УДК 621.391:519.2 **DOI:10.30837/rt.2020.1.200.03**

О.С. ШЕВЧУК

РАНДОМІЗОВАНА СИМЕТРИЧНА КРИПТОСИСТЕМА МАК-ЕЛІСА НА ОСНОВІ УЗАГАЛЬНЕНИХ КОДІВ РІДА - СОЛОМОНА

Вступ

Однією з актуальних проблем сучасної криптографії є створення практичних постквантових криптосистем, стійкість яких базується на складності розв'язання єдиної обчислювально складної задачі, аналогічно тому як стійкість криптосистеми RSA базується на складності факторизації цілих чисел. Перспективний клас таких криптосистем утворюють кодові криптосистеми, найпершою асиметричною з яких є криптосистема Мак-Еліса [1].

На сьогодні відомо декілька конструкцій симетричних кодових криптосистем, аналогічних криптосистемі Мак-Еліса, наприклад схеми шифрування Жордана [2], Рао [3], Рао - Нама [4] та низка вдосконалень останньої криптосистеми [5, 6], однак жодна з них не задовольняє цілком сучасним вимогам щодо стійкості та практичності одночасно.

В [7] запропоновано рандомізовану криптосистему Мак-Еліса, яка відрізняється від оригіналу використанням випадкових секретних даних при зашифруванні. Показано, що, на відміну від оригінальної, рандомізована криптосистема Мак-Еліса за певних умов є обґрунтовано стійкою відносно атак з підібраним відкритим тестом (CPA-стійкою). Іншим прикладом симетричної кодової криптосистеми, для якої відомо обґрунтування стійкості (security proof), є LPN-C [8], проте задача зменшення довжини ключа або підвищення швидкості передачі інформації при збереженні рівня стійкості цієї криптосистеми залишається актуальною.

Дану роботу присвячено створенню та дослідженню симетричної версії криптосистеми з [7], що будується на основі узагальнених кодів Ріда - Соломона (УРС). Вибір цих кодів зумовлено, перш за все, тим, що вони існують для всіх природних значень параметрів (довжини та вимірності коду) і є максимально дистанційно роздільними (МДР), що дозволяє в широких межах змінювати характеристики відповідних криптосистем. Крім того, для зазначених кодів відомі дуже швидкі алгоритми декодування (до половини кодової відстані та, навіть, за її межами) [9]. Нарешті, асиметричні криптосистеми, побудовані на основі кодів УРС [10, 11], є нестійкими, оскільки для них існують ефективні алгоритми відновлення секретних ключів за відкритими [12, 13]. Це викликає додатковий інтерес до стійкості відповідних симетричних криптосистем.

В роботі отримано оцінки стійкості зазначених криптосистем відносно природної атаки з підібраним відкритим текстом та запропоновано алгоритм вибору параметрів для побудови цих криптосистем. Проведено порівняння досліджених криптосистем з відповідними криптосистемами типу LPN-C і показано, що перші мають помітно меншу довжину ключа при заданій стійкості в порівнянні з останніми.

1. Означення основних понять та уточнення постановки задачі

Нехай F – поле з 2^s елементів, $s \geq 2$. Для матриці M над полем F позначимо $< M >$ лінійний код над цим полем, породжений рядками матриці M. Нехай k, n – натуральні числа, $k \leq n \leq 2^s$, $\alpha = (\alpha_1, \ldots, \alpha_n) \in F^n$, $\beta = (\beta_1, \ldots, \beta_n) \in (F \setminus \{0\})^n$, де α_i є попарно різними, $i \in \overline{1,n}$.

Узагальнений код Ріда - Соломона $GRS_{n,k}(\alpha, \beta)$ визначається як лінійний код над полем F з твірною матрицею

$$G_{\alpha,\beta} = \begin{pmatrix} \beta_1 & \beta_2 & \dots & \beta_n \\ \alpha_1\beta_1 & \alpha_2\beta_2 & \dots & \alpha_n\beta_n \\ \dots & \dots & \dots & \dots \\ \alpha_1^{k-1}\beta_1 & \alpha_2^{k-1}\beta_2 & \dots & \alpha_n^{k-1}\beta_n \end{pmatrix}, \quad (1)$$

тобто $GRS_{n,k}(\alpha,\beta) = < G_{\alpha,\beta} >$. Відомо [14], що цей код є *максимально дистанційно розділь-ним*, тобто має найбільшу для заданих n і k мінімальну відстань $d = n-k+1$, а також найбільшу дуальну відстань $d^{\perp} = k+1$. Код УРС дозволяє виправляти будь-яку кількість $t \le \lfloor 1/2 \cdot (d-1) \rfloor$ (адитивних) помилок, використовуючи $O(2^s s^2)$ арифметичних операцій в полі F [9].

Рандомізована симетрична криптосистема Мак-Еліса з параметрами $l, k, n \in \mathbf{N}$, $\varepsilon \in (0,1)$, де $l \le k \le n \le 2^s$, що будується на основі кодів УРС, визначається таким чином.

Секретними ключами у криптосистемі є пари (α,β), де $\alpha, \beta \in F^n$, координати вектора α є попарно різними, а координати вектора β – ненульовими елементами поля F, $i \in \overline{1,n}$.

Для зашифрування відкритого тексту $m \in F^l$ на ключі (α,β) відправник генерує незалежні випадкові вектори r та $\xi = (\xi_1, \dots, \xi_n)$, де вектор r має рівномірний розподіл ймовірностей на множині F^{k-l}, а координати вектора ξ є незалежними випадковими величинами, що розподілені за законом

$$\mathbf{P}(\xi_i = 0) = 2^{-s}(1+(2^s-1)\varepsilon), \ \mathbf{P}(\xi_i = a) = 2^{-s}(1-\varepsilon), \ a \in F \setminus \{0\}, \ i \in \overline{1,n}. \quad (2)$$

Далі відправник обчислює шифротекст за формулою

$$E(m) = (r, m)G_{\alpha,\beta} \oplus \xi. \quad (3)$$

де матриця $G_{\alpha,\beta}$ визначається за формулою (1).

