оценка числа итераций для декодирования квантовых алгоритмом Гровера *(Grover's algorithm)*. Эта оценка имеет вид:

$$C^{\frac{n}{2\log n}},\ C = \frac{1}{(1-R)^{1-R}}, \tag{21}$$

где $R = k/n$ - относительная скорость используемого кода.

Значения, приведенные в таблице 1, рассчитаны по соотношению (21). На практике оценка (21) снижает стойкость криптосистемы *(примерно в два раза уменьшается эквивалентная длина ключа)*, что, впрочем, вполне ожидаемо для надежных постквантовых алгоритмов *(как и для большинства симметричных шифров)*.

Для криптосистемы RSA значения в таблице 1 оценивались следующим образом. Скорость криптопреобразования (шифрования и расшифрования) оценивалась как сложность модульного возведения в степень. В работе [48] показано (стр. 613), что в общем случае для *l*-битных чисел операция модульного возведения в степень требует порядка $\frac{3}{2}l^3$ двоичных операций. Пусть *p* и *q* – два *l*-битных простых числа, модуль преобразования RSA (общесистемный параметр) равен *n* = *pq* (2*l*-битное число), а открытым (секретным) ключом являются 2*l*-битные числа *e* и *d*. Тогда сложность модульного возведения в степень при шифровании (расшифровании) потребует $\frac{3}{2}(2l)^3 = 12l^3$ операций. Более эффективным является последовательное вычисление возведения в степень по модулям (*p – 1*) и (*q – 1*), соответственно. Такой алгоритм потребует в два раза большее число операций, однако в связи с уменьшением размерности модулей общее число операций сократится. Сложность преобразования составит $2\frac{3}{2}l^3 = 3l^3$ и значения, приведенные в таблице 1, соответствуют этой оценке. Размер модуля и соответствующая оценка стойкости (как эквивалентная длина ключа симметричного шифра) указана в работе [49].

Оценки объема квантовых ресурсов, необходимых для решения некоторых асимметричных криптографический задач с помощью алгоритма Шора, при различных параметрах этих задач, и сравнение их со сложностью решения переборного задачи при поиске ключа симметричного шифра приведены в [50]. В частности, для m-битного числа дается оценка $4m^3$ временной сложности квантового алгоритма факторизации Шора и значения, приведенные в таблице 1, соответствуют этой оценке.

Описание блочных симметричных шифров AES (FIPS-197) и Kalyna (DSTU 7624:2014) приведено в [44, 45]. Исследование сложности квантовых алгоритмов криптоанализа симметричных шифров представлено в [51]. В частности, квантовый алгоритм Гровера для решения переборных задач, в том числе, переборного поиска m-битного секретного ключа

симметричного шифра, требует выполнения $\frac{\pi}{4}\sqrt{2^m}$ итераций. На практике же это приводит к соответствующему снижению стойкости *(в два раза уменьшается эквивалентная длина ключа)*.

Следует обратить внимание на высокую скорость криптографического преобразования в схеме Мак-Элиса, которая приближается по скорости шифрования к блочным симметричным шифрам. Действительно, при использовании кода Гоппы с рекомендованными в авторской статье [15] параметрами

$$n = 1024, k = 524, t = 50, d = 2t + 1 = 101,$$

для зашифрования матричным способом (вычисление $IG_X + e$) потребуется выполнить не более 524 операций XOR на одно обрабатываемое слово.

Для примера, один из самых быстрых современных блочных симметричных шифров AES *(американский стандарт шифрования FIPS-197)* требует для зашифрования не менее 4 операций XOR на 32-х битное слово на каждом раунде [44], что при 10 раундах составит не менее 40 операций XOR.

Вторым важным преимуществом схемы Мак-Элиса является возможность совмещать криптографичское преобразование с контролем возникающих ошибок. Действительно, если при формировании криптограммы (18) использовать случайный вектор ошибок e, веса $w(e) < t$, тогда появляется возможность одновременно с криптографическим преобразованием данных контролировать ошибки в пределах исправляющей способности. Уменьшение веса вектора e снизит криптографическую стойкость схемы Мак-Элиса, однако повысит помехоустойчивость передачи данных, т.е. в такой «гибридной» схеме изменяя $w(e)$ можно адаптивно реагировать на потребность в соответствующих услугах безопасности.

Обозначим долю веса вектора ошибок вектора e, приходящегося на искусственное внесение при формировании криптограммы (см. выражение (18)) символом $\rho = w(e) / t$. Тогда стойкость криптосистемы, построенная на алгебраических кодах, будет определяться величиной $\rho \cdot t$, а обеспечиваемая помехоустойчивость передаваемых криптограмм определяться величиной $(1 - \rho) \cdot t$.

Третье и, очевидно, одно из важнейших положительных свойств криптосистемы Мак-Элиса является высокая устойчивость к квантовому криптоанализу. По сравнению с другими несимметричными криптосистемами, например, с RSA, сложность квантового криптоанализа кодовой криптосистемы с увеличением ее параметров возрастает очень быстро. Фактически, сложность криптоанализа квантовыми алгоритмами сопоставима с решением переборных задач поиска эквивалентных ключей симметричных шифров. Данные таблицы 1 наглядно подтверждают эту тенденцию.

