УДК 621.391:519.2 **DOI:10.30837/rt.2021.2.205.02**

А.М. ОЛЕКСІЙЧУК, д-р техн. наук, О.С. ШЕВЧУК

ОЦІНКИ ЕФЕКТИВНОСТІ АТАК НА ОСНОВІ ПІДІБРАНИХ ВІДКРИТИХ ТЕКСТІВ НА КРИПТОСИСТЕМУ РАО-НАМА НАД СКІНЧЕННОЮ АБЕЛЕВОЮ ГРУПОЮ

Вступ

Криптосистема Рао – Нама [1] являє собою симетричну версію кодової криптосистеми Мак-Еліса [2], запропоновану з метою позбутися слабкостей, притаманних найпершим симетричним кодовим схемам шифрування [3, 4]. Майже одразу після опублікування цієї криптосистеми з'явилися атаки на неї на основі підібраних відкритих текстів [5, 6], що привело до появи різноманітних удосконалень та модифікацій оригінальної криптосистеми (див. [7], де можна знайти огляд публікацій, присвячених новітнім версіям криптосистеми Рао – Нама).

В [6] запропоновано так звану еквіваленту криптосистему Рао – Нама, яка описується рівнянням шифрування (над полем з двох елементів) вигляду $c = mA + e$, де відкритий текст m і шифрований текст c є двійковими векторами довжини k і n відповідно, секретний ключ A є двійковою матрицею розміру $k \times n$ і рангу k, а e являє собою випадковий двійковий вектор, який вибирається з певної множини M потужності N, яка зберігається в секреті. При цьому для забезпечення однозначності розшифрування різні елементи множини M повинні мати різні синдроми (тобто бути різними за модулем коду з твірною матрицею G). Така версія криптосистеми Рао – Нама не відрізняється за стійкістю від оригінальної криптосистеми з [1], проте характеризується помітно меншою довжиною ключа [6, с. 126]. В [5, 6] описано низку атак на основі підібраних відкритих текстів як на оригінальну, так і на еквівалентну криптосистеми Рао – Нама, але аналіз ефективності цих атак потребує подальших досліджень.

Метою статті є отримання оцінок ефективності (трудомісткості при заданій верхній межі ймовірності помилки) атак на криптосистему, яка узагальнює еквівалентну схему шифрування Рао – Нама на випадок скінченної абелевої групи (зауважимо, що необхідність дослідження подібних версій криптосистеми Рао – Нама обумовлена їх розглядом у нещодавніх публікаціях; див. [7]). Представлено дві атаки, які будуються на основі підібраних відкритих текстів. Перша з них не згадується у відомих авторам цієї статті працях і за певних (визначених нижче) умов дозволяє відновлювати секретний ключ еквівалентної криптосистеми Рао – Нама за $O(kN^2)$ операцій в середньому.

Друга атака являє собою узагальнено-спрощений варіант відомої атаки Стройка-ван Тілбурга [5, 6]. Показано, що складність цієї атаки залежить від потужності стабілізатора множини M у групі зсувів абелевої групи, над якою розглядається криптосистема Рао – Нама. Отримано оцінку ймовірності тривіальності стабілізатора за умови випадкового вибору множини M за рівноймовірною схемою. З цієї оцінки випливає, що атака Стройка-ван Тілбурга є в середньому помітно більш ефективною в порівнянні із найгіршим випадком, розглянутим в [5, 6].

1. Означення криптосистеми

Нехай G – скінченна абелева група порядку $q > 1$. Секретними ключами криптосистеми, що розглядається, є впорядковані набори $((g_1,...,g_k), M, \sigma)$, де $(g_1,...,g_k) \in G^k$, $M = \{z_1, \dots, z_N\} \subseteq G$, $\sigma : G \to G'$ – епіморфізм груп, ядро якого співпадає з підгрупою H, породженою елементами $g_1,...,g_k$. При цьому вважається, що виконані такі умови:

а) група H є прямою сумою циклічних підгруп, породжених елементами $g_1,...,g_k$ відповідно, тобто кожен елемент $g \in H$ допускає однозначне представлення у вигляді $g = m_1 g_1 + \cdots + m_k g_k$, де $m_i \in \overline{0, q_i - 1}$, q_i – порядок елемента g_i групи G, $i \in \overline{1, k}$;

б) елементи $z_1, \ldots, z_N$ множини M належать різним суміжним класам по підгрупі H, відмінним від цієї підгрупи.

Елементи з M зберігаються у вигляді таблиці $((\sigma(z_i), z_i) : i \in \overline{1, N})$, де кожен елемент z_i записано за адресою $\sigma(z_i)$. Зауважимо, що на підставі умови *б*) усі такі адреси є ненульовими попарно різними елементами групи G'.

