УДК 621.391:519.2

МЕТОД ВОССТАНОВЛЕНИЯ СИСТЕМАТИЧЕСКИХ ЛИНЕЙНЫХ КОДОВ ПО НАБОРАМ ИСКАЖЕННЫХ КОДОВЫХ СЛОВ

А.Н. АЛЕКСЕЙЧУК, А.Ю. ГРЯЗНУХИН

Показано, что задача восстановления систематического линейного кода по набору искаженных кодовых слов, наблюдаемых на выходе двоичного симметричного канала связи, сводится к решению ряда систем линейных уравнений с искаженными правыми частями. Получены оценки сложности решения указанных систем уравнений. Представлены результаты вычислительных экспериментов по восстановлению кодов с малой плотностью проверок на четность.

Ключевые слова: восстановление линейных кодов, система уравнений с искаженными правыми частями, коды с малой плотностью проверок на четность.

ВВЕДЕНИЕ

Одной из практически важных задач в области информационной безопасности является разработка методов восстановления дискретных отображений по наблюдениям за их значениями. Как правило, в реальных условиях такие наблюдения производятся под воздействием шумов (случайных искажений, преднамеренных помех, внутренних сбоев и т.п.), что приводит к специфическим по своей постановке задачам. К их числу относится задача восстановления неизвестных систематических линейных кодов по наборам искаженных кодовых слов, наблюдаемых на выходе двоичного симметричного канала связи. Несмотря на то, что эта задача является естественной как с теоретической, так и с прикладной точек зрения, авторам не удалось найти упоминаний о ней в научных публикациях.

В настоящей статье показано, что эта задача сводится к решению ряда систем линейных уравнений (СЛУ) с искаженными правыми частями (с основами теории таких систем уравнений можно ознакомиться по работам [1, 2]). Показано также, что вероятности искажений в правых частях полученных систем зависят от максимальной плотности проверочных соотношений искомого кода, так что сложность его восстановления растет экспоненциально с ростом указанного параметра.

Проведенные вычислительные эксперименты свидетельствуют о том, что предложенный метод может быть эффективно применен на практике для восстановления линейных кодов с малой плотностью проверок на четность [3], длина и размерность которых не превышают нескольких сотен бит. В частности, если вероятности искажений в канале не превосходят 0,1, то для восстановления одного из таких кодов длины 128 и размерности 80 требуется примерно 15 секунд работы стандартной ЭВМ и не более 385 искаженных кодовых слов.

ПОСТАНОВКА ЗАДАЧИ И ОСНОВНЫЕ ТЕРЕТИЧЕСКИЕ РЕЗУЛЬТАТЫ

Пусть C – неизвестный двоичный линейный (n,k) – код с порождающей матрицей $G=(E_k,X)$, где E_k – единичная матрица порядка k, X – матрица размера $k\times(n-k)$, не содержащая нулевых столбцов. Наблюдается последовательность векторов

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

где U_i – независимые случайные равновероятные двоичные векторы длины k (информационные сообщения), $\eta_i=(\eta_{i,1},\dots,\eta_{i,n})$ – векторы искажений, координаты которых являются независимыми в совокупности и не зависящими от $U_1,\dots,U_m$ случайными величинами, распределенными по законам

$$\mathbf{P}\{\eta_{i,s}=0\}=1-\mathbf{P}\{\eta_{i,s}=1\}=1/2\cdot\left(1+\theta_{i,s}\right), \qquad (2)$$

где

$$\theta_{i,s}\ge\theta>0\ ,\ i\in\overline{1,m}\ ,\ s\in\overline{1,n}\ . \qquad (3)$$

Требуется восстановить матрицу X по известным значениям n,k,θ и последовательности (1).

Предлагаемый метод решения поставленной задачи заключается в построении систем линейных уравнений с искаженными правыми частями относительно столбцов матрицы X и решении этих систем с использованием известных алгоритмов. Для изложения метода введем ряд дополнительных обозначений.

Для любого натурального l обозначим V_l множество двоичных векторов длины l. Обозначим $Y_i^{(1)}$ и $Y_i^{(2)}$ подвекторы вектора Y_i, состоящие из его первых k и последних $n-k$ координат соответственно. Аналогичные обозначения введем для случайного вектора η_i, $i\in\overline{1,m}$. Положим $A_i=U_i\oplus\eta_i^{(1)}$, $\xi_i=\eta_i^{(1)}X\oplus\eta_i^{(2)}$, $i\in\overline{1,m}$.

