УДК 621.391:519.2
DOI: 10.32626/2308-5916.2019-19.88-94

С. В. Мітін

Національного технічного університету України
«Київський політехнічний інститут імені Ігоря Сікорського», м. Київ

ЗАСТОСУВАННЯ АЛГОРИТМУ BKW ДЛЯ ВІДНОВЛЕННЯ СИСТЕМАТИЧНИХ ЛІНІЙНИХ БЛОКОВИХ КОДІВ ЗА НАБОРАМИ СПОТВОРЕНИХ КОДОВИХ СЛІВ

Важливою практичною задачею у галузі інформаційної безпеки є розробка методів відновлення дискретних відображень, які використовуються в сучасних системах передачі, обробки та зберігання даних, за наборами спотворених значень цих відображень, що отримуються під впливом шумів (випадкових спотворень, навмисних перешкод, внутрішніх збоїв тощо). При розв'язанні цієї задачі додаткові складнощі виникають у разі відсутності повних відомостей про алгоритми, що визначають зазначені відображення, та використовуються для перетворення інформації. Окремим випадком поставленої задачі є відновлення систематичних лінійних блокових кодів з невідомими твірними матрицями за наборами спотворених кодових слів, що спостерігаються на виході двійкового симетричного каналу зв'язку. У даній статті запропоновано метод розв'язання останньої задачі, який базується на застосуванні алгоритму BKW, що використовується при побудові кореляційних атак на потокові шифри. Алгоритм застосовується для розв'язання не однієї, а (одночасно) багатьох систем лінійних рівнянь зі спотвореними правими частинами шляхом одноразового перетворення їх спільної матриці коефіцієнтів. Наведено обґрунтування коректності та отримано оцінку ефективності запропонованого методу. Здійснено його порівняння з раніше відомим методом. Показано, що запропонований метод має більшу ефективність за трудомісткістю та обсягом потрібної пам'яті, хоча й потребує більше спотворених кодових слів, які необхідні для відновлення твірної матриці коду. В залежності від параметрів кодів, що відновлюються, та ймовірності спотворення у каналі зв'язку, виграш у трудомісткості запропонованого методу в порівнянні з раніше відомим складає приблизно від 2^{36} до 2^{67} разів. Підтверджено також практичну застосовність запропонованого методу для випадків, коли раніше відомий метод є практично не реалізованим.

Ключові слова: *інформаційна безпека, вивідування інформації, відновлення дискретних відображень, лінійний блоковий код, система рівнянь зі спотвореними правими частинами, алгоритм BKW.*

© С. В. Мітін, 2019

Вступ. Важливою практичною задачею є відновлення невідомого лінійного блокового коду за набором спотворених кодових слів. Ця задача є NP-повною [1], а відомі алгоритми її розв'язання є практично застосовними лише у випадку помірної довжини кодів, що відновлюються, або малої ймовірності спотворень символів у каналі зв'язку [1, 2].

В роботі [3] запропоновано метод вирішення зазначеної задачі для систематичних лінійних блокових кодів з використанням методу максимуму правдоподібності, який за відомих умов характеризується найменшою ймовірністю помилки [4].

Метою даної статті є розробка методу, який удосконалює метод роботи [3] і має в порівнянні з ним більш високу ефективність.

Постановка задачі та основні результати. Використовуватиме такі позначення: C — невідомий двійковий лінійний (n, k)-код з твірною матрицею $G = (I_k, X)$, де I_k — одинична матриця порядку k, X — матриця розміру $k \times (n-k)$ над полем з двох елементів.

Нехай x_j j-й стовпець матриці X. Припустимо, що вага (число ненульових координат) вектора x_j знаходиться в межах від 3 до ρ, де ρ — ціле число, $\rho \geq 3$, $j \in \overline{1, n-k}$. Треба відновити матрицю X за набором m незалежних випадкових рівноймовірних слів коду C, спотворених у двійковому симетричному каналі (ДСК) зв'язку з імовірністю помилки $p \in (0, 1/2)$.