За означенням множина відкритих текстів криптосистеми складається з усіх наборів $(m_1, \ldots, m_k)$, де $m_i \in \overline{0, q_i - 1}$, $i \in \overline{1, k}$. Для зашифрування відкритого тексту на ключі $((g_1,...,g_k), M, \sigma)$ з множини M вибирається випадковий рівноймовірний елемент z і обчислюється шифрований текст

$$c = m_1 g_1 + \cdots + m_k g_k + z. \qquad (1)$$

Для розшифрування цього шифротексту законний отримувач обчислює значення $\sigma(c)$, яке співпадає з елементом $\sigma(z)$ внаслідок означення епіморфізму σ. Далі отримувач знаходить елемент z за елементом $\sigma(z)$, використовуючи таблицю для зберігання множини M, обчислює повідомлення $c - z = m_1 g_1 + \cdots + m_k g_k$, за яким відновлює відкритий текст $(m_1, \ldots, m_k)$, спираючись на умову (а).

Як приклад розглянемо окремий випадок описаної криптосистеми, що будується над групою $G = (\mathbf{GF}(2)^n, \oplus)$ і являє собою еквівалентну версію класичної криптосистеми Рао – Нама, визначену в [6]. В цьому випадку $g_1,...,g_k$ є лінійно незалежними (над полем з двох елементів) двійковими векторами довжини n, які утворюють $k \times n$ матрицю A, що є твірною матрицею деякого двійкового лінійного коду C. Епіморфізм σ групи G в групу $G' = (\mathbf{GF}(2)^{n-k}, \oplus)$ визначається за формулою $\sigma(z) = Bz^T$, де B є перевірочною матрицею коду C, а z^T позначає вектор, транспонований до $z \in G$ (таким чином, в даному випадку $\sigma(z)$ є просто синдромом двійкового слова z). Згідно з формулою (1) шифротекст c отримується в результаті кодування відкритого тексту $(m_1, \ldots, m_k)$ кодом C з подальшим "накладанням" вектора помилок $z \in M$, який можна "зняти" на приймальному кінці системи зв'язку, знаючи множину M та синдром $\sigma(z)$.

В роботі Рао і Нама [1] розглянуто два способи формування множини M, перший з яких полягає у використанні певних заздалегідь визначених векторів, що мають вагу (Геммінга) приблизно $n/2$, а другий – у випадковому виборі цих векторів з урахуванням умови *б*). В [5] показано, що перший варіант не забезпечує належну стійкість криптосистеми Рао – Нама (внаслідок невеликої кількості та простої будови зазначених векторів), проте другий варіант є більш змістовним та потребує додаткових досліджень.

Для наведеної вище криптосистеми над скінченною абелевою групою G зазначений другий спосіб формування множини M узагальнюється таким чином.

Позначимо $r = |G| \cdot |H|^{-1}$ число суміжних класів групи G по підгрупі H. Тоді для формування випадкової множини M спочатку з імовірністю $\binom{r-1}{N}^{-1}$ вибирається множина з N ненульових суміжних класів по підгрупі H, після чого в кожному суміжному класі з імовірністю $|H|^{-1}$ вибирається один випадковий елемент, причому всі ці елементи вибираються

незалежно один від одного. Іншими словами, вважається, що кожна множина M з N елементів, яка задовольняє умові *б*), має однакову ймовірність $p(M)=\binom{r-1}{N}^{-1}|H|^{-N}$. В подальшому така схема формування множини M називається *рівноймовірною*.

Нижче розглядаються атаки на описану криптосистему, при проведенні яких вважається, що супротивник має доступ до оракула зашифрування з невідомим ключем $((g_1,...,g_k), M, \sigma)$. Метою атак є відновлення окремих частин ключа – множини M та впорядкованого набору $(g_1,...,g_k)$. *Ефективність атаки* характеризується її трудомісткістю (середньою або у найгіршому випадку) при заданій верхній межі ймовірності помилки атаки.

2. Алгоритм відновлення множини M

В [5, 6] (для випадку класичної криптосистеми Рао – Нама) описано природний *алгоритм відновлення множини* M, який полягає в наступному:

– вибрати натуральне число t;

– отримати набір $c_1,...,c_t$ шифрованих текстів, подаючи t разів на вхід оракула зашифрування відкритий текст $(m_1,\dots,m_k)=(0,\dots,0)$;

– покласти $M=\{c_1,...,c_t\}$.

Позначимо $t_0=t_0(N)$ найменше натуральне t, для якого множина, що складається з усіх різних елементів $c_1,...,c_t$, дорівнює M. Тоді t_0 є випадковою величною, що дорівнює найменшому числу частинок, які потрібно кинути в N скриньок для того, щоби усі скриньки були заповнені. (Дійсно, елементи множини M можна розглядати як скриньки, а шифровані повідомлення, отримані шляхом зашифрування нульового відкритого тексту, як частинки, що кидаються у скриньки випадково рівноймовірно та незалежно одна від одної) [8].

