УДК 517

АНАЛІЗ ЛАНЦЮГІВ ПІДПРОСТОРІВ «КАЛИНА»-ПОДІБНИХ ШИФРІВ

М. О. Коляда[1, a]

[1] *Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Фізико-технічний інститут*

Анотація

У даній роботі розглянуто та наведено приклади ланцюгів підпросторів для одного, двох, трьох та чатирьох раундів для «Калина»-подібних шифрів. Вперше було розглянуто властивості п'ятого раунду «Калина»-подібних шифрів, що дозволяє побудувати атаку відновлення раундового ключа, для 6го раунду.

Ключові слова: симетрична криптографія, блокові шифри, «Калина», ланцюги підпросторів

Вступ

Наявність у блоковому шифрі характеристик із нерівноімовірним розподілом дозволяє будувати ефективні стастистичні атаки відновлення раундових ключів або ключа шифрування загалом. Пошук та аналіз поведінки таких нерівномірних статистик є постійно актуальною криптографічною задачею. Дослідження ланцюгів підпросторів було вперше запропоноване у 2011 році для криптоаналізу шифру PRINTcipher [1], після чого ідея такого аналізу була ефективно застосована для шифру AES [2], [3]. Для проведення такого аналізу у [2], [3] запропонована техніка використання спеціальним чином підібраних вхідних даних, поведінку яких можна спрогнозувати після декількох раундів шифрування; при цьому від шифру не вимагається наявність спеціальних симетрій або констант. У даній роботі ми розглянемо атаку відновлення раундового ключа, що використовує ланцюги підпросторів, які можна побудувати для блокових шифрів із структурою алгоритму шифрування «Калина» (ДСТУ 7624:2014 [4], [5]).

1. Короткий опис блокового шифру «Калина»

Структура шифру «Калина» подібна до структури шифру Rijndael, але орієнтована на 64-бітні обчислювальні архітектури. Кількість раундів залежить від довжини відкритого тексту та довжини ключа. Відкритий текст подається у вигляді матриці розміром $a \times 8$, де $a \in \{2, 4, 8\}$. У даній роботі будемо розглядати випадок, коли $a = 8$. Базові перетворення для шифрування:

$$E^K_{l,k} = AddKey'^{K_t} \circ MixColumns \circ ShiftRows \circ SubBytes \circ \prod_{i=1}^{t-1} (AddKey^{K_i} \circ MixColumns \circ ShiftRows \circ SubBytes) \circ AddKey^{K_0},$$

де l – розмір внутрішнього стану блокового шифру (у бітах), K_i – раундовий ключ шифрування, k – довжина ключа шифрування (у бітах), $AddKey'^{K_i}$ – функція додавання циклового ключа K_i за модулем 2^{64}, $MixColumns(MC)$ – лінійне перетворення (множення матриці лінійного перетворення на матрицю внутрішнього стану над скінченним полем), $ShiftRows(SR)$ – перестановка елементів $g_{i,j} \in GF(2^8)$ внутрішнього стану (циклічний зсув рядків вправо при матричному поданні), $SubBytes(S-box)$ – шар нелінійного бієктивного відображення, який виконує обробку векторів, заданих над V_8 (байтова підстановка), $AddKey^{K_i}$ – функція додавання циклового ключа K_i за модулем 2.

Надалі введемо позначення. Нехай $R^{(i)}$ – процедура зашифрування перших i раундів.

Більш детальну інформацію про структуру та особливості блокового шифру ДСТУ 7624:2014 «Калина» можна одержати у [4].

2. Ланцюги підпросторів ітеративних блокових шифрів, з розміром блоку 512-біт

Нехай F – раундова функція в ітеративному блоковому шифрі:

$$E_K(m) = k_n \oplus F(...F(k_1 \oplus F(k_0 \oplus m))),$$

де k_i – раундові ключі, отримані з основного ключа K за допомогою певного ключового розкладу. Вхідні повідомлення m розглядаються як бітові вектори із лінійного простору всіх бітових векторів відповідної довжини. Розглянемо пару підпросторів V_1 та V_2 таких, що для довільного вектору a існує

[a] kolyadamariya1710@gmail.com

унікальний вектор b (який залежить від a та ключа) такий, що повинне виконуватись співвідношення $F(V_1 \oplus a) \oplus K \subseteq V_2 \oplus b$, тобто F переводить кожен клас суміжності у якийсь інший клас суміжності.