Поставлену задачу можна формулювати також наступним чином. Розглянемо лінійне відображення $L_C : \{0,1\}^k \to \{0,1\}^n$, що задається кодом C зазначеного вигляду: $L_C(u) = uG$, $u \in \{0,1\}^k$. Спостерігається послідовність, яка складається з m спотворених значень цього відображення:

$$Y_i = U_i G \oplus \eta_i, \ i \in \overline{1, m}, \qquad (1)$$

де U_i — незалежні випадкові рівноймовірні вектори довжини k, $\eta_i = (\eta_{i,1}, \ldots, \eta_{i,n})$ — вектори спотворень, координати яких не залежать від $U_1, \ldots, U_m$ та є незалежними в сукупності випадковими величинами, що розподілені за законом

$$\mathbf{P}\{\eta_{i,s} = 1\} = 1 - \mathbf{P}\{\eta_{i,s} = 0\} \leq p \in (0, 1/2),$$

$s \in \overline{1, n}$, $i \in \overline{1, m}$. Треба відновити відображення L_C (тобто матрицю X) за відомими значеннями n, k, p, ρ та послідовністю (1).

Метод розв'язання поставленої задачі запропоновано в роботі [3]. Сутність методу полягає у складанні та розв'язанні за допомо-

гою методу максимуму правдоподібності [4] систем лінійних рівнянь (СЛР) із спотвореними правими частинами

$$Ax = b^{(j)} = Ax_j \oplus \xi^{(j)}, \ j \in \overline{1, n-k}, \tag{2}$$

де A — відома реалізація випадкової рівноймовірної двійкової матриці розміру $m \times k$, $\xi^{(j)} = (\xi_{1,j}, \dots, \xi_{m,j})^{\mathrm{T}}$ — випадковий вектор з незалежними координатами, що розподілені за законом

$$\mathbf{P}\{\xi_{i,j} = 1\} = 1 - \mathbf{P}\{\xi_{i,j} = 0\} \le \tilde{p} = 1/2 \cdot (1 - (1-2p)^{\rho+1}), \tag{3}$$

$i \in \overline{1, m}$, $j \in \overline{1, n-k}$. В роботі [3] показано, що для відновлення матриці X з імовірністю не менше $1-\delta$, $\delta \in (0, 1/2)$, потрібно не більше ніж

$$m_1 = \left\lceil 8(1-2\tilde{p})^{-2} \ln\left((n-k)\delta^{-1} \sum_{i=3}^{\rho} \binom{k}{i} \right) \right\rceil \tag{4}$$

спотворених кодових слів. При цьому трудомісткість методу складає

$$T_1 = O\left(m_1 (n-k)(\rho+1) \sum_{i=3}^{\rho} \binom{k}{i} \right) \tag{5}$$

операцій.

Далі викладено метод відновлення матриці X, який удосконалює метод роботи [3] і дозволяє суттєво підвищити його ефективність шляхом одночасного розв'язання усіх СЛР (2) із застосуванням модифікованого алгоритму BKW [5].

Алгоритм реалізації методу (алгоритм **B**) залежить від натуральних параметрів k_1, l, t, де $1 \le k_1 \le k-3$, $m \ge lt$, та допоміжного алгоритму **A** розв'язання задачі про адитивне представлення з параметрами $k-k_1, r, l$. Останній являє собою відомий алгоритм BKW [6] і дозволяє знаходити для довільного списку L, що складається з l випадкових незалежних рівноймовірних $(k-k_1)$-вимірних двійкових векторів $z_1, \dots, z_l$ r (не обов'язково різних) номерів $\nu_1, \dots, \nu_r \in \{1, 2, \dots, l\}$ таких, що $z_{\nu_1} + \dots + z_{\nu_r} = 0$.

