УДК 621.391:519.2

А. М. Олексійчук, д-р. техн. наук,
С. М. Ігнатенко, аспірант,
М. В. Поремський, аспірант

Інститут спеціального зв'язку та захисту інформації
Національний технічний університет України
«Київський політехнічний інститут імені І. Сікорського», м. Київ

СИСТЕМИ ЛІНІЙНИХ РІВНЯНЬ ЗІ СПОТВОРЕНИМИ ПРАВИМИ ЧАСТИНАМИ НАД СКІНЧЕННИМИ КІЛЬЦЯМИ

З метою побудови кореляційних атак на сучасні словоорієнтовані потокові шифри досліджуються методи розв'язання систем лінійних рівнянь зі спотвореними правими частинами над довільними скінченними кільцями. Отримано узагальнення й уточнення низки раніше відомих результатів стосовно методів розв'язання зазначених систем рівнянь над полями чи кільцями лишків порядку 2^r.

Ключові слова: *кореляційний криптоаналіз, система лінійних рівнянь зі спотвореними правими частинами, метод максимуму правдоподібності, задача про адитивне k-представлення, алгоритм BKW.*

Вступ. Нехай R — скінченне (асоціативне) кільце з одиницею, $|R| = q$. Розглянемо систему рівнянь (СР) зі спотвореними правими частинами

$$Ax = b, \qquad (1)$$

де A — $m \times n$-матриця над кільцем R, b — вектор довжини m з координатами $b_i = A_i a + \xi_i$, $i = \overline{1,m}$, де $A_1, \dots, A_m$ — рядки матриці A, $a = (a_1, \dots, a_n)^{\mathrm{T}}$ — невідомий вектор-стовпець над кільцем R (істинний розв'язок СР (1)), $\xi_1, \dots, \xi_m$ — незалежні випадкові величини, розподілені за законом $P\{\xi_i = z\} = p(z)$, де $p(z) \ge 0$ для кожного $z \in R$, $\sum\limits_{z \in R} p(z) = 1$. Задача розв'язання СР (1) полягає у відновленні вектора a за відомими матрицею A, вектором b і розподілом ймовірностей $p_\xi = (p(z) : z \in R)$.

До розв'язання цієї задачі приводить, зокрема, побудова кореляційних атак на синхронні потокові шифри, причому, як правило, R є полем з двох елементів [1, 2]. Методи вирішення зазначеної задачі у випадках $R = GF(2^r)$ та $R = Z/(2^r)$, де $r \ge 2$, викладені в роботах [3, 4] і [5, 6] відповідно.

Метою статті є узагальнення та уточнення окремих результатів робіт [4, 6]. Враховуючи обмеження щодо обсягу статті, ми не наводимо тут доведення отриманих теорем.

1. Оцінка кількості рівнянь, необхідних для успішного розв'язання СР (1) із заданою ймовірністю помилки. Припустимо, що матриця A є фіксованою. В цьому випадку будь-який алгоритм відновлення вектора a з системи рівнянь (1) задається певним відображенням $D_A : R^m \to R^n$, яке ставить у відповідність вектору b з координатами (2) «оцінку» вектора a. При цьому (середня) ймовірність помилки алгоритму D_A визначається за формулою $\delta(D_A) = q^{-n} \sum_{a \in R^n} P\{D_A(b) \neq a\}$.

Наступна теорема уточнює теорему 5 роботи [4], яка містить евристичну оцінку числа рівнянь, необхідних для надійного розв'язання СР (1) над полем з 2^r елементів.

Теорема 1. Припустимо, що відображення $x \mapsto Ax$, $x \in R^n$ є ін'єктивним, і m є найменшим числом рівнянь у системі (1), для якого існує алгоритм її розв'язання з ймовірністю помилки не більше ніж $\delta \in (0, 1/2)$. Тоді

$$m \geq \frac{(1-\delta)n \log q - h(\delta)}{\Delta(p_\xi)} \ln 2, \qquad (2)$$

де $\Delta(p_\xi) = q^{-1} \sum_{z \in R} (qp(z) - 1)^2$, $h(\delta) = -\delta \log_2 \delta - (1-\delta) \log_2 (1-\delta)$.