Для розшифрування повідомлення (3) на ключі (α,β) отримувач обчислює вектор $(r, m)G_{\alpha,\beta}$, використовуючи алгоритм декодування коду $GRS_{n,k}(\alpha,\beta)$ зі складністю $O(2^s s^2)$ операцій [9]. Нарешті, на підставі лінійної незалежності рядків матриці (1) отримувач може відновити вектор (r, m), наприклад, за допомогою алгоритму Гаусса.

Зауважимо, що у випадку, коли число ненульових координат вектора ξ є більше ніж $\lfloor 1/2 \cdot (d-1) \rfloor$, можлива помилка розшифрування, ймовірність якої, як показано далі, можна зробити достатньо малою шляхом вибору параметра ε.

Зауважимо також, що означена вище криптосистема є симетричним аналогом окремого випадку рандомізованої (асиметричної) криптосистеми Мак-Еліса, описаної в [7]. В цій криптосистемі замість коду УРС використовується довільний двійковий лінійний код із швидким алгоритмом декодування (на рівні половини мінімальної відстані). У криптосистемах з відкритим ключем, зокрема рандомізованих, побудованих на основі кодів УРС, матриця (1) відіграє роль частини секретного ключа. При цьому відкритим ключем є деяка інша твірна матриця того ж самого коду [10] або його випадково обраного підкоду [11]. Вразливість таких криптосистем є наслідком того, що, знаючи код УРС $GRS_{n,k}(\alpha,\beta)$ (або навіть його підкод відносно невеликої вимірності), можна відновити вектори α, β за поліноміальний від n час [12, 13]. Для симетричної криптосистеми з рівнянням шифрування (3) код $GRS_{n,k}(\alpha,\beta)$ є невідомим, тому що на цю криптосистему є незастосовними атаки, описані в [12, 13].

2. Умова, що забезпечує малість ймовірності помилки розшифрування

Твердження 1. Нехай для зазначених вище $k, n,\ s$ та $\delta \in (0,1)$ виконується умова

$$1-\frac{1}{1-2^{-s}}\left(\frac{1}{2}-\frac{k}{2n}-\sqrt{\frac{\ln(\delta^{-1})}{2n}}\right)\le \varepsilon<1. \tag{4}$$

Тоді ймовірність того, що законний отримувач відновить відкритий текст m за шифротекстом (2), є не менше ніж $1-\delta$.

Доведення. Згідно з означенням криптосистеми, якщо відбувається помилка розшифрування, то число ненульових координат вектора ξ є більше ніж $\lfloor 1/2\cdot(d-1)\rfloor$.

Введемо випадкові величини $\eta_1,\dots,\eta_n$, де $\eta_i=1$, якщо $\xi_i\neq 0$; $\eta_i=0$ – у протилежному випадку. Тоді $\eta_1,\dots,\eta_n$ є незалежними в сукупності та розподілені за законом $\mathbf{P}(\eta_i=1)=1-\mathbf{P}(\eta_i=0)=(1-2^{-s})(1-\varepsilon)$, $i\in\overline{1,n}$. При цьому ймовірність помилки розшифрування не перевищує ймовірності події $\{\eta_1+\dots+\eta_n\ge 1/2\cdot(n-k)\}$, яку можна оцінити за допомогою нерівності Гефдінга [15].

Дійсно, справедливі такі співвідношення:

$$\mathbf{P}(\eta_1+\dots+\eta_n\ge 1/2\cdot(n-k))=\mathbf{P}\left(\sum_{i=1}^{n}\xi_i-\sum_{i=1}^{n}\mathbf{E}\xi_i\ge 1/2\cdot(n-k)-\sum_{i=1}^{n}\mathbf{E}\xi_i\right)=$$

$$=\mathbf{P}\left(\sum_{i=1}^{n}\xi_i-n(1-2^{-s})(1-\varepsilon)\ge 1/2\cdot(n-k)-n(1-2^{-s})(1-\varepsilon)\right)\le$$

$$\le\exp\left\{-2n\left(\frac{1}{2}-\frac{k}{2n}-(1-2^{-s})(1-\varepsilon)\right)^2\right\}\le\delta,$$

де передостання нерівність випливає з нерівності Гефдінга, а остання – з формули (4).

Твердження доведено.

Надалі вважатимемо, що параметри k, n, ε криптосистеми задовольняють умові (4).

3. Атака на криптосистему на основі підібраного відкритого тексту

Припустимо, що супротивник має доступ до оракула зашифрування E вигляду (3) та подає на його вхід t разів відкритий текст $m=0$. В результаті супротивник отримує систему рівнянь

$$r_jG^{(1)}_{\alpha,\beta}\oplus\xi_j=c_j,\ \ j\in\overline{1,t}, \tag{5}$$

де $G^{(1)}_{\alpha,\beta}$ – підматриця, що складається з перших $k-l$ рядків матриці (1), $c_1,\dots,c_t$ – відомі шифротексти, а $r_1,\dots,r_t$ та $\xi_1,\dots,\xi_t$ – незалежні випадкові вектори, де $r_1,\dots,r_t$ мають рівномірний розподіл на множині F^{k-l}, а координати векторів $\xi_1,\dots,\xi_t$ є незалежними випадковими величинами, розподіленими за законом (2).