Алгоритм **B** складається з двох етапів і дозволяє відновлювати перші k_1 координат усіх стовпців матриці X з ймовірністю помилки не вище ніж $\delta' = \delta \lceil k/k_1 \rceil^{-1}$. Застосовуючи цей алгоритм $\lceil k/k_1 \rceil$ разів до наборів координат (зазначених стовпців), що попарно не перетинаються, можна відновити матрицю X в цілому з ймовірністю помилки не вище ніж δ.

Для будь-якого $z \in R^k$ позначимо z' та z'' підвектори вектора z, що складаються з його перших k_1 та останніх $k-k_1$ координат

відповідно. Далі, позначимо A_i i-й рядок матриці A, b_{ij} — i-ту координату вектора $b^{(j)}$ та запишемо системи рівнянь (2) у вигляді

$$A'_i x'_j \oplus A''_i x''_j = b_{ij},\ i = \overline{1,m},\ j \in \overline{1,n-k}.$$

Алгоритм В має такий вигляд.

1. Розіб'ємо систему рядків $A'_1, \dots, A''_m$ на t списків L_s довжини l кожний та застосуємо для кожного $s \in \overline{1,t}$ алгоритм **А** до списку L_s. Якщо хоча б в одному випадку алгоритм **А** завершується неуспішно, то алгоритм **В** припиняє роботу. Інакше отримаємо рівності вигляду $A''_{\nu_1(s)} \oplus \dots \oplus A''_{\nu_r(s)} = 0$, де $A''_{\nu_1(s)}, \dots, A''_{\nu_r(s)} \in L_s$, $s \in \overline{1,t}$.

2. Складемо $n - k$ СЛР із спотвореними правими частинами

$$A'(s)x'_j = b(s,j),\ s \in \overline{1,t},\ j \in \overline{1,n-k}, \qquad (6)$$

де

$$A'(s) = A'_{\nu_1(s)} \oplus \dots \oplus A'_{\nu_r(s)},\ b(s,j) = b_{\nu_1(s),j} \oplus \dots \oplus b_{\nu_r(s),j} =$$
$$= A'(s)x'_j \oplus (\xi_{\nu_1(s),j} \oplus \dots \oplus \xi_{\nu_r(s),j})$$

та розв'яжемо їх методом максимуму правдоподібності.

Усі системи рівнянь (6) мають однакову матрицю коефіцієнтів, яка складається з t рядків $A'(s)$ довжини k_1, $s \in \overline{1,t}$. Внаслідок незалежності випадкових величин $\xi_{i,j}$, $i \in \overline{1,m}$, $j \in \overline{1,n-k}$, та формули (3), спотворення у правих частинах рівнянь системи (6) є незалежними випадковими величинами, розподіленими за законом

$$\mathbf{P}\{\xi_{\nu_1(s),j} \oplus \dots \oplus \xi_{\nu_r(s),j} = 0\} =$$
$$= 1 - \mathbf{P}\{\xi_{\nu_1(s),j} \oplus \dots \oplus \xi_{\nu_r(s),j} = 1\} \le 1/2 \cdot (1 - (1-2p)^{(\rho+1)r}), \qquad (7)$$

де $s \in \overline{1,t}$, $j \in \overline{1,n-k}$.

Для обґрунтування коректності та оцінки ефективності запропонованого методу сформулюємо наступну теорему (доведення якої виходить за межі статті).

Теорема. Нехай $\delta' \in (0, 1/2)$, $k_1 \in \overline{1,k-3}$ і параметри алгоритму **В** визначаються наступним чином:

$$u = \left\lceil \frac{\log(k-k_1)}{2} \right\rceil,\ v = \left\lceil \frac{2(k-k_1)}{\log(k-k_1)} \right\rceil,\ r = 2^{u-1},$$

$$t = \left\lceil 2(1-2p)^{-2r(\rho+1)} \ln\left(2(n-k)(\delta')^{-1} \sum_{i=0}^{\rho} \binom{k_1}{i} \right) \right\rceil,$$

$$l = (u + \left\lceil \ln(2t(\delta')^{-1}) \right\rceil - 1)2^v.$$

Алгоритм відновлює перші k_1 координат усіх стовпців матриці X з ймовірністю помилки не вище ніж δ', використовуючи

$$T(k_1) = O\left(ult + (n-k)(\rho+1)t\sum_{i=0}^{\rho}\binom{k_1}{i} \right), \qquad (8)$$

операцій, $m(k_1) = lt$ спотворених кодових слів та $N(k_1) = O(l + t)$ біт пам'яті.