Зауважимо, що на відміну від теореми 5 в [4], нерівність (2) містить явну залежність параметра m від параметра δ та не базується на будь-яких евристичних припущеннях.

2. Оцінка ймовірності відновлення істинного розв'язку СР (1) методом максимуму правдоподібності. Для будь-якого $x \in R^n$ позначимо $\varepsilon(x) = b - Ax$. Розв'язання СР (1) методом максимуму правдоподібності полягає в знаходженні «оцінки» a^* вектора a за правилом $P\{\xi = \varepsilon(a^*)\} = \max_{x \in R^n} P\{\xi = \varepsilon(x)\}$, де $\xi = (\xi_1, \ldots, \xi_m)$. Якщо вектор a є рівномірно розподіленим на множині R^n, то метод максимуму правдоподібності має найменшу (середню) ймовірність помилки серед усіх методів розв'язання СР (1) (див., наприклад, [7, с. 141].

Аналітичні оцінки ймовірності правильного відновлення істинного розв'язку СР (1) методом максимуму правдоподібності отримані в [8] для випадку $R = GF(2)$ та в [6] для загального випадку. Наступна теорема підсилює основний результат роботи [6].

Теорема 2. Нехай матриця A має рівномірний розподіл ймовірностей на множині усіх матриць розміру $m \times n$ над кільцем R та не залежить від випадкового вектора $\xi = (\xi_1, \dots, \xi_m)$. Позначимо $N_\xi(R) = \{z \in R : p(z) > 0\}$, $p_{\max} = \max\limits_{z \in R} p(z)$, $p_{\min} = \min\limits_{z \in N_\xi(R)} p(z)$ та припустимо, що $p_{\max} \neq p_{\min}$. Тоді для будь-якого $a \in R^n$ справедлива нерівність

$$P_{A,\xi}\{a^* = a\} \geq 1 - q^n \exp\left\{ - \frac{m(D(p_\xi \,\|\, \omega) + D(\omega \,\|\, p_\xi))^2}{2(\log p_{\max} - \log p_{\min})^2} \right\},$$

де

$$D(p_\xi \,\|\, \omega) = \sum_{z \in N_\xi(R)} p(z) \log(qp(z)),\ D(\omega \,\|\, p_\xi) = -q^{-1} \sum_{z \in N_\xi(R)} \log(qp(z)).$$

Отже, за умови теореми 2 метод максимуму правдоподібності надає можливість відновити істинний розв'язок СР (1) з ймовірністю помилки не більше ніж $\delta \in (0, 1)$ (відносно сумісного розподілу матриці A та вектора ξ), якщо кількість рівнянь у системі задовольняє нерівності

$$m \geq \frac{2n \ln(q\delta^{-1})(\log p_{\max} - \log p_{\min})^2}{(D(p_\xi \,\|\, \omega) + D(\omega \,\|\, p_\xi))^2}. \qquad (3)$$

3. Субекспоненційні алгоритми розв'язання СР (1). Найефективніші на сьогодні алгоритми розв'язання систем лінійних рівнянь зі спотвореними правими частинами та випадковими рівноймовірними матрицями коефіцієнтів над полем $GF(2)$ мають субекспоненційну часову складність і, як правило, базуються на розв'язанні задачі про адитивне представлення (див. роботи [9, 10] та наведені там посилання).

Для випадку довільного скінченного кільця R остання задача має таке формулювання. Задано список L, що складається з l випадкових незалежних та рівноймовірних векторів $z_1, \dots, z_l \in R^n$. Потрібно знайти в цьому списку адитивне k-представлення нульового вектора, тобто k (не обов'язково різних) номерів $\nu_1, \dots, \nu_k \in \{1, 2, \dots, l\}$ таких, що $z_{\nu_1} \oplus \cdots \oplus z_{\nu_k} = 0$.

Ефективні алгоритми розв'язання задачі про адитивне представлення відомі для випадків $R = GF(2)$ [11, 12] та $R = GF(2^r)$, де $r \geq 2$ [13]. Наступна теорема узагальнює твердження про властивості одного з таких алгоритмів: алгоритму BKW [11].