Из формулы $G=(E_k,X)$ вытекает, что равенства (1) равносильны соотношениям

$$(Y_i^{(1)},Y_i^{(2)})=(U_i\oplus\eta_i^{(1)},\ U_iX\oplus\eta_i^{(2)})\ ,\ i\in\overline{1,m}\ ,$$

которые могут быть записаны в виде:

$$A_i=Y_i^{(1)}\ ,\ A_iX\oplus\xi_i=Y_i^{(2)}\ ,\ i\in\overline{1,m}\ . \qquad (4)$$

При этом в силу сделанных выше предположений о распределениях случайных векторов U_i, η_i, $i\in\overline{1,m}$, векторы $A_1,\dots,A_m$ независимы в совокупности и равномерно распределены на

множестве V_k, а векторы $\xi_1, \ldots, \xi_m$ независимы в совокупности и не зависят от $A_1, \ldots, A_m$.

Для того, чтобы придать соотношениям (4) более естественный вид, обозначим A матрицу, составленную из строк $A_1, \ldots, A_m$, x_j – j-й столбец матрицы X; положим

$$b^{(j)} = (Y_{1,j}^{(2)}, \ldots, Y_{m,j}^{(2)})^T, \ \xi^{(j)} = (\xi_{1,j}, \ldots, \xi_{m,j})^T,$$

где $Y_{i,j}^{(2)}$ и $\xi_{i,j}$ – j-е координаты векторов $Y_i^{(2)}$ и ξ_i соответственно, $i \in \overline{1,m}$, $j \in \overline{1,n-k}$. На основании вышеизложенного вектор x_j совпадает с истинным решением $x_j^{(0)}$ СЛУ с искаженными правыми частями

$$Ax = b^{(j)} = Ax_j^{(0)} \oplus \xi^{(j)}, \qquad (5)$$

где матрица A и вектор $b^{(j)}$ определяются непосредственно по набору слов вида (1):

$$A_i = Y_i^{(1)}, \ b^{(j)} = (Y_{1,j}^{(2)}, \ldots, Y_{m,j}^{(2)})^T, \ i \in \overline{1,m}, \ j \in \overline{1,n-k}. \qquad (6)$$

Далее, обозначим $\| x_j \|$ вес (число ненулевых координат) вектора x_j, $\rho_C = \max\limits_{1 \le j \le n-k} \| x_j \|$ и предположим, что код C удовлетворяет следующему условию:

$$\rho_C \le \rho, \qquad (7)$$

где $\rho \in \overline{2,k}$. Из равенства $\xi_i = \eta_i^{(1)} X \oplus \eta_i^{(2)}$ и условия (7) вытекает, что случайная величина $\xi_{i,j}$, $j \in \overline{1,n-k}$, является суммой не более $\rho+1$ независимых случайных величин $\eta_{i,s}$, $i \in \overline{1,m}$, $s \in \overline{1,n}$. Отсюда на основании формул (2), (3) следует, что для любых $i \in \overline{1,m}$, $j \in \overline{1,n-k}$ выполняется соотношение

$$\mathbf{P}\{\xi_{i,j} = 0\} = 1 - \mathbf{P}\{\xi_{i,j} = 1\} \ge 1/2 \cdot \left(1 + \theta^{1+\rho}\right). \qquad (8)$$

Итак, доказано следующее утверждение.

Утверждение 1. Пусть выполняются равенства (1), где случайные векторы U_i, η_i, $i \in \overline{1,m}$, удовлетворяют перечисленным выше условиям. Тогда для любого $j \in \overline{1,n-k}$ j-й столбец матрицы X является решением системы уравнений (5), где матрица A и вектор $b^{(j)}$ определяются по формулам (6). При этом A является случайной равновероятной двоичной матрицей размера $m \times k$, а координаты случайного вектора $\xi^{(j)}$ – независимы в совокупности и не зависят от матрицы A, $j \in \overline{1,n-k}$. Кроме того, при выполнении условия (7) справедлива формула (8).