Зауважимо, що, згідно з формулою (8), трудомісткість (при заданій верхній межі ймовірності помилки) алгоритму **В** залежить від допоміжного параметра $k_1 \in \overline{1, k-3}$, який слід вибирати, виходячи з умови $T(k^*) = \min\{T(k_1) : k_1 \in \overline{1, k-3}\}$. Тоді обсяг пам'яті та число спотворених кодових слів, потрібних для виконання алгоритму, складає відповідно $N(k^*) = l(k^*) + t(k^*)$ та $m(k^*) = l(k^*)t(k^*)$.

В таблиці наведені значення (двійкових) логарифмів трудомісткості, обсягу пам'яті та числа спотворених кодових слів, яких достатньо для відновлення матриці X з імовірністю не менше $1 - \delta$ за допомогою запропонованого методу. Значення t, l та m отримані з використанням теореми. В останніх двох колонках таблиці наведені значення параметрів (4), (5).

Таблиця

Чисельні значення параметрів, що характеризують ефективність методів відновлення систематичних лінійних блокових кодів ($n = 128$, $k = 80$, $\delta = \delta' = 0{,}1$)

p	ρ	**Запропонований метод**				**Метод роботи [3]**	
		k^*	$\log T(k^*)$	$\log N(k^*)$	$\log m(k^*)$	$\log T_1$	m_1
0,1000	20	18	88,42	59,36	85,91	94,29	4632991
	30	76	108,75	26,78	35,24	113,42	469177075
	50	76	126,99	39,73	48,69	133,03	3781005896745
0,0500	20	16	59,30	30,79	57,46	87,15	32920
	30	16	71,87	42,86	69,99	102,89	316141
	50	16	96,79	67,18	94,88	115,69	22911318
0,0300	20	16	48,35	26,27	46,41	84,52	5300
	30	16	55,88	27,99	53,89	99,00	21330
	50	16	70,72	41,58	68,68	109,29	271485
0,0100	20	16	37,76	25,58	35,65	81,99	921
	30	16	40,42	25,81	38,20	95,27	1611
	50	16	45,54	26,09	43,15	103,16	3871
0,0010	20	16	33,09	25,17	30,84	80,89	429
	30	16	33,64	25,32	31,23	93,64	521
	50	16	34,45	25,32	31,67	100,48	605
0,0001	20	14	32,63	25,32	30,49	80,78	398
	30	14	32,73	25,32	30,53	93,48	466
	50	14	32,95	25,32	30,64	100,22	504

Висновки. Отримані результати показують, що виграш у трудомісткості запропонованого методу в порівнянні з раніше відомим [3] складає приблизно від 2^{36} до 2^{67} разів в залежності від параметрів кодів, що відновлюються, та ймовірності спотворення у ДСК. Для забезпечення потрібної надійності відновлення кодів запропонований метод потребує більше спотворених слів у порівнянні з методом [3], але характеризується суттєво меншою трудомісткістю.

В окремих (визначених) випадках запропонований метод виявляється практично застосовним, водночас як раніше відомий метод є практично не реалізованим.