Для математичного сподівання випадкової величини t_0 справедливі співвідношення $\mathbf{E}t_0=N\sum_{i=1}^{N}1/i<N(\ln N+C+N^{-1})$, де $C=0{,}5772...$ – константа Ойлера [8, с. 18 ; 9, с. 108]. Це означає, що при $N>2$ середнє число зашифровувань, потрібних для успішного відновлення множини M, не перевищує $N(\ln N+1)$.

Для оцінки трудомісткості наведеного алгоритму відновлення множини M у найгіршому випадку можна скористатися відомою формулою для числа сюр'єктивних відображень t-множини в N-множину: $D(N,t)=\sum_{l=0}^{N}(-1)^l\binom{N}{l}(N-l)^t$ [10]. Для будь-якого $\delta\in(0,1)$ визначимо $t_1=t_1(\delta)$ як найменше натуральне t, для якого $D(N,t)\ge N^t(1-\delta)$. Тоді, застосовуючи наведений алгоритм з параметром $t=t_1$, отримаємо випадкову множину, яка співпадає з M із ймовірністю не менше $1-\delta$.

3. Алгоритм відновлення набору $(g_1,...,g_k)$

Для будь-якого $i\in\overline{1,k}$ позначимо e_i цілочисельний вектор довжини k, який має єдину ненульову, а саме, i-ту, координату, що дорівнює 1. Подаватимемо вектор e_i на вхід оракула зашифрування, поки вперше не отримаємо три різних шифрованих тексти:

$$c_1=g_i+z_{j_1},\ c_2=g_i+z_{j_2}\ \text{та}\ c_3=g_i+z_{j_3}, \tag{2}$$

де $z_{j_1}, z_{j_2}, z_{j_3}$ є незалежними випадковими рівноймовірними елементами множини M. Зауважимо, що на підставі леми 8.2.2 у [6] середнє число зашифрувань, які потрібно зробити для отримання трьох різних шифрованих текстів, дорівнює $3+1/(N-1)+1/(N-2)$.

Віднімаючи з першого рівняння (2) друге та третє відповідно, отримаємо, що

$$c_1 - c_2 = z_{j_1} - z_{j_2},\ c_1 - c_3 = z_{j_1} - z_{j_3}, \tag{3}$$

де $z_{j_1} \neq z_{j_2}$, $z_{j_1} \neq z_{j_3}$, $z_{j_2} \neq z_{j_3}$.

Припустимо, що множина M задовольняє такій *умові однозначності*: для будь-якого ненульового елемента $y \in G$ існує не більше однієї множини $\{z, z'\} \subseteq M$ такої, що $y = z - z'$.

В цьому випадку можна запропонувати наступний *алгоритм знаходження впорядкованого набору* $(g_1,...,g_k)$:

1) відновити множину M за допомогою алгоритму, наведеного в п. 2;

2) занумерувати елементи цієї множини довільним чином та побудувати таблицю T, яка складається з елементів $z_u - z_v$, записаних за адресами $\{u,v\}$, де $1 \le u < v \le N$, а $z_1, \dots, z_N$ є усі попарно різні елементи з M;

3) для кожного $i \in \overline{1,k}$:

– подавати вектор e_i на вхід оракула зашифрування, поки вперше не буде отримано три різних шифрованих тексти c_1, c_2, c_3;

– використовуючи таблицю T, знайти множини $\{u_1,v_1\}$ та $\{u_2,v_2\}$ такі, що $c_1 - c_2 = z_{u_1} - z_{v_1}$ та $c_1 - c_3 = z_{u_2} - z_{v_2}$;

– покласти $g_i = c_1 - z_{j_1}$, де j_1 – єдиний спільний елемент множин $\{u_1,v_1\}$ та $\{u_2,v_2\}$ (такий елемент напевно існує).

Твердження 1. Нехай $r = |G| \cdot |H|^{-1} \ge 14$ і випадкова множина M формується за рівноймовірною схемою (див. п. 1). Тоді ця множина задовольняє умові однозначності з ймовірністю не менше ніж $1 - N^4 |G|^{-1}$. При цьому наведений вище алгоритм відновлює набір $(g_1,...,g_k)$ із середньою трудомісткістю $O(kN^2)$.

Доведення. Перш за все, помітимо, що за умови однозначності система рівнянь (3) має єдиний розв'язок $(z_{j_1}, z_{j_2}, z_{j_3})$, який можна знайти за допомогою наведеного алгоритму, використовуючи рівності $\{u_1,v_1\} = \{j_1, j_2\}$, $\{u_2,v_2\} = \{j_1, j_3\}$. Звідси з урахуванням рівностей (2) випливає, що цей алгоритм вірно знаходить усі елементи набору $(g_1,...,g_k)$.