Ланцюгом підпросторів довжини r назвемо простий кортеж з $r+1$ підпросторів $(V_1, V_2, \ldots, V_{r+1})$, для яких виконуються співвідношення:

$$F(V_i \oplus a_i) \oplus K \subseteq V_{i+1} \oplus a_{i+1}.$$

Позначимо через $E = \{e_{0,0}, \ldots, e_{8,8}\}$ – простір початкових станів шифру «Калина», де $e_{i,j}$ – окремі байти (8-бітові рядки). Визначимо чотири сімейства підпросторів E:

1) *Стовпчиковий простір* C_i визначимо як $C_i = \langle e_{0,i}, e_{1,i}, ..., e_{7,i} \rangle$, $i \in \{0, \ldots, 7\}$.
2) *Діагональний простір* D_i визначимо як $D_i = SR^{-1}(C_i) = \langle e_{7,i}, e_{6,i+1}, ..., e_{0,i+7} \rangle$, де індекс $i+j$ обчислюється за модулем 8, $i \in \{0, \ldots, 7\}$.
3) *Інверсно-діагональний простір* ID_i визначимо як $ID_i = SR(C_i) = \langle e_{0,i}, e_{1,i+1}, ..., e_{7,i+7} \rangle$, де індекс $i+j$ обчислюється за модулем 8, $i \in \{0, \ldots, 7\}$.
4) *Змішанний простір* M_i визначимо таким чином: $M_i = MC(ID_i)$, $i \in \{0, \ldots, 7\}$.

3. Схема атаки відновлення раундового ключа

Зазвичай, в безпечних моделях, зловмисник має чорний ящик (оракул) з доступом, наприклад, до шифруючої функції, котра пов'язана з випадковим секретним ключем, та до її зворотної функції. Мета зловмисника – відновлення раундового ключа. Блокові шифри, зазвичай, намагаються побудувати таким чином, що після останнього раунду шифрування відкритого тексту p_1 зловмисник не може відрізнити шифртекст c_1 від випадкової перестановки. Отже, всі шифртексти є рівноімовірними.

Для наведених вище просторів ми можемо спрогнозувати їх поведінку після декількох раундів їх зашифрування. Це означає, що ми можемо розраховувати статистики для отриманих шифртекстів. Таким чином, для деяких випадків, з певною ймовірністю, ми можемо відрізнити шифртекст від випадкової перестановки. Атака відновлення раундового ключа для $(i+1)$-го раунду виконується за наступним алгоритмом:

1) Генеруємо необхідну кількість пар (p_j, c_j^{i+1}), де p_j – спеціально підібраний відкритий текст, c_j^{i+1} – результат його за шифрування після $(i+1)$-го раунду.
2) Генеруємо випадковий ключ k_{i+1} та застосовуємо до c_j^{i+1} раундове перетворення з ключем k_{i+1}, позначимо $R^{i+1}_{k_{i+1}}(c_j^{i+1}) = c_j$. В результаті ми або отримаємо $c_j = c_j^i$(що означає, що k_{i+1} є раундовим ключем, для $(i+1)$-го раунду), або отримаємо $c_j = c_j^{i+2}$(що означає, що k_{i+1} не є раундовим ключем, для $(i+1)$-го раунду). Відрізнити ці гіпотези ми можемо за допомогою того, що нам відома статистика для шифр текстів c_j^i. Інакше кажучи, якщо отримані c_j – належать відомій нам статистиці, то ми вгадали раундовий ключ.
3) Повторюємо пункт 2, доки не отримаємо k_{i+1} – правильний раундовий ключ.

4. Ланцюги підпросторів шифру «Калина»

Опишемо властивості ланцюгів підпросторів для двох, чотирьох та п'яти раундів шифру «Калина-512».

1) Нехай $I \subseteq \{0, 1, ..., 7\}$, де $0 < |I| < 8$ та $a \in D_I^{\perp}$, тоді існує унікальне $b \in C_I^{\perp}$ таке, що $R_K(D_I \oplus a) = C_I \oplus b$.
2) Нехай $I \subseteq \{0, 1, ..., 7\}$, де $0 < |I| < 8$ та $a \in C_I^{\perp}$, тоді існує унікальне $b \in M_I^{\perp}$ таке, що $R_K(C_I \oplus a) = M_I \oplus b$.
3) $\Pr\left(R^{(2)}(u) \oplus R^{(2)}(v) \in M_i | u \oplus v \in D_I\right) = 1$, де $u \neq v$.
4) $\Pr\left(R^{(4)}(u) \oplus R^{(4)}(v) \in M_j | u \oplus v \in D_I\right) = 0$, де $u \neq v$.