Основными недостатками рассмотренной кодовой крптосистемы являются огромные объемы ключевых данных (десятки мегабит), а также снижение относительной скорости передачи информации (наибольшая стойкость криптосистемы достигается при относительной скорости кодирования $R = k/n \approx 2/3$). Ниже будет показано, что относительную скорость передачи данных можно существенно повысить (рассматриваемые в данной работе кодовые криптосистемы снимают этот конструктивный недостаток схемы Мак-Элиса).

Криптосистема Нидеррайтера. Альтернативным примером криптосистем на кодах есть схема Нидеррайтера, впервые предложенная в [16]. Открытым ключом в этой криптосистеме есть матрица

$$H_X = XHPD, \tag{22}$$

где H – проверочная матрица алгебраического (n,k,d) кода над $GF(q)$ (в оригинальной статье [16] предлагалось использовать обобщенные коды Рида-Соломона), X - невырожденная $(n-k)\times(n-k)$ матрица с элементами из $GF(q)$, P и D – перестановочная и диагональная $n \times n$ матрицы (для двоичных кодов используется только матрица P).

Матрицы X, P и D (как и для криптосистемы Мак-Элиса) являются секретным ключом, который маскирует используемый алгебраический блоковый код под случайный код (код

общего положения), т.е. открытый ключ (22) представляется злоумышленнику как случайно сформированная проверочная матрица некоторого линейного кода, для которого неизвестен алгоритм быстрого декодирования.

Напротив, уполномоченный пользователь, знающий секретный ключ (матрицы X, P и D), может снять действие маскирующих матриц и воспользоваться быстрым алгоритмом декодирования алгебраического кода с проверочной матрицей H.

Криптограмма S_X представляет собой вектор длины $(n-k)$ и вычисляется по правилу

$$S_X = e \cdot H_X^T, \tag{23}$$

где вектор e – вектор длины n и веса $w_h(e) \le t$, который несет конфиденциальную информацию (информационное сообщение, подлежащее зашифрованию). Наибольшая стойкость обеспечивается при $w_h(e) = t$.

Уполномоченный пользователь (имеющий секретный ключ) находит одно из q^k решений выражения $S_X = c_X^* \cdot H_X^T$. Найденное решение - суть кодовое слово с ошибками $c_X^* = I \cdot G_X + e$. Далее, как и в схеме Мак-Элиса, уполномоченный пользователь строит вектор $\bar{c}^* = c_X^* \cdot D^{-1} \cdot P^{-1}$ и декодирует полученное слово. Однако, вместо восстановления информационного слова I', он вычисляет кодовое слово $c' = I' \cdot G$, а затем и вектор ошибок $e' = \bar{c}^* - c'$. На последнем шаге производится вычисление вектора $e = e' \cdot P \cdot D$, который несет конфиденциальную информацию.

Таким образом, в криптосистеме Нидеррайтера основным средством маскировки линейного кода под случайный код являются (как и в криптосистеме Мак-Элиса) матрицы *X, P, D*. Если использовать коды Гоппы, тогда многочлен *G(x)* может выступать дополнительным секретным параметром.

В работе [19] показано, что стойкости криптосистем Мак-Элиса и Ниддерайтера эквивалентны и эффективную атаку на одну из схем можно легко трансформировать в атаку на другую схему. В этом смысле оценки стойкости криптосистемы Мак-Элиса, приведенные в таблице 1, справедливы и по отношению к криптосистеме Ниддерайтера. Другие характеристики этих криптосистем (скорость шифрования/расшифрования, объемы закрытого и открытого ключа) также сопоставимы.

Очевидным преимуществом теоретико-кодовой схемы Нидеррайтера по сравнению с криптосистемой Мак-Элиса является потенциально большая относительная скорость передачи данных. Действительно, относительная скорость в криптосистеме Мак-Элиса определяется относительной скоростью используемого (n,k,d) кода, т.е. равна $R = k/n$, причем наибольшая стойкость достигается при $R = k/n \approx 2/3$ (см. рис. 1). Информационное сообщение в системе Нидеррайтера сперва преобразуется в равновесную последовательность e длины n и веса $w_h(e) \le t$, а затем умножается на проверочную матрицу как в (23).

Положим $w_h(e) = t$ (в этом случае будет обеспечена максимальная стойкость криптосистемы для заданных (n,k,d) параметров кода). Тогда максимальное число бит информационных данных, которые можно зашифровать в системе Нидеррайтера при использовании двоичного (n,k,d) кода, будет определятся выражением:

$$l_{\inf} = \left\lfloor \log_2 C_n^t \right\rfloor = \left\lfloor \log_2 \left(\frac{n!}{t!(n-t)!} \right) \right\rfloor,$$

где $\lfloor x \rfloor$ - наибольшее целое число, меньшее x.