Теорема 3. Нехай u, v, λ — натуральні числа і $n \leq uv$, $k = 2^{u-1}$, $l = (u + \lambda - 1)q^v$. Тоді існує алгоритм, який знаходить адитивне k-пред-

ставлення нульового вектора у випадковому рівноймовірному списку L довжини l з ймовірністю не менше ніж $1-e^{-\lambda}$, використовуючи $O(u(u+\lambda)q^v)$ операцій над n-вимірними векторами над кільцем R.

Зауважимо, що алгоритм, зазначений у формулюванні теореми 3, не відрізняється за сутністю від класичного алгоритму BKW.

Можливість ефективного розв'язання задачі про адитивне представлення дозволяє узагальнити низку відомих субекспоненційних алгоритмів розв'язання СР (1) над полями порядку 2^r на системи лінійних рівнянь зі спотвореними правими частинами над довільним скінченним кільцем.

Розглянемо докладніше один з таких алгоритмів, який є безпосереднім узагальненням алгоритму з [4].

Алгоритм B, що пропонується, залежить від натуральних параметрів n_1, l, t, де $1 \le n_1 \le n-1$, $m \ge lt$, та допоміжного алгоритму А розв'язання задачі про адитивне представлення з параметрами $n-n_1, k, l$ і за певних умов дозволяє відновлювати перші n_1 координат істинного розв'язку СР (1).

Для будь-якого $z \in R^n$ позначимо z' та z'' підвектори вектора z, що складаються з його перших n_1 та останніх $n-n_1$ координат відповідно. Запишемо СР (1) у вигляді $A_i'x' \oplus A_i''x'' = b_i$, $i = \overline{1,m}$.

Алгоритм B має такий вигляд.

1. Розіб'ємо систему рядків $A_1'', \dots, A_m''$ на t списків L_j довжини l кожний та застосуємо для кожного $j = \overline{1,t}$ алгоритм А до списку L_j. Якщо хоча б в одному випадку алгоритм А завершується неуспішно, то алгоритм B припиняє роботу. Інакше отримаємо рівності вигляду $A''_{\nu_1(j)} + \dots + A''_{\nu_k(j)} = 0$, де $A''_{\nu_1(j)}, \dots, A''_{\nu_k(j)} \in L_j$, $j = \overline{1,t}$.

2. Складемо СР зі спотвореними правими частинами

$$A'(j)x' = b(j), \ j = \overline{1,t},$$

де

$$A'(j) = A'_{\nu_1(j)} + \dots + A'_{\nu_k(j)},$$

$$b(j) = b_{\nu_1(j)} + \dots + b_{\nu_k(j)} = A'(j)a' + (\xi_{\nu_1(j)} + \dots + \xi_{\nu_k(j)}),$$

та розв'яжемо її методом максимуму правдоподібності.

Теорема 4. Нехай матриця A має рівномірний розподіл ймовірностей на множині усіх матриць розміру $m \times n$ над кільцем R та не залежить від випадкового вектора $\xi = (\xi_1, \dots, \xi_m)$. Нехай, далі,

$$u = \left\lceil \frac{\log(n-n_1)}{2} \right\rceil,\ v = \left\lceil \frac{2(n-n_1)}{\log(n-n_1)} \right\rceil,$$

$$k = 2^{u-1},\ l = (u\ +\ \left\lceil \ln(2t\delta^{-1}) \right\rceil\ -\ 1)q^v,\ m \geq lt,$$

де t задається виразом у правій частині нерівності (3) з заміною n на n_1, δ на $\delta/2$, а розподілу p_ξ на його k-й степінь відносно операції згортки розподілів ймовірностей на адитивній групі кільця R. Тоді, використовуючи в ролі A алгоритм, зазначений у формулюванні теореми 3, можна відновити з ймовірністю не менше ніж $1-\delta$ перші n_1 координат вектора a за допомогою алгоритму B, використовуючи $2n_1tq^{n_1} + O(ult)$ операцій над n-вимірними векторами над кільцем R.

Зокрема, при фіксованих q, n_1 та $n \to \infty$ трудомісткість алгоритму B складає $O(tq^{\frac{2n}{\log n}} \log t \log n)$ зазначених операцій.