Для наступних тверджень відмітимо:

Нехай, ненульовий елемент матриці розміру 8×8 знаходиться у рядку s та стовпчику c, та для r, c виконується: $s - c = i \mod 8$, для $n \leq 7$(залежить від розмірності D), тоді така матриця має i-ту діагональ.

Нехай, ненульовий елемент матриці розміру 8×8 знаходиться у рядку r та стовпчику c, та для r, c виконується: $r + c = i \mod 8$, для $n \leq 7$(залежить від розмірності D), тоді така матриця має i-ту антидіагональ.

Два тексти t_1 та t_2 належать одному і тому ж класу суміжності D, якщо байти їх різниці $t_1 \oplus t_2$, що лежать на n-діагоналі рівні нулю. Аналогічно, два тексти t_1 та t_2 належать одному і тому ж класу суміжності M, якщо байти їх різниці $MC^{-1}(t_1 \oplus t_2)$, що належать n-тій антидіагоналі для $n \leq 7$(залежить від розмірності M) рівні нулю.

Теорема. Візьмемо підпростори D_I та M_J для фіксованих I та J, де $|I| = 1$. Візьмемо який-небудь клас суміжності D_I такий, що є $D_I \oplus a$ для фіксованого $a \in D_I^{\perp}$, що містить всі $(2^8)^8$ відкритих текстів та пов'язані з ними після 5-го раунду шифртексти, тобто такі пари (p^i, c^i), для $i = 0, \ldots, (2^8)^{8|I|} - 1$, де $p^i \in D_I \oplus a$ та $C^i = R^5(p^i)$. Нехай n – кількість різних пар шифртекстів (c^i, c^j), $i \neq j$ такі, що $c^i \oplus c^j \in M_J$, тобто

$$n := |\{(p^i, c^i), (p^j, c^j) \mid \forall p^i, p^j \in D_I \oplus a, \\ p^i < p^j, c^i \oplus c^j \in M_J\}|.$$

Тоді n ділиться на 128.

Ідея доведення: Як ми бачили вище, клас суміжності D_I завжди відображається в клас суміжності M_I після 2-го раунду, тобто для кожного $a \in D_I^{\perp}$ існує єдине $b \in M_I^{\perp}$ таке, що $R^2(D_I \oplus a) = M_I \oplus b$. Має місце і зворотне твердження, тобто для довільного $b' \in M_I^{\perp}$ існує єдине $a' \in D_I^{\perp}$ таке, що $R^{-2}(M_I \oplus b') = D_I \oplus a'$. Отримуємо ланцюг:

$$D_I \oplus a \xrightarrow[prob.1]{R^2(.)} M_I \oplus b \xrightarrow{R(.)} D_J \oplus a' \xrightarrow[prob.1]{R^2(.)} M_J \oplus b'.$$

Ідея аналізу цього ланцюга підпросторів полягає в

обґрунтуванні центрального переходу $M_I \oplus b \xrightarrow{R(.)} D_J \oplus a'$.

Лема. Візьмемо D_J та M_I підпростори визначені вище для фіксованих I та J де $|I| = 1$. Візьмемо довільний клас суміжності M_I , розглянемо усі $(2^8)^8$ відкритих текстів та відповідні їм шифртексти після одного раунду зашифрування, тобто $(\hat{p^i}, \hat{c^i})$ для $i = 0, ..., (2^8)^8 - 1$, де $\hat{c^i} = R(\hat{p^i})$. Кількість різних пар шифртекстів $(\hat{c^i}, \hat{c^j})$ для $i \neq j$ такі, що $\hat{c^i} \oplus \hat{c^j} \in D_J$ (вони належать одному класу суміжності D_J) кратна 128, тобто існує n' таке, що $n = n' \cdot 8$.

Доведення: Візьмемо два елементи p_1 та p_2 з одного класу суміжності $M_i \oplus a$, для $a \in M_i^{\perp}$. Без втрати загальності будемо вважати, $i = 0$. За означенням M_i, можна представити як вектор $\{x_0, x_1, ..., x_7\}$ помножений на відому нам матрицю, тобто кожний стовпчик M_i залежить від відповідної координати вектора. Нехай елементу p_1 буде відповідати вектор $\{x_0, x_1, ..., x_7\}$, а елементу p_2 буде відповідати вектор $\{x'_0, x'_1, ..., x'_7\}$. Тоді можливі випадки, коли елементи p_1 та p_2 мають одну, дві, три, чотири, п'ять, шість, сім та вісім нерівних коорднат. Розглянемо найпростіші з них.