З формул (1) і (5) випливає, що $c_1,\dots,c_t$ є словами невідомого коду УРС $GRS_{n,k-l}(\alpha,\beta)$, спотвореними у 2^s-му симетричному каналі зв'язку. Отже, відновивши цей код за набором його спотворених слів, можна знайти вектори α,β за поліноміальний від n час за допомогою алгоритму з [12]. Оскільки $GRS_{n,k-l}(\alpha,\beta)$ є кодом МДР, він є систематичним і для його відновлення можна скористатися методом, викладеним в [16].

Позначимо $G=(I_{k-l},\,X)$ невідому канонічну твірну матрицю коду $GRS_{n,k-l}(\alpha,\beta)$, де I_{k-l} – одинична матриця порядку $k-l$, X – матриця розміру $(k-l)\times(n-k)$. Існує оборотна матриця U над полем F така, що $G^{(1)}_{\alpha,\beta}=UG$. Підставляючи зазначену рівність у формулу (5) та позначаючи $v_j=r_jU$, отримаємо систему рівнянь вигляду

$$v_j(I_{k-l},\,X)\oplus\xi_j=c_j,\ j\in\overline{1,t}, \tag{6}$$

де $v_1,\dots,v_t$ є незалежними випадковими векторами з рівномірним розподілом на множині F^{k-l}.

Слідуючи методу [16], перетворимо систему рівнянь (6) наступним чином. Позначимо $c_j^{(1)}$ та $c_j^{(2)}$ підвектори вектора c_j, що складаються з його перших $k-l$ та останніх $n-(k-l)$ координат відповідно. Аналогічні позначення введемо для випадкового вектора ξ_j, $j\in\overline{1,t}$. Покладемо $A_j=v_j\oplus\xi_j^{(1)}$, $\zeta_j=\xi_j^{(1)}X\oplus\xi_j^{(2)}$, $j\in\overline{1,t}$. Рівність (6) рівносильна співвідношенням $(c_j^{(1)},\,c_j^{(2)})=(v_j\oplus\xi_j^{(1)},\ v_jX\oplus\xi_j^{(2)})$, $j\in\overline{1,t}$, які можуть бути записані у вигляді:

$$A_j=c_j^{(1)},\ A_jX\oplus\zeta_j=c_j^{(2)},\ j\in\overline{1,t}. \tag{7}$$

При цьому на підставі зазначених вище припущень щодо розподілів випадкових векторів $r_1,\dots,r_t$ та $\xi_1,\dots,\xi_t$ вектори $A_1,\dots,A_t$ є незалежними в сукупності та мають рівномірний розподіл на множині F^{k-l}, а вектори $\zeta_1,\dots,\zeta_t$ є незалежними в сукупності та не залежать від $A_1,\dots,A_t$.

Нарешті, як і в [16] позначимо A матрицю, що складається з рядків $A_1,\dots,A_t$, x_i – i-й стовпець матриці X; покладемо $b^{(i)}=(c_{1,i}^{(2)},\dots,c_{t,i}^{(2)})^{\mathrm{T}}$, $\zeta^{(i)}=(\zeta_{1,i},\dots,\zeta_{t,i})^{\mathrm{T}}$, де $c_{j,i}^{(2)}$ та $\zeta_{j,i}$ – i-ті координати векторів $c_j^{(2)}$ та ζ_j відповідно, $j\in\overline{1,t}$, $i\in\overline{1,n-(k-l)}$. На підставі рівностей (7) вектор x_i співпадає з істинним розв'язком $x_i^{(0)}$ системи лінійних рівнянь зі спотвореними правими частинами

$$Ax=b^{(i)}=Ax_i^{(0)}\oplus\zeta^{(i)}, \tag{8}$$

над полем F, де матриця A і вектор $b^{(i)}$ визначаються безпосередньо за набором слів $c_1,\dots,c_t$:

$$A_j=c_j^{(1)},\ b^{(i)}=(c_{1,i}^{(2)},\dots,c_{t,i}^{(2)})^{\mathrm{T}},\ j\in\overline{1,t},\ i\in\overline{1,n-(k-l)}.$$

Отже, атака на криптосистему, що розглядається, полягає у побудові для кожного $i\in\overline{1,n-(k-l)}$ системи лінійних рівнянь зі спотвореними правими частинами вигляду (8) та її розв'язанні за допомогою відомих методів. Відновивши істинні розв'язки зазначених систем рівнянь, знайдемо канонічну твірну матрицю $G=(I_{k-l},\,X)$ коду $GRS_{n,k-l}(\alpha,\beta)$, за якою відновимо вектори α,β, використовуючи алгоритм, описаний в [12].

4. Оцінка ефективності наведеної атаки

Для того, щоб оцінити трудомісткість атаки та обсяг матеріалу t, потрібного для її надійного виконання, треба спочатку знайти розподіл ймовірностей спотворень у правих частинах рівнянь системи (8).

Для будь-якого $\varepsilon \in (0,1)$ назвемо ε-*нерівномірним розподілом* на полі F розподіл ймовірностей вигляду (2).

Наступна лема випливає безпосередньо з наведеного означення та формули повної ймовірності.

Лема 1. Нехай ξ та η є незалежними випадковими величинами на полі F, що мають ε_1-нерівномірний та ε_2-нерівномірний розподіли ймовірностей відповідно. Тоді для будь-якого $c \in F \setminus \{0\}$ випадкова величина $c\xi$ має ε_1-нерівномірний, а випадкова величина $\xi + \eta$ має $\varepsilon_1\varepsilon_2$-нерівномірний розподіл ймовірностей на полі F.

Твердження 2. Для будь-якого $i \in \overline{1, n-(k-l)}$ координати випадкового вектора $\zeta^{(i)}$ у правій частині системи рівнянь (8) мають ε^{k-l+1}-нерівномірний розподіл ймовірностей, де ε визначається згідно з умовою (2).