Следствие. Пусть столбцы x_1, ..., x_{n-k} матрицы X удовлетворяют условию $\| x_j \| = \rho \ge 2$, $j \in \overline{1,n-k}$. Тогда восстановление этой матрицы по набору, состоящему из m слов кода C, искаженных в двоичном симметричном канале с вероятностью ошибки $p \in (0, 1/2)$, сводится к решению $n-k$ СЛУ с искаженными правыми частями (5), где $\xi^{(j)} = (\xi_{1,j}, \ldots, \xi_{m,j})^T$ – случайный вектор с независимыми в совокупности координатами, распределенными по закону

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

$$i \in \overline{1,m}, \ j \in \overline{1,n-k}.$$

Таким образом, для нахождения по наблюдаемой последовательности (1) неизвестной матрицы X следует составить СЛУ с искаженными правыми частями (5) и решить их одним из известных методов. Отметим, что эти методы и алгоритмы можно разделить на универсальные (применимые к любым системам уравнений, независимо от вида их левых и правых частей) и алгоритмы, использующие особенности строения матриц коэффициентов и/или множеств искомых решений рассматриваемых СЛУ. К универсальным относятся метод максимума правдоподобия [1], алгоритмы "декодирования по информационным совокупностям" [4 – 6], трудоемкость которых зависит экспоненциально от числа неизвестных системы, а также ряд субэкспоненциальных алгоритмов [7 – 12], лучшие из которых требуют порядка $2^{O(k/\log k)}$ операций при том же количестве уравнений, где k – число неизвестных в системе. Если параметр ρ_C искомого кода удовлетворяет условию (7), где ρ – небольшая константа, то для решения систем уравнений (5) можно использовать метод максимума правдоподобия (трудоемкость которого в этом случае полиномиально зависит от k и экспоненциально от ρ) или один из недавно разработанных алгоритмов [13, 14], предназначенных для нахождения истинных решений СЛУ с искаженными правыми во множестве векторов заданного (малого) веса.

Остановимся подробнее на модификации метода максимума правдоподобия, состоящей в применении процедуры декодирования в ближайшее кодовое слово [1].

Рассмотрим СЛУ с искаженными правыми частями

$$Ax = b = Ax^{(0)} \oplus \xi, \qquad (9)$$

где A – случайная равновероятная двоичная матрица размера $m \times k$, $x^{(0)}$ – фиксированный неизвестный вектор, принадлежащий множеству $M \subseteq V_k$, а $\xi = (\xi_1, \ldots, \xi_m)^T$ – случайный вектор с независимыми в совокупности координатами, распределенными по закону $\mathbf{P}\{\xi_i = 0\} = 1 - \mathbf{P}\{\xi_i = 1\} = p_i = 1/2 \cdot (1 + \theta_i)$, где $\theta_i \ge \tilde{\theta} > 0$ для любого $i \in \overline{1,m}$. Докажем следующее вспомогательное утверждение.

Лемма. Пусть $\hat{x} \in M$ – вектор, удовлетворяющий условию

$$\nu_m(\hat{x}) = \min_{x \in M} \nu_m(x),$$

где $\nu_m(x) = \| Ax \oplus b \|$ для любого $x \in M$. Тогда при указанных выше предположениях относительно системы уравнений (9) справедливо неравенство

$$\mathbf{P}\{\hat{x} \ne x_0\} \le |M| \exp\{-1/8 \cdot \tilde{\theta}^2 m\}. \qquad (10)$$

Доказательство. Проводится по схеме, аналогичной доказательству теоремы 5.1 в [1]. Заметим, что для любого $C > 0$

$$\{\hat{x} \neq x^{(0)}\} \subseteq \{\nu_m(x^{(0)}) \geq C\} \cup \bigcup_{x \in M: x \neq x^{(0)}} \{\nu_m(x) < C\},$$

откуда следует, что

$$\mathbf{P}\{\hat{x} \neq x^{(0)}\} \leq \mathbf{P}\{\nu_m(x^{(0)}) \geq C\} + \\ + (|M|-1) \max_{x \in M: x \neq x^{(0)}} \mathbf{P}\{\nu_m(x) < C\}. \quad (11)$$