Далі, згідно з результатами п. 2, середнє число операцій, потрібних для побудови множини M на кроці 1) алгоритму, дорівнює $O(N \log N)$. Для побудови таблиці T на кроці 2) треба виконати $1/2 \cdot N(N-1)$ операцій. Нарешті, на кроці 3) для знаходження множини $\{u_1,v_1\}$ треба відшукати в таблиці T елемент $c_1 - c_2$, а у випадку його відсутності – елемент $c_2 - c_1$ (хоча б один з цих двох елементів обов'язково є в таблиці), що потребує не більше ніж $N(N-1)$ операцій. Таку ж кількість операцій треба виконати для знаходження множини $\{u_2,v_2\}$. Отже, середня трудомісткість кроку 3) не перевищує $k\left(3+1/(N-1)+1/(N-2)+3+2N(N-1)\right) = O(kN^2)$, і такою ж є оцінка середньої трудомісткості алгоритму в цілому.

Переконаємося зараз у справедливості першої частини твердження.

Позначимо: p_N ймовірність того, що випадкова множина M не задовольняє умові однозначності; m_N – число усіх множин потужності N, які задовольняють умові б) з п. 1 та

не задовольняють умові однозначності; $\tilde{m}_N$ – число усіх впорядкованих наборів $(z_1, \dots, z_N)$, елементи яких належать різним ненульовим суміжним класам групи G по підгрупі H, й таких, що існує принаймні дві різні множини $\{u_1, v_1\}$, $\{u_2, v_2\}$ чисел від 1 до N такі, що $z_{u_1} - z_{v_1} = z_{u_2} - z_{v_2}$. З наведених означень випливає, що

$$p_N = m_N \binom{r-1}{N}^{-1} |H|^{-N} = \frac{\tilde{m}_N}{N!} \binom{r-1}{N}^{-1} |H|^{-N}. \quad (4)$$

При цьому

$$\tilde{m}_N \le \sum_{\{\{u_1,v_1\},\{u_2,v_2\}\}} \tilde{m}_N(u_1, v_1; u_2, v_2), \quad (5)$$

де $\tilde{m}_N(u_1, v_1; u_2, v_2)$ – число впорядкованих наборів $(z_1, \dots, z_N)$, елементи яких належать різним ненульовим суміжним класам групи G по підгрупі H, й таких, що $z_{u_1} - z_{v_1} = z_{u_2} - z_{v_2}$, а підсумування здійснюється за всіма множинами $\{\{u_1, v_1\}, \{u_2, v_2\}\}$ такими, що $\{u_1, v_1\}$ та $\{u_2, v_2\}$ є різними підмножинами множини чисел від 1 до N

Запишемо суму у правій частині нерівності (5) у вигляді $\tilde{m}_{N,1} + \tilde{m}_{N,2}$, де $\tilde{m}_{N,1}$ та $\tilde{m}_{N,2}$ є сумами чисел $\tilde{m}_N(u_1, v_1; u_2, v_2)$ за всіма такими підмножинами $\{\{u_1, v_1\}, \{u_2, v_2\}\}$, що $\{u_1, v_1\} \cap \{u_2, v_2\} = \varnothing$ та $|\{u_1, v_1\} \cap \{u_2, v_2\}| = 1$ відповідно.

Помітимо, що за умови $\{u_1, v_1\} \cap \{u_2, v_2\} = \varnothing$ число $\tilde{m}_N(u_1, v_1; u_2, v_2)$ дорівнює добутку двох чисел, перше з яких є кількістю всіх наборів $(z_{u_1}, z_{v_1}, z_{u_2}, z_{v_2})$, елементи яких належать різним ненульовим суміжним класам групи G по підгрупі H та задовольняють умові $z_{u_1} - z_{v_1} = z_{u_2} - z_{v_2}$, а друге є кількістю всіх впорядкованих наборів довжини $N-4$, елементи яких належать різним ненульовим суміжним класам групи G по підгрупі H. Перше з цих двох чисел не перевищує кількості розв'язків $(z_{u_1}, z_{v_1}, z_{u_2}, z_{v_2})$ рівняння $z_{u_1} - z_{v_1} = z_{u_2} - z_{v_2}$, яка дорівнює $|G|^3$, а друге дорівнює $(r-5)_{N-4} |H|^{N-4}$, де $(r-5)_{N-4}$ позначає число розміщень з $r-5$ по $N-4$. Таким чином, кожен доданок у сумі, що дорівнює $\tilde{m}_{N,1}$, не перевищує $|G|^3 (r-5)_{N-4} |H|^{N-4}$, в той час як число доданків дорівнює $\frac{1}{2}\binom{N}{2}\binom{N-2}{2} \le \frac{N^4}{8}$. Отже, $\tilde{m}_{N,1} \le \frac{N^4}{8} |G|^3 (r-5)_{N-4} |H|^{N-4}$.

Аналогічно доводиться, що кожен доданок в сумі, яка дорівнює $\tilde{m}_{N,2}$, не перевищує $|G|^2 (r-4)_{N-3} |H|^{N-3}$, в той час як число доданків є $\frac{1}{2}\left(\binom{N}{2}^2 - \binom{N}{2}\binom{N-2}{2} - \binom{N}{2}\right) \le \frac{N^4}{8}$.

Отже, $\tilde{m}_{N,2} \le \frac{N^4}{8} |G|^2 (r-4)_{N-3} |H|^{N-3}$.