Доведення. За означенням j-та координата випадкового вектора $\zeta^{(i)}$ має вигляд $\zeta_{j,i} = \xi_j^{(1)} x_i \oplus \xi_j^{(2)}$. При цьому всі координати вектора x_i є ненульовими елементами поля F. Дійсно, оскільки (I_{k-l}, X) – твірна матриця коду МДР $GRS_{n,k-l}(\alpha, \beta)$, то припущення про те, що деяка координата вектора x_i дорівнює нулю, тягне за собою існування слова над полем F, яке анулює зазначений код та має вагу, меншу ніж $k-l+1$, що протирічить тому, що $k-l+1$ є дуальною відстанню цього коду.

Таким чином, $\zeta_{j,i}$ є сумою точно $k-l+1$ незалежних випадкових величин, кожна з яких, згідно з умовою (2), має ε-нерівномірний розподіл ймовірностей на полі F. Отже, на підставі леми 1 $\zeta_{j,i}$ має ε^{k-l+1}-нерівномірний розподіл на цьому полі. Твердження доведено.

Отримаємо оцінку трудомісткості розв'язання системи рівнянь вигляду (8), вважаючи, що для цього використовується один з найефективніших на сьогодні алгоритмів [17].

Зазначений алгоритм залежить від параметрів $k' \geq 2$, що є степенем двійки, та $l' \in \overline{1, k-l}$, які задовольняють умові

$$\log(2\Delta(k')^{-1} l' s \ln 2) \leq s(k-l-l')(\log k')^{-1}, \tag{9}$$

і складається з двох етапів.

На першому етапі за допомогою алгоритму Вагнера (k'-tree algorithm) [18] здійснюється виключення з вхідної системи рівнянь (8) останніх $k-l-l'$ невідомих. В результаті отримується нова система лінійних рівнянь зі спотвореними правими частинами від l' невідомих над полем F, кожне рівняння якої є сумою певних k' рівнянь вхідної системи. На другому етапі отримана система рівнянь розв'язується методом максимальної правдоподібності із застосуванням швидкого перетворення Адамара.

Таким чином, зазначений алгоритм дозволяє відновити перші l' невідомих системи рівнянь (8). Застосовуючи його $\lceil l/l' \rceil$ разів до різних наборів невідомих, що не перетинаються, можна знайти шуканий вектор $x_i^{(0)}$.

В [17] показано, що середня (відносно незалежного випадкового рівноймовірного вибору рядків матриці A) трудомісткість алгоритму розв'язання системи рівнянь (8) визначається за формулою

$$T(k', l') = (m(k', l'))^{\frac{1}{\theta}} k' 2^{\frac{s(k-l-l')}{\theta}} + s(m(k', l') + sl' 2^{sl'}) + 2^{s(l'+1)}, \tag{10}$$

а обсяг матеріалу, потрібного для успішного розв'язання цієї системи, – за формулою

$$t = t(k', l') = k' 2^{\frac{s(k-l-l')}{\theta}} (2l' s \ln 2)^{\frac{1}{\theta}} \Delta(k')^{-\frac{1}{\theta}}, \tag{11}$$

де

$$\theta = 1 + \log k',\ m(k',l') = 2\Delta(k')^{-1} l' s \ln 2, \tag{12}$$

$$\Delta(k') = 2^{-s} \sum_{z \in F} (2^s \mathbf{P}\{\mu_1 + \cdots + \mu_{k'} = z\} - 1)^2,$$

причому $\mu_1, \ldots, \mu_{k'}$ є незалежними випадковими величинами, розподіленими за тим самим законом, що і спотворення у правій частині системи рівнянь (8).

Твердження 3. Справедлива рівність

$$\Delta(k') = (2^s - 1)\varepsilon^{2k'(k-l+1)}. \tag{13}$$

Доведення. На підставі твердження 2 випадкові величини $\mu_1, \ldots, \mu_{k'}$ мають ε^{k-l+1}-нерівномірний розподіл ймовірностей на полі F. Отже, згідно з лемою 1, випадкова величина $\mu_1 + \cdots + \mu_{k'}$ має $\varepsilon^{k'(k-l+1)}$-нерівномірний розподіл ймовірностей. Звідси безпосередньо випливає нерівність (13).

Таким чином, для оцінювання ефективності наведеної атаки можна використовувати наступний алгоритм.

Алгоритм 1.

Вхідні дані:

– верхня межа δ ймовірності помилкового розшифрування відкритого тексту законним отримувачем, $\delta \in (0, 1)$;

– довжина n та вимірність k коду УРС над полем з 2^s елементів, де $s \geq 2$, $k < n - \sqrt{2n \ln(\delta^{-1})}$, $n \leq 2^s$;

– довжина l відкритих текстів у криптосистемі, що розглядається, $l \leq k$.

1. Покласти $\varepsilon = 1 - \dfrac{1}{1 - 2^{-s}} \left(\dfrac{1}{2} - \dfrac{k}{2n} - \sqrt{\dfrac{\ln(\delta^{-1})}{2n}} \right)$.

2. Для будь-яких $l' \in \overline{1, k-l}$ та $k' = 2, 4, 8, 16, \ldots$, що задовольняють умові (9), обчислити значення $T(k', l')$, використовуючи формули (10), (12), (13).

3. Обрати k^* та $l^* \in \overline{1, k-l}$ такі, що $T(k^*, l^*) = \min\{T(k', l')\}$, де мінімум береться за всіма зазначеними вище парами (k', l').