Далее, по условию леммы $\nu_m(x^{(0)})$ является суммой независимых случайных величин $\xi_1, \ldots, \xi_m$. Следовательно, полагая

$$C = 1/4 \cdot m\left(2 - \tilde{\theta}\right), \quad (12)$$

на основании неравенства Гефдинга [15] получим следующие соотношения:

$$\mathbf{P}\{\nu_m(x^{(0)}) \geq C\} = \mathbf{P}\{\sum_{i=1}^{m} \xi_i - \sum_{i=1}^{m} \mathbf{E}\xi_i \geq C - \sum_{i=1}^{m} p_i\} \leq \\ \leq \mathbf{P}\{\sum_{i=1}^{m} \xi_i - \sum_{i=1}^{m} \mathbf{E}\xi_i \geq 1/4 \cdot m\tilde{\theta}\} \leq \exp\{-1/8 \cdot m\tilde{\theta}^2\}. \quad (13)$$

Пусть теперь $x \in M$, $x \neq x^{(0)}$; тогда по условию леммы $\nu_m(x)$ является суммой независимых случайных величин $\eta_1, \ldots, \eta_m$, равномерно распределенных на множестве $\{0,1\}$. Следовательно, на основании формулы (12) и неравенства Гефдинга

$$\mathbf{P}\{\nu_m(x) < C\} = \mathbf{P}\{\sum_{i=1}^{m} \eta_i - \sum_{i=1}^{m} \mathbf{E}\eta_i < C - 1/2 \cdot m\} = \\ = \mathbf{P}\{\sum_{i=1}^{m} \eta_i - \sum_{i=1}^{m} \mathbf{E}\eta_i < -1/4 \cdot m\tilde{\theta}\} \leq \exp\{-1/8 \cdot m\tilde{\theta}^2\}. \quad (14)$$

Подставляя оценки (13), (14) в формулу (11), получим неравенство (10). Лемма доказана.

Предположим теперь, что столбцы $x_1, \ldots, x_{n-k}$ порождающей матрицы кода C принадлежат известному множеству $M \subseteq \{x \in V_k : 1 \leq \| x \| \leq \rho\}$, где $\rho \in \overline{2,k}$, и требуется восстановить их, решая СЛУ с искаженными правыми частями (5).

Алгоритм решения указанных СЛУ путем декодирования в ближайшее кодовое слово [1] состоит в вычислении для каждого $j \in \overline{1,n-k}$ всех значений $\| Ax \oplus b^{(j)} \|$, где $x \in M$, и нахождении вектора $\hat{x}_j \in M$ такого, что $\| A\hat{x}_j \oplus b^{(j)} \| = \min_{x \in M} \| Ax \oplus b^{(j)} \|$.

Следующее утверждение позволяет оценить эффективность этого алгоритма.

Утверждение 2. Пусть выполнено условие утверждения 1, $\delta \in (0,1)$ и

$$m = \left\lceil 8 \cdot \theta^{-2(1+\rho)} \ln\left((n-k)\,\delta^{-1}\,|M|\right)\right\rceil. \quad (15)$$

Тогда описанный алгоритм восстанавливает все столбцы матрицы X с вероятностью не менее $1-\delta$, используя $O\left(m(n-k)(\rho+1)|M|\right)$ двоичных операций. В частности, если $\| x_j \| = \rho$ для любого $j \in \overline{1,n-k}$, то двоичная временная сложность алгоритма равна

$$T = O\left(m(n-k)(\rho+1)\binom{k}{\rho}\right). \quad (16)$$

Доказательство. Положим $\tilde{\theta} = \theta^{1+\rho}$; тогда на основании леммы и соотношений (8)

$$\mathbf{P}\left(\bigcup_{j=1}^{n-k} \{\hat{x}_j \neq x_j\}\right) \leq \sum_{j=1}^{n-k} \mathbf{P}\{\hat{x}_j \neq x_j\} \leq \\ \leq (n-k)\,|M| \exp\{-1/8 \cdot \tilde{\theta}^2 m\} \leq \delta,$$

где последнее неравенство следует непосредственно из формулы (15). Таким образом, вероятность ошибки алгоритма не превосходит δ. Далее, для нахождения вектора $\hat{x}_j$, $j \in \overline{1,n-k}$, необходимо вычислить векторы $Ax \oplus b^{(j)}$ для всех $x \in M$ (что потребует не более $\rho m |M|$ двоичных операций) и найти их веса (что займет еще $O(m|M|)$ операций). Следовательно, суммарная трудоемкость алгоритма составляет $O\left(m(n-k)(\rho+1)|M|\right)$ двоичных операций, что и требовалось доказать.