Підставляючи наведені оцінки у формулу (4), отримаємо, що

$$p_N \le \frac{\tilde{m}_{N,1} + \tilde{m}_{N,2}}{N!} \binom{r-1}{N}^{-1} |H|^{-N} \le$$

$$\le \frac{N^4 |G|^3}{8} \frac{(r-5)_{N-4} |H|^{N-4}}{(r-1)_N |H|^N} + \frac{N^4 |G|^2}{8} \frac{(r-4)_{N-3} |H|^{N-3}}{(r-1)_N |H|^N} =$$

$$= \frac{N^4 |G|^3}{8|H|^4 r^4} \frac{r^4}{(r-1)(r-2)(r-3)(r-4)} + \frac{N^4 |G|^2}{8|H|^3 r^3} \frac{r^3}{(r-1)(r-2)(r-3)}.$$

Нарешті, враховуючи рівність $r = |G| \cdot |H|^{-1}$, отримаємо, що кожен з двох доданків у наведеній сумі не перевищує $\frac{N^4}{8|G|} \cdot \frac{1}{(1-4r^{-1})^4}$, що у свою чергу, не перевищує $\frac{N^4}{2|G|}$ через умову $r \ge 14$.

Таким чином, справедлива нерівність $p_N \le \frac{N^4}{|G|}$, що й треба було довести.

Зауважимо, що у випадку $G = (\mathbf{GF}(2)^n, \oplus)$ отриману оцінку ймовірності p_N можна декілька підсилити, враховуючи той факт, що $\tilde{m}_{N,2} = 0$. Дійсно, як випливає з означення цього параметра у викладеному доведенні, справедлива рівність $|\{u_1, v_1\} \cap \{u_2, v_2\}| = 1$, оскільки в протилежному випадку множини $\{u_1, v_1\}$ та $\{u_2, v_2\}$ співпадають внаслідок рівності $z_{u_1} + z_{v_1} = z_{u_2} + z_{v_2}$. Таким чином, якщо $G = (\mathbf{GF}(2)^n, \oplus)$, то за умови твердження $p_N \le \frac{N^4}{2|G|}$.

Отриманий результат показує, що у випадку, коли N є не надто великим числом (а саме, $N < \delta |G|^{1/4}$, $\delta \in (0,1)$), і множина M формується за рівноймовірною схемою, описана криптосистема може бути зламана з ймовірністю не менше ніж $1-\delta$ за $O(kN^2)$ операцій в середньому. Для підвищення стійкості криптосистеми треба збільшити параметр N, що негативно відіб'ється на її практичності, оскільки таблицю з N елементів множини M, треба зберігати в пам'яті.

4. Атака Стройка-ван Тілбурга

В даному пункті наведено узагальнення на випадок криптосистеми над скінченною абелевою групою атаки, запропонованої в [5] та вдосконаленої в [6]. Зауважимо, що у відзначених публікаціях для опису атаки використовуються помічені графи та їх групи автоморфізмів, що не є обов'язковим і, на нашу думку, ускладнює викладення. Нижче наведено більш простий опис атаки, аналогічної за сутністю запропонованій в [6]. Показано, що складність атаки залежить від потужності стабілізатора множини M у групі зсувів абелевої групи G. Основним результатом цього пункту є твердження про те, що зазначений стабілізатор є з високою ймовірністю тривіальним у випадку, коли випадкова множина M формується за рівноймовірною схемою (див. п. 1). Звідси випливає, що у зазначеному випадку трудомісткість наведеної атаки дорівнює $O(kN^2\gamma(N))$, де $\gamma(N)$ – трудомісткість пошуку елемента в масиві, який складається з N елементів групи G.

Атака, що розглядається, проводиться в два етапи, на першому з яких відновлюється множина M за допомогою алгоритму з п. 2. На другому етапі відновлюється набір $(g_1, ..., g_k)$. З цією метою для кожного $i \in \overline{1,k}$ на вхід оракула зашифрування подається повідомлення e_i поки вперше не буде отримано N різних шифротекстів $c_1, ..., c_N$. На підставі формули (1) невідомий елемент g_i задовольняє системі рівнянь

$$x + z_{\pi(j)} = c_j, \ j \in \overline{1,N}, \tag{6}$$

де $M=\{z_1,...,z_N\}$, π – деяка підстановка на множині $\overline{1,N}$. Неважко знайти один з можливих розв'язків системи рівнянь (6), перебираючи всі значення $\pi(1)$ та обчислюючи $x=c_1-z_{\pi(1)}$, $z_{\pi(2)}=c_2-x$, …, $z_{\pi(N)}=c_N-x$. Оскільки обчислені таким чином елементи $z_{\pi(1)}$, $z_{\pi(2)}$, …, $z_{\pi(N)}$ є попарно різними, то необхідною й достатньою умовою правильності вибору значення $\pi(1)$ є приналежність елементів $z_{\pi(2)}$, …, $z_{\pi(N)}$ множині M, тобто умова

$$c_j-c_1+z_{\pi(1)}\in M,\ j\in\overline{2,N}. \tag{7}$$

Якщо ця умова виконується, то шуканий розв'язок g_i системи рівнянь (6) визначається за формулою $g_i=x=c_1-z_{\pi(1)}$.