Результат:

– значення k^* та l^* (останнє з яких дорівнює кількості координат довільного фіксованого стовпця канонічної твірної матриці коду УРС, які відновлюються за допомогою атаки);

– середня часова складність атаки $T(k^*, l^*)$;

– обсяг матеріалу $t(k^*, l^*) = k^* 2^{\frac{s(k-l-l^*)}{\theta}} (2l^* s \ln 2)^{\frac{1}{\theta}} \Delta(k^*)^{-\frac{1}{\theta}}$, потрібного для успішної реалізації атаки.

Зауважимо, що нерівності $s \geq 2$ та $k < n - \sqrt{2n \ln(\delta^{-1})}$ тягнуть умову $\varepsilon \in (0, 1)$, яка є необхідною для коректного вибору параметра ε. При цьому значення $T(k^*, l^*)$ визначає нижню межу стійкості криптосистеми відносно розглянутої атаки.

В табл. 1, 2 наведено чисельні значення параметрів та оцінки стійкості деяких криптосистем, що будуються на основі кодів УРС.

Таблиця 1

Результати виконання алгоритму 1 при $\delta = 10^{-8}$, $s = 9$, $n = 512$, $k = 374$ ($\varepsilon = 0{,}9994$)

l	l^*	k^*	$\log T(k^*, l^*)$	$\log t(k^*, l^*)$
1	38	256	362,77	362,76
2	38	256	361,72	361,71
10	37	256	354,28	354,18
50	33	256	316,16	316,15
60	32	256	306,63	306,62
70	31	256	297,10	297,08
100	28	256	268,52	268,48
110	27	256	259,00	258,94
120	26	256	249,49	249,41
200	18	256	174,04	173,12
210	16	256	164,58	164,57
220	15	256	155,04	155,03
245	13	256	130,85	130,69
246	7	512	129,94	129,92
247	7	512	128,94	128,93
365	1	128	18,43	15,63
366	1	64	18,31	14,39
367	1	64	18,25	13,09

Як видно з табл. 1, при $s = 9$, $l = 110$ часова складність розглянутої атаки на криптосистему становить не менше ніж 2^{259}, а обсяг потрібного для реалізації атаки матеріалу складає не менше ніж $2^{258{,}94}$. При цьому швидкість передачі інформації в системі є $l/n = 110/512 = 0{,}2148...$, а довжина ключа складає $2ns = 9216$ біт. В той же час, згідно з табл. 2, майже таку ж саму стійкість відносно розглянутої атаки можна отримати, вважаючи $s = 10$, $l = 600$; при цьому швидкість передачі складає $l/n = 600/1024 = 0{,}5859...$, а довжина ключа збільшується до $2ns = 20480$ біт.

Таблиця 2

Результати виконання алгоритму 1 при $\delta = 10^{-8}$, $s = 10$, $n = 1024$, $k = 827$ ($\varepsilon = 0{,}9986$)

l	l^*	k^*	$\log T(k^*, l^*)$	$\log t(k^*, l^*)$
1	90	256	917,81	917,74
2	90	256	916,67	916,52
10	89	256	907,91	907,85
50	85	256	864,31	863,40
70	82	256	842,28	842,28
100	79	256	808,96	808,94
160	72	256	743,37	743,36
200	68	256	698,93	698,91
260	61	256	633,33	633,33
300	57	256	588,90	588,88
370	49	256	512,18	512,18
420	44	256	456,68	456,61
500	35	256	368,80	368,79
560	29	256	303,01	302,09
600	24	256	258,73	258,72
640	20	256	214,45	214,25
700	13	256	148,63	148,62
720	11	256	126,40	126,37
780	4	256	60,65	60,65
790	3	256	49,49	49,49
800	2	256	38,32	38,31
810	1	256	27,10	27,09
812	1	256	24,71	24,64
813	1	128	23,55	23,41
820	1	64	20,15	13,97

Зауважимо, що в обох випадках ймовірність правильного прийому повідомлень законним отримувачем криптосистеми є не менше ніж $1-10^{-8}$. При цьому застосування швидких алгоритмів кодування та декодування кодів УРС зі складністю $O(2^s s^2)$ операцій в полі F [9]) дозволяє отримувати швидкі програмні реалізації розглянутих криптосистем.

5. Порівняння з криптосистемою LPN-C

Розглянемо окрему версію криптосистеми LPN-C [8], яка будується на основі коду УРС над тим самим полем F, що й досліджена вище (рандомізована симетрична) криптосистема Мак-Еліса.

Зазначена криптосистема LPN-C залежить від параметрів $l_1, l_2, n_1 \in \mathbf{N}$, $\varepsilon \in (0,1)$, де $l_1 \le n_1 \le 2^s$, і визначається за допомогою рівняння шифрування

$$\tilde{E}(m) = (mG_1 \oplus rG_2 \oplus \xi, r), \tag{14}$$

де $m \in F^{l_1}$ – відкритий текст, $\tilde{E}(m)$ – шифрований текст, G_1 – твірна матриця коду УРС довжини n_1 та вимірності l_1 над полем F (яка є загальнодоступною), G_2 – $l_2 \times n_1$-матриця над цим полем, що є секретним ключем, а r та $\xi = (\xi_1, \ldots, \xi_n)$ – незалежні випадкові вектори, де вектор r має рівномірний розподіл ймовірностей на множині F^{l_2}, а координати вектора ξ є незалежними випадковими величинами, розподіленими за законом (2).