Криптограмма (23) представляет собой синдромный вектор длины $n-k$, т.е. относительная скорость передачи данных в криптосистеме Нидеррайтера (для двоичных кодов) будет определяться выражением:

$$R^* = \frac{\left\lfloor \log_2\left(\frac{n!}{t!(n-t)!}\right)\right\rfloor}{n-k}.$$

Последнее выражение легко обобщается на случай недвоичных кодов с основанием q:

$$R^* = \frac{\left\lfloor \log_q\left((q-1)^t \frac{n!}{t!(n-t)!}\right)\right\rfloor}{n-k}. \quad (24)$$

Если предположить, что информационная последовательность будет преобразовываться во все возможные вектора e длины n и веса $0 \le w_h(e) \le t$, тогда последнее выражение примет вид:

$$R^* = \frac{\left\lfloor \log_q\left(\sum_{i=0}^{t}(q-1)^i \frac{n!}{i!(n-i)!}\right)\right\rfloor}{n-k}. \quad (25)$$

Алгоритм кодирования информационной последовательности в равновесную последовательность e длины n и веса $w_h(e)$ для произвольного основания q приводится, например, в работе [52].

Выражение (25) достигает максимума для т.н. *совершенных кодов* (perfect codes), кодовые (n,k,d) параметры которых удовлетворяют верхней границе Хемминга для мощности (числа кодовых слов) $A_q(n,d)$ произвольного линейного q-ичного кода [38-40]:

$$A_q(n,d) \le \frac{q^n}{\sum_{i=0}^{t}(q-1)^i \frac{n!}{i!(n-i)!}}. \quad (26)$$

Мощность линейного (n,k,d) кода над $GF(q)$ равна q^k, следовательно из (26) следует ограничение на число информационных символов кода

$$k \le n - \log_q\left(\sum_{i=0}^{t}(q-1)^i \frac{n!}{i!(n-i)!}\right).$$

Если параметры (n,k,d) кода удовлетворяют верхней границе Хемминга, т.е. достигается равенство в (26) и код совершенен, тогда

$$\log_q\left(\sum_{i=0}^{t}(q-1)^i \frac{n!}{i!(n-i)!}\right) = n-k, \quad (27)$$

что после подстановки в (25) дает $R^* = 1$, т.е. относительная скорость передачи максимальна и криптограмма в схеме Нидеррайтера не будет содержать избыточных символов.

В качестве примера приведем совершенный двоичный ($q = 2$) код Хемминга исправляющий одну ошибку ($t = 1$). Он определен для любого положительного целого $m > 2$ и имеет кодовые параметры $(2^m - 1, 2^m - m - 1, 3)$ [38 – 40]. Очевидно, что для этих значений

$$\sum_{i=0}^{t}\frac{n!}{i!(n-i)!} = 2^m,\ n-k = m$$

и относительная скорость (25) равна 1.

Другим примером является совершенный двоичный код Голея *(perfect binary Golay code)* с параметрами $(23, 12, 7)$ [38-40]. Он позволяет исправить $t = 3$ ошибки и для этих значений имеем

$$\sum_{i=0}^{t}\frac{n!}{i!(n-i)!} = 2048,\ n-k = 12,$$

т.е. относительная скорость (25) также равна 1.

К сожалению, в [53-54] показано, что любой нетривиальный совершенный код имеет параметры кода Хэмминга или кода Голея, т.е. достижение максимальной относительной скорости в системе Нидеррайтера ограничивается только этими конструкциями.

Большинство других кодов, в том числе и коды Гоппы, обладают конструктивными (n,k,d) параметрами, лежащими существенно ниже верхней границы (26) *(реальные дистанционные характеристики кодов Гоппы выше)*. Например, для кода Гоппы с параметрами *n = 1024, k = 524, t = 50* (использован в авторском варианте [15] схемы Мак-Элиса) относительная скорость шифрования (24) в схеме Нидеррайтера равна $R^* \approx 0{,}57$, что несущественно больше по сравнению со скоростью $R \approx 0{,}51$ в схеме МакЭлиса. С увеличением длины кода конструктивные параметры кодов Гоппы ухудшаются, что приводит к снижению скорости R^*. Эту тенденцию наглядно демонстрируют результаты расчетов, представленные в таблице 2, в которой приводятся оценки относительной скорости передачи данных для криптосистем Мак-Элиса и Нидеррайтера при использовании кодов с параметрами из таблицы 1.

Таблица 2 – Относительная скорость передачи данных для различных крипто-кодовых схем с двоичными кодами Гоппы

Кодовые (n,k,d) параметры	(1 024, 524, 101)	(2 048, 1 300, 137)	(4 096, 2 584, 253)	(16 384, 10 322, 867)
Схема Мак-Элиса	≈ 0,51	≈ 0,63	≈ 0,63	≈ 0,63
Схема Нидеррайтера	≈ 0,57	≈ 0,57	≈ 0,53	≈ 0,48
Предлагаемая схема	≈ 0,79	≈ 0,84	≈ 0,83	≈ 0,81