Позначимо $I(M)=\{x\in G: x+M=M\}$ стабілізатор множини M в групі зсувів абелевої групи G. Відзначимо, що $I(M)$ є найбільшою за включенням підгрупою I групи G такою, що M є об'єднанням деяких суміжних класів G по I. Зокрема, число $|I(M)|$ є дільником числа $N=|M|$.

Твердження 2. Множина розв'язків системи рівнянь (6) має вигляд $x_0+I(M)$, де x_0 – довільний фіксований розв'язок цієї системи рівнянь.

Доведення. Якщо x є розв'язком системи (6), то для деяких підстановок π та π' на множині $\overline{1,N}$ мають місце рівності $x+z_{\pi(j)}=c_j=x_0+z_{\pi'(j)}$, $j\in\overline{1,N}$. Отже, $x-x_0+z_{\pi(j)}=z_{\pi'(j)}$, $j\in\overline{1,N}$, звідки випливає, що $x-x_0\in I(M)$. Навпаки, якщо $y\in I(M)$, то $x_0+y=c_j-z_{\pi'(j)}+y$ і, отже, $(x_0+y)+z_{\pi'(j)}=c_j$, $j\in\overline{1,N}$, тобто x_0+y задовольняє системі рівнянь (6) для підстановки $\pi=\pi'$.

Твердження доведено.

Таким чином, на підставі викладеного можна запропонувати наступний *алгоритм, який реалізує розглянуту атаку на криптосистему*:

1) відновити множину M за допомогою алгоритму, наведеного в п. 2;

2) для кожного $i\in\overline{1,k}$:

– подавати вектор e_i на вхід оракула зашифрування поки вперше не буде отримано N різних шифротекстів $c_1,...,c_N$;

– знайти шляхом перебору значення $\pi(1)\in\overline{1,N}$, для якого виконується умова (7) та покласти $g_i^{(0)}=c_1-z_{\pi(1)}$.

З означення криптосистеми і твердження 2 випливає, що наведений алгоритм завжди знайде певний набір $(g_1^{(0)},...,g_k^{(0)})\in G^k$ такий, що сукупність усіх шуканих наборів $(g_1,...,g_k)$ визначається за формулою

$$\{(g_1^{(0)},...,g_k^{(0)})+(x_1,...,x_k):(x_1,...,x_k)\in I(M)^k\}. \tag{8}$$

Іншими словами, кожен набір, який належить множині (8), може використовуватися (з погляду супротивника, який має доступ тільки до оракула зашифрування) в ролі частини $(g_1,...,g_k)$ секретного ключа криптосистеми. І навпаки, будь-який набір з останньою властивістю належить множині (8).

Таким чином, вся інформація про набір $(g_1,...,g_k)$, яку може отримати супротивник, маючи доступ до оракула зашифрування, полягає в тому, що цей набір належить множині (8), і наведений вище алгоритм дозволяє знайти один з елементів цієї множини.

Як видно з опису алгоритму, його середня трудомісткість складає $O(kN^2\gamma(N))$ операцій, де $\gamma(N)$ – складність пошуку одного елемента групи G в масиві, що використовується для зберігання множини M. Зокрема, $\gamma(N)=O(N)$ для довільної групи G і $\gamma(N)=O(\log N)$ у випадку, коли $G=(\mathbf{GF}(2)^n,\oplus)$ і елементи множини M розташовані в масиві в лексикографічному порядку. Крім того, середня трудомісткість побудови всієї множини (8) дорівнює $O(kN^2\gamma(N)+k\,|\,I(M)\,|^k)$, що при $k\geq 3$ є величиною порядку kN^k.

Знаючи набір $(g_1^{(0)},...,g_k^{(0)})$ і множину M, можна реалізувати на криптосистему атаку на основі відомого шифротексту, викладену в [5, 6] для випадку класичної криптосистеми Рао – Нама.

Дійсно, нехай c є шифрованим текстом вигляду (1), отриманим з деякого невідомого відкритого тексту $(m_1,...,m_k)$. Тоді