Для зашифрування кожного відкритого тексту m вектори r та ξ генеруються незалежно від решти даних так, як це робиться у рандомізованій криптосистемі Мак-Еліса. Для розшифрування повідомлення (14) на ключі G_2 отримувач, знаючи r, обчислює вектор $mG_1 \oplus \xi$, за яким відновлює відкритий текст m, використовуючи швидкий алгоритм декодування коду УРС з відомою твірною матрицею G_1 [11]. Таким чином, на відміну від розглянутої вище криптосистеми Мак-Еліса, де матриця $\begin{pmatrix} G_1 \\ G_2 \end{pmatrix}$ є невідомою твірною матрицею коду УРС, в LPN-C матриця G_1 є загальнодоступною, а матриця G_2 генерується випадково рівноймовірно та відіграє роль секретного ключа. При цьому, на відміну від криптосистеми Мак-Еліса, для забезпечення можливості розшифрування повідомлень в LPN-C у складі шифротексту передається випадковий вектор r.

Зауважимо також, що параметр ε вибирається, виходячи з вимоги до ймовірності правильного відновлення відкритого тексту законним отримувачем. Зокрема, на підставі твердження 1 за умови $l_1 < n_1 - \sqrt{2n_1 \ln(\delta^{-1})}$ значення $\varepsilon = 1 - \frac{1}{1-2^{-s}}\left(\frac{1}{2} - \frac{l_1}{2n_1} - \sqrt{\frac{\ln(\delta^{-1})}{2n_1}}\right)$ забезпечує відновлення відкритого тексту за шифрованим із ймовірністю не менше $1-\delta$.

На криптосистему LPN-C можна здійснити атаку, аналогічну розглянутій вище атаці на криптосистему Мак-Еліса, зашифровуючи t_1 разів повідомлення $m = 0$ та формуючи n_1 систем лінійних рівнянь зі спотвореними правими частинами над полем F відносно стовпців матриці G_2. Кількість невідомих у кожній системі рівнянь дорівнює l_2, а спотворення у їх правих частинах розподілені за тим самим законом, що й координати випадкового вектора ξ (див. формулу (2)). Отже, для оцінювання ефективності такої атаки можна використовувати наступний алгоритм, аналогічний алгоритму 1.

Алгоритм 2.

Вхідні дані:

– верхня межа δ ймовірності помилкового розшифрування відкритого тексту законним отримувачем, $\delta \in (0,1)$;

– довжина n_1 та вимірність l_1 коду УРС над полем з 2^s елементів, де $s \geq 2$, $l_1 < n_1 - \sqrt{2n_1 \ln(\delta^{-1})}$, $n_1 \leq 2^s$;

– число l_2 рядків матриці G_2.

1. Покласти $\varepsilon = 1 - \dfrac{1}{1-2^{-s}}\left(\dfrac{1}{2} - \dfrac{l_1}{2n_1} - \sqrt{\dfrac{\ln(\delta^{-1})}{2n_1}}\right)$.

2. Для будь-яких $l' \in \overline{1, l_2}$ та $k' = 2, 4, 8, 16, \ldots$, що задовольняють умові

$$\log(2\Delta_1(k')^{-1} l' s \ln 2) \leq s(l_2 - l')(\log k')^{-1},$$

де $\Delta_1(k') = (2^s - 1)\varepsilon^{2k'}$, обчислити

$$T_1(k', l') = (m_1(k', l'))^{\frac{1}{\theta}} k' 2^{\frac{s(l_2 - l')}{\theta}} + s(m_1(k', l') + sl'2^{sl'}) + 2^{s(l'+1)},$$

де $\theta = 1 + \log k'$, $m_1(k', l') = 2\Delta_1(k')^{-1} l' s \ln 2$,

3. Обрати k^* та $l^* \in \overline{1, l_2}$ такі, що $T_1(k^*, l^*) = \min\{T_1(k', l')\}$, де мінімум береться за всіма зазначеними вище парами (k', l').

Результат:

– значення k^* та l^* (останнє з яких дорівнює кількості елементів довільного фіксованого стовпця матриці G_2, які відновлюються за допомогою атаки);

– середня часова складність атаки $T_1(k^*, l^*)$;

– обсяг матеріалу $t_1(k^*, l^*) = k^* 2^{\frac{s(l_2 - l^*)}{\theta}} (2l^* s \ln 2)^{\frac{1}{\theta}} \Delta_1(k^*)^{-\frac{1}{\theta}}$, потрібного для успішної реалізації атаки.

Зауважимо, що середню часову складність відновлення всієї матриці G_2 за допомогою наведеної атаки можна оцінити за формулою $T_1^{(\text{tot})}(k^*, l^*) = n_1 \lceil l_2 / l^* \rceil T_1(k^*, l^*)$; при цьому обсяг матеріалу, потрібний для відновлення цієї матриці, співпадає з $t_1(k^*, l^*)$.

Використовуючи алгоритми 1 і 2, порівняємо розглянуті криптосистеми за довжиною ключа при заданій нижній межі стійкості відносно наведеної атаки, вважаючи, що в обох випадках (як криптосистеми Мак-Еліса, так і криптосистеми LPN-C) зашифровуються відкриті тексти однакової довжини $l = l_1$.