Очевидно, что с увеличением длины кода Гоппы скорость (24), (25) для схемы Нидеррайтера снижается и не превосходит относительной скорости кодирования $R = k/n$. В авторской статье [16] в схеме Нидеррайтера предлагалось использовать обобщенные коды Рида-Соломона, их (n,k,d) параметры связаны соотношением $d = n - k + 1$, т.е. удовлетворяют верхней границе Синлгтона [38 – 40]. Тогда, например, для $q = 1024$ расширенный код Рида-Соломона будет иметь параметры (1 024, 524, 501) и оценка (24) дает относительную скорость для схемы Нидеррайтера $R^* \approx 0{,}66$, что на 30% выше по сравнению с $R \approx 0{,}51$ для схемы Мак-Элиса. Однако в работе [19] предложена эффективная атака на криптосистемы с обобщенными кодами Рида-Соломона, т.е. применение этого класса кодов несостоятельно. Таким образом, стойкие ко всем известным атакам криптосистемы Мак-Элиса и Нидеррайтера на двоичных кодах Гоппы сравнимы по относительной скорости передачи данных. Наибольшая стойкость обеспечивается для относительной скорости передачи данных 1/2 .. 2/3 и этот существенный недостаток частично снимается в предлагаемой ниже криптосистеме.

Предлагаемая криптосистема. По своей сути предлагаемая криптосистема является дальнейшим развитием схемы Мак-Элиса с дополнительным кодированием информационных данных по схеме Нидеррайтера. На рис. 2 схематично изображен процесс криптографического преобразования с использованием кодов:

– в схеме Мак-Элиса информация размещается в кодовом слове замаскированного кода. Зашифрование состоит в добавлении случайного вектора ошибок, который можно интерпретировать как сеансовый (одноразовый) ключ. Расшифрование состоит в декодировании кодового слова, т.е. снятия действия случайного вектора ошибок с кодового слова, содержащего информационную последовательность;