$$c=m_1g_1+\cdots+m_kg_k+z=m_1g_1^{(0)}+\cdots+m_kg_k^{(0)}+z', \tag{9}$$

де $z'=z+(m_1g_1+\cdots+m_kg_k)-(m_1g_1^{(0)}+\cdots+m_kg_k^{(0)})$. Оскільки набір $(g_1,...,g_k)$ належить множині (8), а $I(M)$ є підгрупою групи G, то $(m_1g_1+\cdots+m_kg_k)-(m_1g_1^{(0)}+\cdots+m_kg_k^{(0)})\in I(M)$ і, отже, $z'\in M$, оскільки $z\in M$. Перебираючи усі значення z', можна знайти з формули (9) (можливо, у варіантах) значення $m_1g_1^{(0)}+\cdots+m_kg_k^{(0)}$ за $O(N\theta(k,N))$ операцій, де $\theta(k,N)$ позначає складність перевірки приналежності елемента групи G підгрупі, породженій елементами $g_1^{(0)},...,g_k^{(0)}$. При цьому, якщо $I(M)=\{0\}$, то $(g_1,...,g_k)=(g_1^{(0)},...,g_k^{(0)})$ і на підставі умов (а), (б) з означення криптосистеми (п. 1) значення $m_1g_1^{(0)}+\cdots+m_kg_k^{(0)}$, а отже, і невідомий відкритий текст, можна відновити однозначно.

Отримаємо зараз верхню оцінку ймовірності події $I(M)\neq\{0\}$, вважаючи, що M є випадковою множиною, яка формується за рівноймовірною схемою (див. п. 1).

Твердження 3. Нехай $\left\lceil\frac{N}{2}\right\rceil<\frac{|G|}{2\exp G}$, де $\exp G$ – експонента (максимальний порядок елементів) групи G. Тоді для випадкової множини M потужності N, яка формується за рівноймовірною схемою, справедлива нерівність

$$\mathbf{P}(I(M)\neq\{0\})\leq\frac{1}{2}\left(\frac{N}{2}\right)^{\frac{N}{2}}\left(1-\frac{r}{N}\right)^{-N}|\,G\,|^{2-\frac{N}{2}}, \tag{10}$$

де $r=|\,G\,||\,H\,|^{-1}$.

Доведення. Нехай $I(M)\neq\{0\}$. Тоді існує ненульовий елемент $a\in G$ такий, що $a+M=M$, звідки випливає, що M є об'єднанням деяких циклів підстановки $t_a(x)=x+a$, $x\in G$.

Позначимо l порядок елемента a групи G. Підстановка t_a є добутком $\frac{|G|}{l}$ циклів, кожен з яких має довжину l, а кількість усіх множин потужності N, які є об'єднанням циклів цієї підстановки, дорівнює біноміальному коефіцієнту $\binom{\frac{|G|}{l}}{\frac{N}{l}}$, якщо l ділить N, та нулю – у протилежному випадку. При цьому, оскільки множина M формується за рівноймовірною схемою, то

$$\mathbf{P}(I(M)\neq\{0\})\le\sum_{a\in G\setminus\{0\}}\mathbf{P}(a+M=M)\le|G||H|^{-N}\binom{r-1}{N}^{-1}\binom{\frac{|G|}{l}}{\frac{N}{l}}. \qquad (11)$$

Далі, з умови $\left\lceil\frac{N}{2}\right\rceil<\frac{|G|}{2\exp G}$ випливає, що $\frac{N}{l}\le\left\lceil\frac{N}{2}\right\rceil<\frac{|G|}{2l}$ і, отже,

$$\binom{\frac{|G|}{l}}{\frac{N}{l}}\le\binom{\frac{|G|}{l}}{\left\lceil\frac{N}{2}\right\rceil}\le\frac{\left(\frac{|G|}{2}\right)^{\left\lceil\frac{N}{2}\right\rceil}}{\left\lceil\frac{N}{2}\right\rceil!}\le\frac{\left(\frac{|G|}{2}\right)^{\frac{N}{2}+1}}{\left\lceil\frac{N}{2}\right\rceil!}.$$

Підставляючи зазначену оцінку в формулу (11) та використовуючи рівність $r=|G||H|^{-1}$, отримаємо, що

$$\mathbf{P}(I(M)\neq\{0\})\le|G||H|^{-N}\frac{N!}{(r-1)_N}r^N\frac{|H|^N}{|G|^N}\frac{\left(\frac{|G|}{2}\right)^{\frac{N}{2}+1}}{\left\lceil\frac{N}{2}\right\rceil!}=$$

$$=N(N-1)\cdots\left(\left\lceil\frac{N}{2}\right\rceil+1\right)2^{-\frac{N}{2}-1}\frac{1}{(1-r^{-1})(1-2r^{-1})\cdots(1-Nr^{-1})}|G|^{2-\frac{N}{2}}\le$$

$$\le\frac{1}{2}\left(\frac{N}{2}\right)^{\frac{N}{2}}\left(1-\frac{r}{N}\right)^{-N}|G|^{2-\frac{N}{2}}.$$

Таким чином, справедлива нерівність (10). Твердження доведено.