Розглянемо, наприклад, рандомізовану симетричну криптосистему Мак-Еліса з параметрами $l = 110$, $k = 374$, $n = 512$, $\varepsilon = 0{,}9986$ (де $s = 9$, $\delta = 10^{-8}$), яка забезпечує стійкість на рівні $T(k^*, l^*) = 2^{259}$ (див. табл. 1). Вважаючи $l_1 = l$, переберемо всі значення n_1, що задовольняють умові $l_1 < n_1 - \sqrt{2n_1 \ln(\delta^{-1})}$, $n_1 \leq 2^s$, для кожного з яких переберемо усі значення $l_2 = 2, 3, \ldots$, поки не знайдемо найменше $l_2 = l_2(n_1)$, для якого повна середня часова складність $T_1^{(\text{tot})}(k^*, l^*)$ атаки на криптосистему LPN-C з параметрами l_1, n_1 та l_2 (при тих самих s та δ) є не менше ніж 2^{259}. Для кожної пари (n_1, l_2) позначимо $d(n_1, l_2) = sn_1 l_2$ та $\rho(n_1, l_2) = l_1/(n_1 + l_2)$ довжину ключа (у бітах) та швидкість передачі інформації за допомогою криптосистеми LPN-C. Нарешті, визначимо $n_1^{(\text{opt})}$, виходячи з умови $d(n_1^{(\text{opt})}, l_2(n_1^{(\text{opt})})) = \min\limits_{n_1}\{d(n_1, l_2(n_1))\}$. Тоді криптосистема LPN-C з параметрами l_1, $n_1^{(\text{opt})}$, $l_2(n_1^{(\text{opt})})$ має найменшу довжину ключа серед усіх подібних криптосистем з такою ж довжи-

ною відкритого тексту, які забезпечують стійкість відносно розглянутої атаки на рівні $T_1^{(\mathrm{tot})}(k^*,l^*) \geq 2^{259}$.

Результати розрахунків показують (табл. 3, 4), що при $s=9$, $l_1=110$ значення $n_1^{(\mathrm{opt})}$ та $l_2(n_1^{(\mathrm{opt})})$ дорівнюють 219 та 300 відповідно. Отже, при довжині відкритого тексту $l_1=110$ найменша довжина ключа криптосистеми LPN-C, що забезпечує стійкість на рівні 2^{259}, складає 591300 біт (табл. 3). При цьому для забезпечення такої ж самої стійкості за допомогою криптосистеми Мак-Еліса (табл. 1) достатньо використовувати помітно коротший ключ довжини $2ns=9216$ біт. Крім того, обидві криптосистеми мають близькі швидкості передачі, що дорівнюють приблизно $0{,}21$.

Аналогічні результати отримаємо при $s=10$, $l_1=600$: тут значення $n_1^{(\mathrm{opt})}$ та $l_2(n_1^{(\mathrm{opt})})$ дорівнюють 844 та 257 відповідно. При цьому найменша довжина ключа криптосистеми LPN-C, потрібна для забезпечення її стійкості на рівні $2^{258{,}73}$, складає 2253480 біт (табл. 4), в той час як довжина ключа відповідної криптосистеми Мак-Еліса дорівнює 20480 бітам (табл. 2).

Таблиця 3

Значення параметрів, що характеризують стійкість та практичність криптосистем LPN-C

при $s=9$, $l_1=110$, $n_1=219$ ($\delta=10^{-8}$)

l_2	l^*	k^*	$\log t_1(k^*,l^*)$	$\log T_1(k^*,l^*)$	$\log T_1^{(\mathrm{tot})}(k^*,l^*)$	$\rho(n_1,l_2)$	$d(n_1,l_2)$
2	1	2	2,95	18,21	36,17	0,49	3942
5	1	16	10,54	18,21	37,49	0,49	9855
20	2	128	28,77	29,24	47,66	0,46	39420
50	5	256	56,34	56,78	75,06	0,40	98550
80	7	512	81,07	81,07	93,86	0,36	157680
110	10	512	105,42	105,46	120,73	0,33	216810
150	14	512	137,87	138,39	156,99	0,29	295650
190	16	1024	164,28	164,29	176,05	0,26	374490
200	17	1024	171,66	171,66	184,60	0,26	394200
240	21	1024	201,14	201,71	220,52	0,23	473040
290	25	1024	238,80	239,03	256,72	0,21	571590
299	25	1024	246,16	246,16	257,71	0,21	589329
300	26	1024	246,17	246,81	265,76	0,21	591300

Таблиця 4

Значення параметрів, що характеризують стійкість та практичність криптосистем LPN-C

при $s=10$, $l_1=600$, $n_1=844$ ($\delta=10^{-8}$)

l_2	l^*	k^*	$\log t_1(k^*,l^*)$	$\log T_1(k^*,l^*)$	$\log T_1^{(\mathrm{tot})}(k^*,l^*)$	$\rho(n_1,l_2)$	$d(n_1,l_2)$
2	1	4	3,42	20,13	40,97	0,70	16880
5	1	16	11,13	20,13	42,30	0,70	42200
20	2	128	30,73	31,51	53,41	0,69	168800
80	7	512	87,71	87,72	104,22	0,64	675200
110	10	512	114,76	114,86	134,24	0,62	928400
130	11	1024	128,93	128,94	142,25	0,61	1097200
170	15	1024	161,70	162,51	184,66	0,59	1434800
200	17	1024	187,17	187,19	204,80	0,57	1688000
220	19	1024	203,54	203,86	224,87	0,56	1856800
240	20	1024	220,82	220,83	235,30	0,55	2025600
260	22	1024	237,20	237,23	255,03	0,54	2194400
266	22	1024	242,65	242,66	256,14	0,54	2245040
267	23	1024	242,66	243,26	265,04	0,54	2253480

Висновки

1. Досліджена рандомізована кодова криптосистема є симетричним аналогом окремого випадку рандомізованої (асиметричної) криптосистеми Мак-Еліса [7] та являє собою удосконалену версію симетричної кодової криптосистеми LPN-C [8], що будується на основі кодів УРС. На відміну від LPN-C (див. формулу (14)), де матриця G_1 вигляду (1) є загальнодоступною, а матриця G_2 генерується випадково рівноймовірно та відіграє роль секретного ключа, у рандомізованій симетричній криптосистемі Мак-Еліса матриця $\begin{pmatrix} G_1 \\ G_2 \end{pmatrix}$ є невідомою твірною матрицею коду УРС, що утворює секретний ключ. При цьому, на відміну від криптосистеми Мак-Еліса, для забезпечення можливості розшифрування повідомлень в LPN-C у складі шифротексту передається випадковий вектор r.