Перший випадок. Розглянемо випадок, коли у елементів p_1 та p_2 лише одна нерівна координата, тобто два тексти p_1 та p_2 належать одному і тому класу суміжності $M_0 \cap C_0 \oplus a$, де $a \in (M_0 \cap C_0)^{\perp}$. Оскільки $M_0 \cap C_0 \subseteq C_0$, з цього випливає, що якщо $p_1 \oplus p_2 \in C_0$, то $R(p_1) \oplus R(p_2) \in M_0$. Оскільки $M_I \cap D_I = \{0\}$, для будь-яких I та J таких, що $|I| + |J| \leq 8$, з цього випливає що $R(p_1) \oplus R(p_2) \notin D_J$, для будь-яких J таких, що $|J| \leq 7$. Іншими словами, якщо прийняти дану гіпотезу, то два тексти ніколи не потраплять в той самий клас суміжності D_J після одного раунду зашифрування.

Другий випадок. Розглянемо випадок, коли у елементів p_1 та p_2 нерівні між собою дві координати, тобто два тексти p_1 та p_2 належать одному і тому класу суміжності $M_0 \cap C_0 \oplus a$, де $a \in (M_0 \cap C_0)^{\perp}$. Нехай ці два елементи, після одного раунду зашифрування потрапили в один і той самий клас суміжності D_J, де $|J| = 7$. Інакше кажучи, нехай $j = \{0, ..., 7\}$ та існують такі координати x_0, x_1 та x'_0, x'_1 та j такі, що для елементів p_1 та p_2 виконується: $(R(p_1) \oplus R(p_2))_i, i - j = 0$ для будь-яких $i = 0, 1, ..., 7$, де сумма береться за модулем 8. Надалі візьмемо два тексти $\hat{p_1}$ та $\hat{p_1}$, котрі можна отримати, якщо у елементів p_1 та p_2 змінити місцями координати x_0 та x'_0 відповідно. Тоді, після одного раунду зашифрування елементи $\hat{p_1}$ та $\hat{p_1}$ також потраплять у клас суміжності D_J, оскільки $R(p_1) \oplus R(p_2) = R(\hat{p_1}) \oplus R(\hat{p_2})$. А отже, ми можемо стверждувати, що кількість пар шифртекстів, що потрапили до класу суміжності D_J буде кратна двом.

Інші випадки розглядаються аналогічно. В результаті, буде доведено, що у випадку, коли у елементів p_1 та p_2 всі координати різні, то можна побудувати 2^7 пар елементів, для котрих буде виконуватися $R(p_1) \oplus R(p_2) = R(\hat{p_1}) \oplus R(\hat{p_2})$. А отже, кількість пар шифртекстів, що потрапили до класу суміжності D_J буде кратна $2^7 = 128$.

Висновки

Отже, в результаті даної роботи було запропоновано та проведено аналіз ланцюгів підпросторів для перших п'яти раундів для «Калина»-подібних шифрів, які можна використати для побудови атаки розпізнавання на шестираундовий шифр «Калина». В майбутньому ми плануємо отримати ланцюги підпросторів для вхідних даних, що подаються у вигляді не квадратної матриці 8×8, а у вигляді прямокутних матриць розмірами 4×8 та 2×8 для «Калина»-подібних шифрів.

Перелік використаних джерел

1. Leander Gregor, Abdelraheem Mohamed Ahmed, AlKhzaimi Hoda, Zenner Erik. A Cryptanalysis of PRINTcipher: The Invariant Subspace Attack. — 2011.
2. Grassi Lorenzo, Rechberger Christian, Rønjom Sondre. Subspace Trail Cryptanalysis and its Applications to AES. — 2016. — Access mode: `http://eprint.iacr.org/2016/592`.
3. Grassi Lorenzo, Rechberger Christian, Rønjom Sondre. New Structural-Differential Property of 5-Round AES. — 2017. — Access mode: `http://eprint.iacr.org/2017/118`.
4. Oliynykov Roman, Gorbenko Ivan, Kazymyrov Oleksandr. A New Encryption Standard of Ukraine: The Kalyna Block Cipher. — 2015. — Access mode: `http://eprint.iacr.org/2015/650`.
5. Горбенко І.Д., Тоцький О.С., ін. С.В. Казьміна та. Перспективний блоковий шифр "Калина" – основні положення та специфікація. — 2007.