Список використаних джерел:

1. Valembios A. Detection and recognition of a binary linear code. Discrete Applied Mathematics. 2001. Vol. 111 (1-2). P. 199–218.
2. Cluzeau M., Finiasz M. Recovering a code's length and synchronization from a noisy intercepted bitstream. IEEE Conference ISIT'09. *Proc. IEEE Press.* 2009. P. 2737–2731.
3. Алексейчук А. Н., Грязнухин А. Ю. Метод восстановления систематических линейных кодов по наборам искаженных кодовых слов. *Прикладная радиоэлектроника*. 2013. Т. 12. № 2. С. 313–318.
4. Балакин Г. В. Введение в теорию случайных систем уравнений. *Труды по дискретной математике.* М. : ТВП. 1997. Т. 1. С. 1–18.
5. 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. URL: http://eprint.iacr.org/2016/311.
6. 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.

APPLICATION OF BKW ALGORITHM FOR RECOVERING SYSTEMATIC LINEAR BLOCK CODES FROM SAMPLES OF NOISY CODEWORDS

The important practical problem in the information security sphere is the development of methods for recovering discrete mappings, which are used in modern systems for transmitting, processing and storing data, from samples of noisy values of these mappings caused by noise impact (random distortion, deliberate interference, internal faults, etc.). In solving this problem additional difficulties arise in the absence of complete information about the algorithms, which define these mappings and used to transform information. A special case of the problem is systematic linear block codes recovering with unknown generating matrix from samples of corrupted codewords observed at the output of a binary symmetric channel. In this paper, the problem-solving method, which based on the BKW algorithm application, which is used for building the correlation attack on streams ciphers, is suggested. The algorithm is applied for solving not one but (simultaneously) many systems of linear equations with noised right-hand sides by single transformation of their co-coefficients matrix. The justification for the correctness is given and

an estimation of the proposed method efficiency is obtained. Its comparison with the previously known method is made. It is shown that the proposed method has greater efficiency in terms of the complexity and volume of necessary memory, although it requires more noised codewords that are necessary for code generating matrix recovering. Depending on recovered codes parameters and the probabilities of distortion in the communication channel, benefits in terms of the complexity of the proposed method in comparison with the previously known is from 2^{36} up 2^{67} once. The practical applicability of the proposed method for cases, where the previously known method is practically not realizable, is confirmed.

Key words: *information security, deducing of information, discrete mappings recovering, linear block code, system of liner equations with noised right-hand sides, BKW algorithm.*

Одержано 21.01.2019

УДК 004.383.3:004.9.347
DOI: 10.32626/2308-5916.2019-19.94-100

Л. М. Николайчук*, канд. юрид. наук,
А. Р. Воронич*, канд. техн. наук,
Т. О. Заведюк,** канд. техн. наук

* Івано-Франківський національний
технічний університет нафти і газу м. Івано-Франківськ,
** Надвірнянський коледж Національного
транспортного університету м. Надвірна

МЕТОДИ НЕЙРОПРОЦЕСОРНОГО ОПРАЦЮВАННЯ СИГНАЛІВ ТА КОМУНІКАЦІЙНИХ ВЗАЄМОДІЙ У СЕРЕДОВИЩІ СУБ'ЄКТІВ ПРАВА

Обгрунтована концепція адекватності моделей нейропроцесорного опрацювання сигналів та комунікаційних взаємодій в інформаційному середовищі суб'єктів права. Показано взаємозв'язок понять ймовірнісної та суб'єктивної ентропії в теорії інформації та юриспруденції. Запропоновані моделі імпульсно-квадратичного перетворення гармонічних сигналів на вході формального нейрона, модель аксона нейрона, рекурентного корелляційного нейрона та інформаційної нейромоделі суб'єкта права.

Ключові слова: *нейропроцесори, компоненти біологічних нейронів, ймовірнісна та суб'єктивна ентропія, моделі компонентів нейрона та суб'єктів права.*

Вступ. Широкомасштабне застосування ІТ-технологій та комп'ютеризованих систем керування є одним з істотних факторів соціально-економічного і технологічного розвитку. Комп'ютеризовані та хмарні

 © Л. М. Николайчук, А. Р. Воронич, Т. О. Заведюк, 2019