2. Стійкість рандомізованої симетричної криптосистеми Мак-Еліса відносно розглянутої атаки з підібраним відкритим текстом базується на складності відновлення невідомого коду УРС за набором його спотворених кодових слів. На сьогодні ця задача є обчислювально складною, а відомий метод її розв'язання полягає у складанні та розв'язанні систем лінійних рівнянь зі спотвореними правими частинами вигляду (8) над полем визначення коду УРС. Зауважимо, що оскільки секретний ключ (тобто код УРС) є невідомим, на рандомізовану симетричну криптосистему є незастосовними атаки [12, 13], відомі для асиметричних криптосистем, що будуються на основі кодів УРС.

3. В порівнянні з LPN-C, рандомізована симетрична криптосистема Мак-Еліса характеризується помітно меншою довжиною ключа при заданій стійкості (та однаковій довжині вхідних повідомлень). Зокрема, при $s = 9$, $l_1 = 110$ (див. табл. 1, 3) для забезпечення стійкості на рівні 2^{259} відносно розглянутої атаки за допомогою криптосистеми Мак-Еліса потрібен ключ, довжина якого є понад у 64 рази менше довжини секретного ключа криптосистеми LPN-C, а при $s = 10$, $l_1 = 600$ (див. табл. 2, 4) скорочення довжини ключа криптосистеми Мак-Еліса становить понад 110 разів.

4. На сьогодні залишається відкритим запитання про те, чи є досліджена криптосистема CPA-стійкою (відомо [8], що для LPN-C відповідь на це запитання є позитивною). Зокрема, важливо з'ясувати, чи дозволяють структурні особливості кодів УРС зменшити складність алгоритмів їх відновлення за наборами спотворених кодових слів у порівнянні з алгоритмами відновлення довільних лінійних блокових кодів.

Список літератури:

1. McEliece R.J. A public-key cryptosystem based on algebraic coding theory // Prog. Rep., Jet Prop. Lab., California Inst. Technol, 1978. P. 114 – 116.
2. Jordan J.P. A variant of public-key cryptosystem based on Goppa codes // Sigact news, 1983. P. 61 – 66.
3. Rao T.R.N. Cryptosystems using algebraic codes // Int. Conf on Computer Systems & Signal Processing, 1984.
4. Rao T.R.N. Private-key algebraic code encryption / T.R.N. Rao, K.H. Nam // IEEE Trans. on Inform Theory, 1987. P. 829 – 833.
5. Sobhi Afshar A.A. Efficient secure channel coding based on quasy-ciclic low-density parity-check codes / A.A. Sobhi Afshar, T. Eghlidos, M.R. Aref // Journal of IET-Communications, 2009. P. 279 – 292.
6. Hooshmand R. Improving the Rao-Nam secret key cryptosystem using regular EDF-QC-LDPC codes / R. Hooshmand, T. Eghlidos, M.R. Aref // ISC Journal of Information sequrity, 2012. P. 3 – 14.
7. Nojima R. Semantic security for the McElice cryptosystem without random oracles / R. Nojima, H. Imai, K. Kobara, K. Morozov // Des. Codes Cryptography, 2008. P. 289 – 305.
8. Gilbert H. How to Encrypt with the LPN Problem / H. Gilbert, J.B. Mattew, M.J.B. Robshaw, Y. Seurin // ICALP (2), Proceedings, Springer Verlag, 2008. P. 679- 690.
9. Федоренко С.В. Методы быстрого декодирования линейных блоковых кодов. С.-Петербург : ГУАПб, 2008. 199 с.
10. Niederreiter N. Knapsack-type cryptosystems and algebraic coding theory // Problems of Control and Information Theory, 1986. P. 159 – 166.

11. Berger T. How to mask the structure of codes for a cryptographic use / T. Berger, P. Loidreau // Designs, Codes and Cryptography, 2005. P. 63.

12. Сидельников В.М. О системе шифрования, построенной на основе обобщенных кодов Рида-Соломона / В.М. Сидельников, С.О. Шестаков // Дискретная математика, 1992. Т. 4. Вып. 3. С. 57 – 63.

13. Wieschebrink Ch. Cryptanalysis of the Niederreiters public key scheme based on GRS sub-codes // Post-Quantum Cryptography, 2010 Proceedings. Springer Verlag, 2010. P. 61 – 72.

14. Мак-Вильямс Ф.Дж. Теория кодов, исправляющих ошибки ; пер. с англ. / Ф.Дж. Мак-Вильямс, Н. Дж. А. Слоэн. Москва : Связь, 1979. 743 с.

15. Hoeffding W. Probability inequalities for sums of bounded random variables // J. Amer. Statist. Assoc, 1963. Vol. 58. № 301. P. 13 – 30.

16. Алексейчук А. Н. Метод восстановления систематических линейных кодов по наборам искаженных кодовых слов / А.Н. Алексейчук, А.Ю. Грязнухин // Прикладная радиоэлектроника. 2013. Т. 12. № 2. С. 313 – 318.

17. Zhang B. Fast correlation attacks over extension fields, large-unit linear approximation and cryptanalysis of SNOW 2.0 / B. Zhang, C. Xu, W. // Meier–Cryptology ePrint Archive, 2016/311. http://eprint.iacr.org/2016/311.

18. Wagner D. A generalized birthday problem // Advances in Cryptology – CRYPTO'02, Proceedings. Springer Verlag, 2002. P. 288 – 303.

Інститут спеціального зв'язку та захисту інформації
Національного технічного університету України "КПІ"
імені Ігоря Сікорського

Надійшла до редколегії 11.01.2020