Наслідок. Нехай $G=(\mathbf{GF}(2)^n,\oplus)$ і N є непарним числом. Тоді $I(M)=\{0\}$. Якщо ж N є парним і $N<2^{n-k-1}$, а множина M формується за рівноймовірною схемою, то

$$\mathbf{P}(I(M)\neq\{0\})\le2^{-n\left(\frac{N}{2}-2-\frac{N}{2n}(\log n+1)\right)-1}. \qquad (12)$$

Доведення. Якщо $N=|M|$ є непарним числом, то рівність $a\oplus M=M$ є неможливою для будь-якого ненульового елемента $a\in\mathbf{GF}(2)^n$. Отже, в цьому випадку $I(M)=\{0\}$. Для парного N нерівність (12) є безпосереднім наслідком формули (10) та нерівності $\left(1-\frac{r}{N}\right)^{-N}\le2^N$, яка випливає з оцінки $N<2^{n-k-1}=\frac{r}{2}$ (зауважимо також, що в даному випадку $\exp G=2$ і нерівність $\left\lceil\frac{N}{2}\right\rceil<\frac{|G|}{2\exp G}$ напевно виконується).

Наслідок доведено.

Висновки

Основним результатом статті є аналітичні оцінки ефективності двох атак на симетричну криптосистему, яка узагальнює відому еквівалентну схему Рао – Нама [6] на випадок скінченної абелевої групи. Секретними ключами такої криптосистеми над групою G є набори $((g_1,\dots,g_k),M,\sigma)$, де $(g_1,\dots,g_k)\in G^k$, $M=\{z_1,\dots,z_N\}\subseteq G$, $\sigma:G\to G'$ – епіморфізм груп,

ядро якого співпадає з підгрупою H, породженою елементами $g_1,...,g_k$. При цьому вважається, що виконані умови (а), (б) з п. 1.

Перша атака, запропонована в цій статті, дозволяє зламувати зазначену криптосистему з ймовірністю не менше ніж $1-\delta$ за $O(kN^2)$ операцій в середньому за умови $N < \delta |G|^{1/4}$, де $\delta \in (0,1)$.

Друга атака являє собою узагальнено-спрощений варіант атаки Стройка-ван Тілбурга [5, 6] і за певних умов (див. твердження 3) дозволяє зламувати зазначену криптосистему за $O(kN^2\gamma(N))$ операцій, де $\gamma(N)$ – трудомісткість пошуку елемента в масиві, який складається з N елементів групи G.

На відміну від [5, 6], при аналізі ефективності атаки Стройка-ван Тілбурга не використовуються деякі "зайві" поняття (поміченого графу та його групи автоморфізмів), що суттєво спрощує викладення. Показано, що складність цієї атаки залежить від потужності стабілізатора множини M у групі зсувів абелевої групи G, який є тривіальним з ймовірністю, зазначеній у формулюванні твердження 3.

Для підвищення стійкості розглянутої криптосистеми до наведених атак треба збільшити параметр N, що негативно відіб'ється на її практичності, оскільки треба зберігати в пам'яті таблицю з N елементів множини M.

Список літератури:

1. Rao T.R.N., Nam K.H. Private-key algebraic code encryption // IEEE Trans. on Inform Theory, 1989. P. 829 – 833.
2. McEliece R.J. A public-key cryptosystem based on algebraic coding theory // Prog. Rep., Jet Prop.Lab., California Inst. Technol, 1978. P. 114 – 116.
3. Jordan J.P. A variant of public-key cryptosystem based on Goppa codes // Sigact news, 1983. P. 61 – 66.
4. Rao T.R.N. Cryptosystems using algebraic codes // Int. Conf on Computer Systems & Signal Processing, 1984.
5. Struik R., van Tilburg J. Thr Rao-Nam scheme is insecure against a chosen plaintext attack // Advances in Cryptology-CRYPTO'87. Proc. Springer, 1988. P. 445 – 457.
6. Van Tilburg J. Security-analyses of a class of cryptosystems based on linear error-correcting codes. PhD Thesis. Technische Universiteit Eindhoven, 1994.
7. Bagheri K., Eghlidos T., Sadeghi M.-R., Panario D. Lattice based join encryption, encoding, and modulation scheme // arXiv: 1906.06280v1[cs.IT] 4 Juni 2019. P. 1 – 30.
8. Колчин В.Ф., Севастьянов Б.А., Чистяков В.П. Случайные размещения. Москва : Наука, 1976. 223 с.
9. Гельфонд А.О. Исчисление конечных разностей. Москва : Наука, 1967. 376 с.
10. Сачков В.Н. Введение в комбинаторные методы дискретной математики. Москва : Наука, 1982. 384 с.

Надійшла до редколегії 25.02.2021

Відомості про авторів:

Олексійчук Антон Миколайович – д-р техн. наук, доцент, Інститут спеціального зв'язку та захисту інформації Національного технічного університету України "КПІ", професор кафедри Кібербезпеки; Україна; e-mail: alex-dtn@ukr.net; ORCID: https://orcid.org/0000-0003-4385-4631

Шевчук Ольга Сергіївна – Інститут спеціального зв'язку та захисту інформації Національного технічного університету України "КПІ", інженер кафедри Кібербезпеки; Україна; e-mail: olia13511@gmail.com; ORCID: https://orcid.org/0000-0002-2866-439X