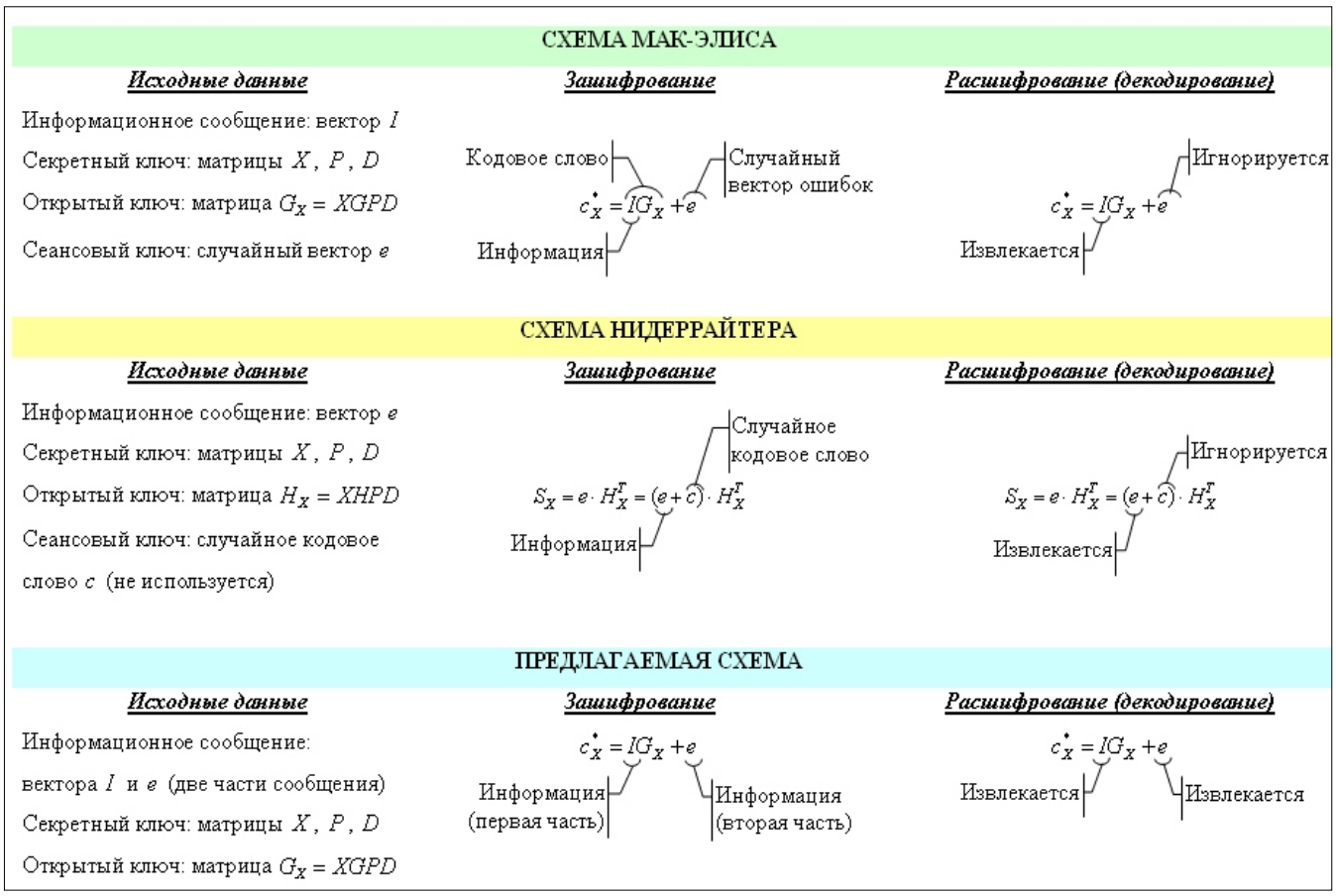

Рис. 2 – Криптографическое преобразование с использованием кодов

- в схеме Нидеррайтера информация размещается в векторе ошибок (посредством равновесного кодирования). По сформированному вектору ошибок вычисляется синдромная последовательность замаскированного кода, которая не зависит от кодового слова, т.е. к исходному вектору ошибок может быть добавлено произвольное кодовое слово (по умолчанию считается нулевым). Синдромный вектор позволяет однозначно декодировать это слово на приемной стороне, только теперь информация извлекается не из кодового слова, а из вектора ошибок;
- в предлагаемой схеме информационная последовательность разбивается на две части. Первая часть помещается в кодовое слово, вторая – в вектор ошибок (посредством равновесного кодирования). Для повышения стойкости эти две части могут быть дополнительно преобразованы (перемешаны, зашифрованы и т.д.). Дальше все преобразования выполняются как в схеме Мак-Элиса. Но на приемной стороне информация извлекается как из кодового слова (1-я часть), так и из вектора ошибок (2-я часть).

Таким образом, предлагаемая схема объединяет способы преобразования информационных данных схем Мак-Элиса и Нидеррайтера, что позволяет существенно повысить относительную скорость передачи данных. Зашифрование осуществляется по правилу (19), где I – первая информационная часть сообщения (как в схеме Мак-Элиса) и e – вторая часть информационного сообщения (как в схеме Нидеррайтера). Если предположить, что $w_h(e) = t$, тогда относительная скорость будет определяться выражением:

$$R^{**} = \frac{k + \left\lfloor \log_q \left((q-1)^t \frac{n!}{t!(n-t)!} \right) \right\rfloor}{n}, \qquad (28)$$

где в числителе первое слагаемое соответствует первой части информационных данных I, а второе слагаемое – второй части e.

Для случая $0 \le w(e) \le t$ выражение (28) перепишется в виде

$$R^{**} = \frac{k + \left\lfloor \log_q \left(\sum_{i=0}^{t} (q-1)^i \frac{n!}{i!(n-i)!} \right) \right\rfloor}{n}. \quad (29)$$

Для случая использования совершенных кодов выражение (29), как и (25), достигает максимума. Действительно, подставив (27) в (29) получим:

$$R^{**} = \frac{k+n-k}{n} = 1.$$

Кроме того, предлагаемая схема за счет информационного кодирования вектора e позволяет повысить относительную скорость и для несовершенных кодов. В качестве примера в таблице 2 приведены оценки относительной скорости передачи информации при использовании различных двоичных кодов Гоппы. Очевидно, что использование предлагаемой схемы шифрования увеличивает относительную скорость передачи данных на 30-40% по сравнению с лучшим показателем среди схем Мак-Элиса и Нидеррайтера.

4 Выводы

Несимметричные криптосистемы на основе алгебраических блоковых кодов были предложены около 40 лет назад и воспринимались тогда большинством исследователей как некое экзотическое и малоприменимое направление в криптографии. Очевидные недостатки (огромные объемы ключевых данных и снижение относительной скорости передачи) в течение длительного времени сдерживали их дальнейшее развитие и практическое использование. И только в последние годы, когда стало понятно, что многие существующие, стандартизированные и широко используемые на практике криптоалгоритмы могут оказаться беззащитными против атак квантового криптоанализа, кодовые криптосистемы получили заслуженное внимание исследователей. Декодирование случайного кода – чрезвычайно сложная вычислительная задача и переборный поиск при ее решении – вероятно лучшее из известных на сегодняшний день решение. Квантовые алгоритмы ускоряют этот процесс, что снижает временны*е* затраты криптоанализа, но это снижение не является критичным (примерно в два раза уменьшается эквивалентная длина ключа).

Фактически следует признать, что кодовые криптосистемы являются реальной альтернативой современных несимметричных криптосистем (RSA, ECC, или других) в части построения надежных постквантовых алгоритмов. Приведенные в таблице 1 расчеты наглядно подтверждают этот вывод. Кроме того, особенности построения кодовых схем позволяют одновременно с криптозащитой реализовать дополнительную услугу контроля возникающих ошибок, что, безусловно, представляет интерес для телекоммуникационных систем специального назначения.

Для практического применения кодовых криптосистем необходимо решить (или смирится с их существованием) несколько конструктивных проблем. Первая, и наиболее очевидная, - огромные объемы ключевых данных. В связи с возможностью использования квантовых вычислительных систем эти объемы придется значительно увеличить (примерно в четыре раза). Так например, для рассмотренных вариантов (таблица 1), объемы ключей достигают сотен мегабит и пока не представляется возможным их уменьшить без снижения стойкости криптосистемы. Ключи в кодовых схемах – это генераторные (порождающие и/или проверочные) матрицы линейного кода, которые должны выглядеть для злоумышленника как случайный набор линейнонезависимых векторов. Сжать или каким-то образом уменьшить этот набор не представляется возможным.