Ниже, в табл. 1, 2 приведены численные значения (двоичного) логарифма трудоемкости алгоритма и объема материала, достаточного для восстановления с вероятностью не менее $1-\delta$ систематического линейного кода с параметрами n, k, ρ по набору кодовых слов, искаженных в двоичном симметричном канале с вероятностью ошибки $p = 1/2 \cdot (1-\theta)$.

Как видно из таблиц, количество слов, достаточное для восстановления кодов с указанными параметрами, во многих случаях заметно меньше их размерности k. Это объясняется высокой избыточностью описания класса рассматриваемых кодов, каждый из которых задается некоторой $k \times (n-k)$-матрицей X с малым числом ρ единиц в каждом столбце. Согласно табл. 2, для надежного (с вероятностью не менее 0,9) восстановления кода с параметрами $n = 4000$, $k = 2000$, $\rho = 3$ требуется выполнить не более 2^{53} двоичных операций, что находится в пределах возможностей современных супер-ЭВМ. При $n = 500$, $k = 300$, $\rho = 3$ трудоемкость изложенного алгоритма составляет не более 2^{41} операций, что позволяет говорить о возможности восстановления кодов с такими параметрами в реальном времени с помощью стандартных вычислительных средств.

РЕЗУЛЬТАТЫ ВЫЧИСЛИТЕЛЬНЫХ ЭКСПЕРИМЕНТОВ

Для проверки изложенных выше теоретических выводов были проведены вычислительные эксперименты по восстановлению ряда случайно сгенерированных кодов с малой плотностью проверок на четность. Эксперименты проводились с различными источниками сообщений, обладающими естественной (малой) избыточностью, и различными кодами, параметры которых удовлетворяют условиям $n \leq 256$, $k \leq 100$, $\rho \leq 5$.

В табл. 3, 4 показаны типичные результаты экспериментальных исследований, полученные для двух таких кодов C и C' с порождающими матрицами (E_{80}, X) и (E_{80}, X') соответственно, где X' и X'' – случайно сгенерированные 80×48 -матрицы, содержащие, соответственно, $\rho=3$ и $\rho=5$ единиц в каждом столбце.

Для каждой пары значений (p, m'), указанных в табл. 3, 225 раз выполнялась следующая процедура: m' независимых случайных сообщений источника кодировались кодом C; координаты полученных кодовых слов искажались независимо друг от друга с вероятностью p; полученный список из m' искаженных слов подавался на вход алгоритма, описанного в предыдущем пункте.

В результате выполнения алгоритма восстанавливались столбцы матрицы X. В последних трех колонках табл. 3 указаны средние значения (по всем 225 запускам) числа правильно

Таблица 1

Численные значения параметров (15), (16) ($n=500$, $k=300$, $\delta=0{,}1$)

p	ρ					
	3		4		5	
	m	$\log T$	m	$\log T$	m	$\log T$
0,100	1093	40,84	2029	47,95	3646	54,69
0,050	426	39,48	625	46,25	888	52,65
0,030	301	38,98	405	45,63	527	51,90
0,010	216	38,50	267	45,03	320	51,18
0,001	187	38,30	223	44,77	257	50,87

Таблица 2

Численные значения параметров (15), (16) ($n=4000$, $k=2000$, $\delta=0{,}1$)

p	ρ					
	3		4		5	
	m	$\log T$	m	$\log T$	m	$\log T$
0,100	1475	52,81	2767	62,68	5020	72,18
0,050	575	51,45	852	60,98	1222	70,14
0,030	406	50,95	552	60,35	725	69,39
0,010	291	50,47	364	59,75	440	68,67
0,001	252	50,26	304	59,49	354	68,35