Висновки. Отримані результати показують, що відомі методи та алгоритми розв'язання систем лінійних рівнянь зі спотвореними правими частинами над кільцями лишків або полями порядку 2^r допускають узагальнення на випадок систем над довільними скінченними кільцями. Зокрема, це відноситься до найкращих на сьогодні субекспоненційних алгоритмів розв'язання систем лінійних рівнянь зі спотвореними правими частинами над полем з 2^r елементів [4, 10]. Теореми 2 та 3 підсилюють раніше відомі оцінки кількості рівнянь, що необхідні та, відповідно, достатні для розв'язання СР вигляду (1) із заданою ймовірністю помилки і можуть бути використані для оцінювання стійкості сучасних словоорієнтованих потокових шифрів відносно кореляційних атак.

Розробка методу обґрунтування стійкості зазначених шифрів відносно атак, що базуються на розв'язання СР (1) над кільцями порядку $q > 2$, є предметом подальших досліджень.

Список використаних джерел:

1. Canteaut A. Fast correlation attacks against stream ciphers and related open problems. *The 2005 IEEE Information Theory Workshop on Theory and Practice in Information-Theoretic Security.* ITW 2005, E-Proc. 2005. P. 49–54.
2. Meier W. Fast correlation attacks: methods and countermeasures. *Lecture Notes in Computer Science — zaFSE'2011, Proceedings.* Springer Verlag, 2011. P. 55–67.
3. Johansson T., Jonsson F. Correlation attacks on stream ciphers over GF(2^n). *The 2001 International Symposium on Information Theory — ISIT'2001, Proceedings.* Springer Verlag, 2001. P. 140.

4. Zhang B., Xu C., Meier W. Fast correlation attacks over extension fields, large-unit linear approximation and cryptanalysis of SNOW 2.0. Cryptology ePrint Archive, Report 2016/311. http://eprint.iacr.org/2016/311.
5. Алексейчук А. Н., Игнатенко С. М. Метод оптимизации алгоритмов решения систем линейных уравнений с искаженной правой частью над кольцом вычетов по модулю 2^N. *Реєстрація, зберігання і обробка даних.* 2005. Т. 7. № 1. С. 21–29.
6. Алексейчук А. Н., Игнатенко С. М. Нижняя граница вероятности восстановления истинного решения системы линейных уравнений с искаженной правой частью над кольцом вычетов по модулю 2^N. *Захист інформації.* 2006. № 4. С. 6–12.
7. Чечёта С. И. Введение в дискретную теорию информации и кодирования: учебное издание. М.: МЦНМО, 2011. 224 с.
8. Балакин Г. В. Введение в теорию случайных систем уравнений. Труды по дискретной математике. М.: ТВП, 1997. Т. 1. С. 1–18.
9. Олексійчук А. М. Субекспоненційні алгоритми розв'язання систем лінійних булевих рівнянь зі спотвореними правими частинами. *Прикладная радиоэлектроника.* 2012. Т. 11. № 2. С. 3–11.
10. Bogos S., Tram′er F., Vaudenay S. On solving LPN using BKW and variants. Implementation and analysis. Cryptology ePrint Archive, Report 2015/049. http://eprint.iacr.org/2015/049.
11. Blum A., Kalai A., Wasserman H. Noise-tolerant learning, the parity problem, and the statistical query model. J. ACM. 2003. Vol. 50. N 3. P. 506–519.
12. Wagner D. A generalized birthday problem. *Advances in Cryptology — CRYPTO'02, Proceedings.* Springer Verlag, 2002. P. 288–303.
13. Minder L., Sinclair A. The extended k-tree algorithm. *The 19th Annual ACM-SIAM Sympposium on Discrete Algorithms, Proceedings.* 2009. P. 586–595.

In order to build correlation attacks on modern word-oriented stream ciphers, methods for solving systems of linear equations corrupted by noise over arbitrary finite rings are investigated. Generalizations and refinements of earlier known results about methods for solving such systems of equations over the fields or residue rings of order 2^r are obtained.

Key words: *correlation cryptanalysis, system of linear equations corrupted by noise, maximum likelihood method, k-sum problem, BKW algorithm.*

Одержано 30.01.2017