Вторая проблема – низкая относительная скорость передачи данных – в данной работе частично решена. Предложена новая схема шифрования, которая, фактически, объединяет известные способы кодирования информационных данных *(применяются в схемах Мак-Элиса и Нидеррайтера)*. В результате относительная скорость увеличивается, что повышает эффек-

тивность криптосистемы в целом. Если использовать эффективные *(в смысле дистанционных свойств)* коды, относительная скорость будет близка к 100%.

Даже для конструкций, которые лежат значительно ниже верхних кодовых границ наблюдается существенное (на 30-40%) повышение относительной скорости передачи данных (см. таблицу 2). По мнению авторов, это улучшение позволяет приступить к разработке конкретных протоколов криптографической защиты с использованием кодовых схем и начать их практическое внедрение.

Ссылки

[1] Alfred J. Menezes, Paul C. van Oorschot, Scott A. Vanstone. Handbook of Applied Cryptography – CRC Press, 1997. – 794 p.
[2] Niels Ferguson and Bruce Schneier. Practical Cryptography. – John Wiley & Sons, 2003. – 432 pp.
[3] Arto Salomaa. Public-Key Cryptography, Second, Enlarged Edition. – Springer-Verlag, Berlin, Heidelberg, New York, 1996. – x+271 pp.
[4] Nigel Smart. Cryptography: An Introduction (3rd Edition). – 432 pp. https://www.cs.umd.edu/~waa/414-F11/IntroToCrypto.pdf
[5] David Deutsch and Richard Jozsa. Rapid solutions of problems by quantum computation. // Proceedings of The Royal Society of London A: Mathematical, Physical and Engineering Sciences, vol. 439, no. 1907. – 1992. – P. 553-558.
[6] Cleve R., Ekert A., Macchiavello C., Mosca M. Quantum algorithms revisited // Proceedings of The Royal Society of London A: Mathematical, Physical and Engineering Sciences, vol. 454, no. 1969. – 1998. – P. 339-354.
[7] Simon D. R. On the power of quantum computation // Foundations of Computer Science, 1994 Proceedings., 35th Annual Symposium. – P. 116-123.
[8] Grover L. A fast quantum mechanical algorithm for database search. // Proceedings of the 28th annual ACM symposium on the theory of computing (STOC, 96). ACM Press, New York. – 1996. – P. 212–219.
[9] Grover L. A framework for fast quantum mechanical algorithms. // Proceedings of the 13th annual ACM symposium on theory of computing (STOC' 98). ACM Press, New York. – 1998. – P. 53–62.
[10] Shor P. W. Algorithms for quantum computation: discrete logarithms and factoring // Foundations of Computer Science : Conference Publications. – 1994. – P. 124-134.
[11] Shor P. W. Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer // Foundations of Computer Science: Conference Publications. – 1997. – P. 1484-1509.
[12] Neal Koblitz and Alfred J. Menezes. A Riddle Wrapped in an Enigma. https://eprint.iacr.org/2015/1018.pdf.
[13] Committee on National Security Systems, Use of public standards for the secure sharing of information among national security systems, Advisory Memorandum 02-15, July 2015. https://cryptome.org/2015/08/CNSS_Advisory_Memo_02-15.pdf.
[14] Bernstein, Daniel J., Buchmann, Johannes, and Dahmen, Erik. Post-Quantum Cryptography. – 2009, Springer-Verlag, Berlin-Heidleberg. – 245 p.
[15] McEliece R. J. A public-key cryptosystem based on algebraic coding theory. DSN Progress Report 42-44, Jet Propulsion Lab., Pasadena, CA, January-February, 1978. P. 114-116.
[16] Niederreiter H. Knapsack-type cryptosystems and algebraic coding theory // Problem Control and Inform Theory, 1986, v. 15. P. 19-34.
[17] T. R. N. Rao and K. H. Nam. Private-key algebraic-coded cryptosystem. Advances in Cryptology – CRYPTO 86, New York. – NY: Springer. – pp. 35–48.
[18] Yu. V. Stasev, A. A. Kuznetsov. Asymmetric Code-Theoretical Schemes Constructed with the Use of Algebraic Geometric Codes // Cybernetics and Systems Analysis, Volume 41, Issue 3, May 2005, Pages 354 – 363.
[19] Sidel'nikov V.M. Kriptografiya i teoriya kodirovaniya. Materialy konferentsii «Moskovskii universitet i razvitie kriptografii v Rossii», MGU. – 2002. – 22 s.
[20] Sidel'nikov V.M., Shestakov S.O. O sisteme shifrovaniya, postroennoi na osnove obobshchennykh kodov Rida-Solomona. // Diskretnaya matematika. – 1992. – T.4.№3. – S. 57-63.
[21] Kuznetsov A.A. Algebraicheskaya teoriya blokovykh kodov i ee prilozheniya v kriptografii // Persha mizhnarodni naukova konferencija 25–27 travnja 2005r. „Teorija ta metody obrobky sygnaliv". Tezy dopovidej. – K.: NAU. – 2005. – S. 6-8.
[22] Kuznetsov A.A. Issledovanie effektivnosti kriptosistem na algebraicheskikh blokovykh kodakh // Systemy obrobky informacii'. – Kharkiv: KhUPS. –2005 – Vyp. 4. – S. 202–206.
[23] Kuznetsov A.A. Issledovanie pomekhoustoichivosti i kriptostoikosti teoretiko-kodovykh skhem. // Modeljuvannja ta informacijni tehnologii'. – Kyi'v: NANU. – 2005. – №33. – S. 81-84.
[24] Fisher Jean-Dernard, Jacques Stern. An efficient Pseudo-Random Generator Provably as Secure as Syndrome Decoding / Jean-Dernard Fisher, Stern Jacques // EUROCRYPT'96 Proceeding, LNCS 1070. P. 245-255.
[25] Kuznetsov A.A., Korolev R.V., Ryabukha Yu.N. Usovershenstvovannyi metod bystrogo formirovaniya posledovatel'nostei psevdosluchainykh chisel // Zbirnyk naukovyh prac' HUPS. – Kharkiv: KhUPS. – 2008. – Vyp. 3 (18). – S. 101-104.
[26] Gaborit, P., Laudaroux, C., and Sendrier, N.: Synd: a very fast code-based cipher stream with a security reduction. In IEEE Conference, ISIT'07, pages 186–190.
[27] Kirill Morozov, Tsuyoshi Takagi. Zero-Knowledge Protocols for the McEliece Encryption // Information Security and Privacy Volume 7372 of the series Lecture Notes in Computer Science pp 180-193.
[28] Bernstein D. Post-quantum cryptography [Text] / D. Bernstein, J. Buchmann, E. Dahmen. – Berlin: Springer, 2009. – 246 p.
[29] Courtois, N., Finiasz M., Sendrier N.: How to achieve a McEliece-based digital signature scheme. In Advances in Cryptology - ASIACRYPT 2001, volume 2248, pages 157–174.
[30] Finiasz M.: Parallel-CFS: Strengthening the CFS McEliece-based signature scheme. In Biryukov A., Gong G., Stinson D., eds.: Selected Areas in Cryptography. Volume 6544 of LNCS., Springer (2010) pp. 159-170.
[31] Stern J.: A new identification scheme based on syndrome decoding. In Advances in Cryptology - CRYPTO'93, volume 773 of LNCS. Springer Verlag (1994).