Таблица 3

Результаты вычислительных экспериментов ($n=128$, $k=80$, $\rho=3$, $\delta=0{,}1$)

p	Теория		Эксперимент			
	m	$\log T$	m'	Среднее число восстановленных столбцов (%)	Среднее число восстановленных матриц (%)	Среднее время выполнения алгоритма (сек.)
0,100	836	32,67	304	99,81	92,00	13,96
			376	99,98	99,11	14,98
			384	100	100	15,55
			832	100	100	22,40
0,050	326	31,32	104	99,78	89,33	4,55
			120	99,98	99,11	4,72
			128	100	100	5,05
			320	100	100	10,00
0,030	231	30,82	64	99,70	87,11	3,38
			72	99,93	96,44	4,08
			80	100	100	4,12
			224	100	100	7,60
0,010	165	30,34	40	99,87	49,60	3,04
			48	99,99	93,00	3,11
			56	100	100	3,18
			160	100	100	6,00
0,001	143	30,13	24	99,16	68,89	2,28
			32	99,99	99,56	2,69
			40	100	100	3,75
			136	100	100	6,00

Таблица 4

Результаты вычислительных экспериментов ($n=128$, $k=80$, $\rho=5$, $\delta=0{,}1$)

p	Теория		Эксперимент			
	m	$\log T$	m'	Среднее число восстановленных столбцов (%)	Среднее число восстановленных матриц (%)	Среднее время выполнения алгоритма (сек.)
0,050	659	40,56	656	100	100	741,53
0,030	391	39,81	384	100	100	297,86
0,010	237	39,09	232	100	100	197,97
0,001	191	38,78	184	100	100	155,14

восстановленных столбцов, числа правильных восстановлений всей матрицы X и времени выполнения алгоритма. Для сравнения в левой части табл. 3 приведены значения параметров m и $\log T$, рассчитанные по формулам (15) и (16) соответственно. Вычисления проводились на ЭВМ с процессором Intel Pentium G620 (2,6 ГГц) и объемом оперативной памяти 2 Гб RAM (DDR3) на базе Windows XP (использовалась среда разработки Microsoft Visual C++ Studio 2008).

Отметим, что для моделирования источника использовалась заранее сформированная и протестированная случайная последовательность достаточно высокого качества. Количество запусков (равное 225) выбрано так, чтобы отклонение значений, приведенных в предпоследней колонке табл. 3, от теоретической вероятности правильного восстановления матрицы X не превышало 0,1 с надежностью не менее 0,9973 (см., например, [6], с. 87 – 88). В частности, если все 225 запусков завершены успешно, то при указанном в таблице количестве слов m' вероятность правильного восстановления матрицы X больше либо равна 0,9 с надежностью не менее 99,73 %.

Данные в табл. 4 получены аналогично, лишь с тем отличием, что вместо 225 запусков процедуры (для каждой пары входных значений (p, m')) выполнялось только 10. Последнее обстоятельство обусловлено заметным увеличением времени выполнения алгоритма с ростом параметра ρ .

Как видно из таблиц, для восстановления искомых матриц X' и X'' с заданной надежностью во всех случаях требуется меньше данных по сравнению с теоретической верхней границей (15). При этом время восстановления занимает от нескольких секунд до нескольких минут в зависимости от значений параметров p , ρ и m' .

Таким образом, предложенный метод может быть эффективно применен на практике для восстановления кодов с малой плотностью проверок на четность, длина и размерность которых не превышают нескольких сотен бит. Разработка более эффективных, по сравнению с описанным выше, алгоритмов восстановления кодов является задачей дальнейших исследований.

Литература

[1] Балакин Г.В. Введение в теорию случайных систем уравнений // Труды по дискретной математике. – М.: ТВП. – 1997. – Т. 1. – С. 1–18.

[2] Левитская А.А. Системы случайных уравнений над конечными алгебраическими структурами // Кибернетика и системный анализ. – 2005. – Т. 41, № 1. – С. 82–116.

[3] Галлагер Р.Г. Коды с малой плотностью проверок на четность / Р.Г. Галлагер // Сб. Теория кодирования. Пер. с англ. – М.: Мир, 1964. – С. 139–165.

[4] Евсеев С.Г. О сложности декодирования линейных кодов / С.Г. Евсеев // Проблемы передачи информации. – 1983. – Т. 19. – Вып. 1. – С. 1–8.

[5] Coffey J.T. The complexity of information set decoding / J.T. Coffey, R.M. Goodman // IEEE on Inform. Theory. – 1990. – Vol. 36. – P. 1031–1037.