[32] Veron P.: Improved identification schemes based on error-correcting codes. Appl. Algebra Eng. Commun. Comput., 8(1):57–69 (1996).
[33] Anne Canteaut and Nicolas Sendrier. Cryptanalysis of the original McEliece cryptosystem. In Kazuo Ohta and Dingyi Pei, editors, Advances in cryptology — ASIACRYPT'98, volume 1514 of Lecture Notes in Computer Science, pages 187– 199.
[34] Vladimir M. Sidelnikov and Sergey O. Shestakov. On insecurity of cryptosystems based on generalized Reed-Solomon codes. Discrete Mathematics and Applications, 1992. – 439–444.
[35] Minder L., Shokrollahi A. Cryptanalysis of the Sidelnikov Cryptosystem // Advances in Cryptology — EUROCRYPT 2007: 26th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Barcelona, Spain, May 20-24, 2007. Proceedings — Springer Berlin Heidelberg, 2007. — P. 347—360.
[36] Daniel J. Bernstein and Tanja Lange and Christiane Peters. Attacking and defending the McEliece cryptosystem. https://cr.yp.to/codes/mceliece-20080807.pdf.
[37] E. Berlekamp, R. McEliece, H. van Tilborg. On the Inherent Intractability of Certain Coding Problems. // IEEE Transactions on Information Theory, vol. IT-24, No. 3, May 1978. – P. 384-386.
[38] Clark G.C., Cain J.B. Error-Correction Coding for Digital Communications. – Springer, 1981, - 432 p.
[39] Blahut R. E. Theory and Practice of Error Control Codes. – Addison Wesley Publishing Company, Inc., Reading, Massachusetts, 1983, 1983, – 500 pp.
[40] F. J. MacWilliams and N. J. A. Sloane. The theory of error-correcting codes. – North-Holland, Amsterdam, New York, Oxford, 1977, – 762 pp.
[41] Claude E. Shannon. Communication in the Presence of Noise. Proceedings of the IRE, vol. 37, no. 1, pp. 10–21, Jan. 1949.
[42] V. D. Goppa. Novyi klass lineinykh korrektiruyushchikh kodov // Probl. peredachi inform., 1970, tom 6, vypusk 3, S. 24–30.
[43] V. D. Goppa. Na neprivodimykh kodakh dostigaetsya propusknaya sposobnost' DSK. // Probl. peredachi inform., 1974, tom 10, vypusk 1, S. 111–112.
[44] National Institute of Standards and Technology, "FIPS-197: Advanced Encryption Standard", November 2001: http://csrc.nist.gov/publications/ fips/fips197/fips-197.pdf.
[45] A New Encryption Standard of Ukraine: The Kalyna Block Cipher. https://eprint.iacr.org/2015/650.pdf.
[46] Raphael Overbeck, Nicolas Sendrier, Code-based cryptography. In: Daniel J. Bernstein, et al. (eds). First International Workshop on Post-quantum Cryptography, PQ Crypto 2006, Leuven, The Netherland, May 23-26, 2006. Selected papers, pp. 95-145.
[47] D. J. Bernstein. Grover vs. McEliece. In N. Sendrier, editor, Post-Quantum Cryptography, Third International Workshop, PQCrypto 2010, Darmstadt, Germany, May 25-28, 2010. Proceedings, volume 6061 of Lecture Notes in Computer Science, pages 73–80. Springer, 2010.
[48] A. Menezes, P. van Oorschot, S. Vanstone. Chapter 14. Efficient Implementation // Handbook of Applied Cryptography. – CRC-Press, 1996. – 816 p.
[49] Kerry Maletsky. RSA vs ECC Comparison for Embedded Systems. White Paper. Atmel Corporation – 2015, 5p. http://www.atmel.com/images/atmel-8951-cryptoauth-rsa-ecc-comparison-embedded-systems-whitepaper.pdf.
[50] John Proos and Christof Zalka. Shor's discrete logarithm quantum algorithm for elliptic curves. arxiv.quant-ph/0301141 v2, 2004.
[51] Ziatdinov M. Using frequency analysis and Grover's algorithm to implement known ciphertext attack on symmetric ciphers // Lobachevskii Journal of Mathematics 2013 vol.34 N4, pages 313-315.
[52] Metod nedvijkovogo rivnovagovogo koduvannja / V. B. Dudykevych, O. O. Kuznjecov, B. P. Tomashevs'kyj // Suchasnyj zahyst informacii'. - 2010. - № 3. - S. 57-68. - Rezhym dostupu: http://nbuv.gov.ua/UJRN/szi_2010_3_10.
[53] Tietavainen A., Perko A. There are no unknown perfect binary codes. - Annales Universitatis Turkuensis. - Ser. A, I 148, 3-10[6], 1971.
[54] Lint van J. H. Nonexistence theorems for perfect error-correcting codes. - Computers in Algebra and Number Theory. - Vol. IV [6], 1971.

Reviewer: Roman Oliynikov, Doctor of Sciences (Engineering), Full Professor, V. N. Karazin Kharkiv National University, Kharkiv, Ukraine.
E-mail: roliynykov@gmail.com

Received: August 2016.

Authors:
Alexandr Kuznetsov, Doctor of Sciences (Engineering), Full Prof., V.N. Karazin Kharkiv National University, Kharkiv, Ukraine.
E-mail: kuznetsov@karazin.ua
Andriy Pushkar'ov, Director of department, State Service of Special Communication and Information Protection of Ukraine, Kyiv, Ukraine.
Sergii Kavun, Doctor of Sciences (Economics), Ph.D. (Engineering), Full Prof., Department of Information Technologies, Kharkiv Educational and Research Institute of the University of Banking, Kharkiv, Ukraine.
E-mail: kavserg@gmail.com
Vyacheslav Kalashnikov, Doctor of Sciences (Physics and Mathematics), Full Prof., Department of Systems and Industrial Engineering, Campus Monterrey, Monterrey, Mexico. E-mail: kalash@itesm.mx

Code-Based Public-Key Cryptosystems: the current state, the existing contradictions and prospects of practical use for the post-quantum period.

Abstract. Discusses the asymmetric cryptosystem based on algebraic coding, are investigated with temporary status, controversies and prospects of practical use for the post-quantum period. We propose a new scheme of crypto-transformation, significantly increases the relative rate of information transmission. Shown the advantage of the proposed design compared to the already known schemes Mac-Elisa and Niederreiter.

Keywords: Public-Key Cryptosystems, Post-Quantum Cryptography, Code-Based Cryptography.

Рецензент: Роман Олійников, д.т.н., проф., Харківський національний університет імені В.Н. Каразіна, Харків, Україна.
E-mail: roliynykov@gmail.com

Надійшло: Серпень 2016.

Автори:
Олександр Кузнецов, д.т.н., проф., академік Академії наук прикладної радіоелектроніки, Харківський національний університет імені В.Н. Каразіна, Харків, Україна.
E-mail: kuznetsov@karazin.ua
Андрій Пушкарьов, директор департаменту Державної служби спеціального зв'язку та захисту інформації України, Київ, Україна.
Сергій Кавун, доктор економічних наук, к.т.н., проф., Харківський навчально-науковий інститут ДВНЗ "Університет банківської справи", Харків, Україна.
E-mail: kavserg@gmail.com
В'ячеслав Калашников, д.ф.-м.н., проф., департамент систем і промислового виробництва Технологічного університету Монтеррея, Монтеррей, Мексика.
E-mail: kalash@itesm.mx

Несиметричні криптосистеми на основі алгебраїчного кодування: сучасний стан, наявні суперечності і перспективи практичного використання на пост-квантовий період.

Анотація. Розглядаються несиметричні криптосистеми на основі алгебраїчного кодування, досліджуються сучасний стан, наявні суперечності і перспективи практичного використання на пост-квантовий період. Пропонується нова схема криптоперетворення, яка істотно підвищує відносну швидкість передачі інформації. Показано перевагу запропонованої конструкції в порівнянні з вже відомими схемами Мак-Еліса та Нидеррайтера.

Ключові слова: несиметричні криптосистеми, пост-квантова криптографія, криптографія на основі кодів.