[6] Becker A. Decoding random binary linear codes in $2^{n/20}$: how 1 + 1 = 0 improves information set decoding / A. Becker, A. Joux, A. May, A. Meurer // Cryptology ePrint Arhive, Report 2012/026, http: // eprint.iacr.org/2012/026.

[7] Коваленко І.М. Про алгоритм суб'експоненційної складності декодування сильно спотворених лінійних кодів / І.М. Коваленко // Доп. АН УРСР. Сер. А. – 1988. – № 10. – С. 16–17.

[8] Blum A. Noise-tolerant learning, the parity problem, and the statistical query model / A. Blum, A. Kalai, H. Wasserman // J. ACM. – 2003. – Vol. 50. – № 3. – P. 506–519.

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

[10] Lyubashevsky V. The pariry problem in the presence of noise, decoding random linear codes, and the subset sum problem / V. Lyubashevsky // APPROX and RANDOM'05, Proceedings. – Springer Verlag, 2005. – P. 378–389.

[11] Fossorier M.P.C. A novel algorithm for solving the LPN problem and its application to security evaluation of the HB protocol for RFID authentication / M.P.C. Fossorier, M.J. Mihaljević, H. Imai, Y. Cui, K. Matsuura // Cryptology ePrint Arhive, Report 2006/197, http://eprint.iacr.org/2006/197.

[12] Minder L. The extended k-tree algorithm / L. Minder, A. Sinclair // The 19th Annual ACM-SIAM Symposium on Discrete Algorithms, Proceedings, 2009. – P. 586–595.

[13] Grigorescu E. On noise-tolerant learning of sparse parities and related problems / E. Grigorescu, L. Reysin, S. Vempala // The 22nd Internationale Conf. of Algorithmic learning Theory, Proceedings, 2011. – P. 413–424.

[14] Valliant G. Finding correlation in subquadratic time, with applications to learning parities and juntas with

noise / G. Valliant // Electronic Colloquium on Computational Complexity, Report 2012/006, http://eccc.hpi-eb.dc/report/2012/006.

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

[16] Ширяев А.Н. Вероятность: Учеб. пособие для вузов / А.Н. Ширяев. – М.: Наука, 1989. – 640 с.

Поступила в редколлегию 15.03.2013

Алексейчук Антон Николаевич, доктор технических наук, профессор кафедры Института специальной связи и защиты информации Национального технического университета Украины “КПИ”. Научные интересы: теоретическая криптография.

Грязнухин Александр Юрьевич, аспирант Института специальной связи и защиты информации Национального технического университета Украины “КПИ”. Научные интересы: корреляционный криптоанализ.

УДК 621.391:519.2

Метод відновлення систематичних лінійних кодів за наборами систематичних кодових слів / А.М. Олексійчук, О.Ю. Грязнухін // Прикладна радіоелектроніка: наук.-техн. журнал. – 2013. – Том 12. – № 2. – С. 313–318.

Показано, що задача відновлення систематичного лінійного коду за набором спотворених кодових слів, що спостерігаються на виході двійкового симетричного каналу зв’язку, зводиться до розв’язання ряду систем лінійних рівнянь зі спотвореними правими частинами. Отримано оцінки складності розв’язання зазначених систем рівнянь. Подано результати обчислювальних експериментів щодо відновлення кодів з малою щільністю перевірок на парність.

Ключові слова: відновлення лінійних кодів, система рівнянь зі спотвореними правими частинами, коди з малою щільністю перевірок на парність.

Бібліогр.: 16 найм.

UDC 621.391:519.95

A method of restoring systematic linear codes by samples of noisy code words / A.N. Alekseychuk, A.Yu. Gryaznukhin // Applied Radio Electronics: Sci. Journ. – 2013. Vol. 12. – № 2. – P. 313–318.

It is shown that the problem of restoring a systematic linear code by samples of noisy code words observed at the output of a binary symmetric channel can be reduced to solving of some systems of linear equations with noised right-hand sides. Estimations of comlexity of an algorithm for solving the said systems of equations are obtained. Results of experiments with computer simulation on retrieving codes with low-density parity-checks in the presence of noise are obtained.

Keywords: restoring of linear codes, system of linear equations with noised right-hand sides, low-density parity-check codes.

Ref.: 16 items.