A. M. Zubkov, V. I. Kruglov, Statistical characteristics of weight spectra of random linear codes over GF(p),
Mat. Vopr. Kriptogr., 2014, Volume 5, Issue 1, 27–38

МАТЕМАТИЧЕСКИЕ ВОПРОСЫ КРИПТОГРАФИИ
2014 Т. 5 № 1 С. 27–38

УДК: 519.212.2

Статистические характеристики весовых спектров случайных линейных кодов над GF(p)

А. М. Зубков, В. И. Круглов

Математический институт им. В. А. Стеклова РАН, Москва

Получено 22.IV.2013

Получены формулы для первых двух моментов элементов весового спектра равновероятно выбранного линейного подкода фиксированного линейного кода над конечным полем $\mathbf{F}_p$ в терминах его весового спектра, а также оценки для распределения минимального веса ненулевого кодового слова в выбранном подкоде. Выведены формулы для распределения веса суммы двух независимых случайных векторов над $\mathbf{F}_p$ с заданными весами, вычислены математическое ожидание и дисперсия этого распределения.

Ключевые слова: линейные коды, случайные подкоды, весовой спектр, слово минимального веса

Statistical characteristics of weight spectra of random linear codes over GF(p)

A. M. Zubkov, V. I. Kruglov

Steklov Mathematical Institute of RAS, Moscow

Abstract. For a random uniform subcode of fixed linear code over the finite field $\mathbf{F}_p$ its weight spectrum is considered. Formulas for the first two moments of the weight spectrum elements and estimates for the minimal nonzero weight distribution of subcode elements are derived in terms of weight spectrum of the code. Formulas for the first two moments and the weight distribution of sum of two independent random vectors having fixed weights are given also.

Key words: linear codes, random subcodes, weight spectrum, word of minimal weight

Citation: *Mathematical Aspects of Cryptography*, 2014, vol. 5, no. 1, pp. 27–38 (Russian).

1. Введение

Работа продолжает исследования весовых спектров линейных подпространств дискретных пространств, начатые в [1]. Результаты, полученные в [1] для линейных пространств (далее — кодов) над $\mathbf{F}_2$, переносятся на линейные коды над $\mathbf{F}_p$, p — простое, а вместо линейных случайных равновероятных кодов рассматриваются случайные равновероятные линейные подкоды заданной размерности в фиксированном линейном коде с известным весовым спектром. В работе получены формулы для первых моментов элементов весового спектра, явные выражения для типичных значений минимального веса ненулевых кодовых слов, формулы для математического ожидания и дисперсии веса суммы случайных независимых векторов с заданными весами. Отметим, что свойства весового спектра случайных равновероятных кодов над $\mathbf{F}_q$, $q = p^k$, изучались и ранее (см. [11] и цитированную там литературу).

Пусть p — фиксированное простое число. Будем обозначать через $\mathbf{F}_p^N = \{X = (x_1, \ldots, x_N)\colon x_1, \ldots, x_N \in \mathbf{F}_p\}$ линейное N-мерное пространство над простым полем $\mathbf{F}_p$. Любое k-мерное ($k < N$) подпространство $L \subset \mathbf{F}_p^N$, будем называть k-мерным линейным кодом.

Весом вектора $X = (x_1, \ldots, x_N) \in \mathbf{F}_p^N$ назовем число $w(X) = \sum_{k=1}^N I\{x_k \neq 0\}$ его ненулевых координат.

Через $\left(\mathbf{F}_p^N\right)_s$ и $\left(\mathbf{F}_p^N\right)_{\leq s}$ будем обозначать соответственно множество векторов фиксированного веса s и множество ненулевых векторов веса, не превосходящего s, в $\mathbf{F}_p^N$:

$$\left(\mathbf{F}_p^N\right)_s = \left\{X \in \mathbf{F}_p^N \mid w(X) = s\right\}, \quad \left(\mathbf{F}_p^N\right)_{\leq s} = \left\{X \in \mathbf{F}_p^N \mid 0 < w(X) \leq s\right\};$$

тогда $\mathbf{F}_p^N = \bigsqcup_{s=0}^N \left(\mathbf{F}_p^N\right)_s$.

Пусть $v_s(L) = |L \cap \left(\mathbf{F}_p^N\right)_s|$ и $v_{\leq s}(L) = |L \cap \left(\mathbf{F}_p^N\right)_{\leq s}|$ — соответственно количество векторов веса s и количество ненулевых векторов веса не больше s в линейном коде L; набор $\{v_s(L)\}_{s=0}^N$ называют *весовым спектром* кода L.

Предельные пуассоновские теоремы для случайных величин $v_s(L)$ и аналогичных им доказывались в [2, 3, 5, 6].

2. Моменты элементов весового спектра случайного подкода

Пусть L — линейный k-мерный код в $\mathbf{F}_p^N$ с весовым спектром $\{v_s\}_{s=0}^N$. Очевидно, линейное подпространство L изоморфно $\mathbf{F}_p^k$. Следующие утверждения являются переформулировками хорошо известных результатов (см.,

например, [8, п. 12.2] или [7, гл. 3, § 4], или [4, гл. 15, § 15.2, теорема 9]) и для случая $p = 2$ приводились в [1].

Лемма 1. а) *Количество линейных k-мерных кодов в $\mathbf{F}_p^N$ равно*

$$\frac{(p^N-1)(p^N-p^1)\dots(p^N-p^{k-1})}{(p^k-1)(p^k-p^1)\dots(p^k-p^{k-1})}.$$

б) *Вероятность того, что случайный линейный код, имеющий равномерное распределение на множестве всех k-мерных кодов в $\mathbf{F}_p^N$, содержит фиксированные линейно независимые векторы $X_1, X_2, \dots, X_m \in \mathbf{F}_p^N$, равна*

$$\frac{(p^k-1)(p^k-p^1)\dots(p^k-p^{m-1})}{(p^N-1)(p^N-p^1)\dots(p^N-p^{m-1})}.$$

Согласно лемме 1, в k-мерном линейном коде $L \subset \mathbf{F}_p^N$ существует

$$\frac{(p^k-1)(p^k-p^1)\dots(p^k-p^{k^*-1})}{(p^{k^*}-1)(p^{k^*}-p^1)\dots(p^{k^*}-p^{k^*-1})}$$

k^*-мерных линейных подкодов. Следующая теорема содержит формулы для первых двух моментов элементов

$$v_s(L^*) = |\{X \in L^* \mid w(X) = s\}|, \quad s = 0, \dots, N,$$

весового спектра k^*-мерного случайного кода L^*, выбранного случайно и равновероятно из множества всех k^*-мерных подкодов k-мерного кода L.

Теорема 1. *Если $L \subset \mathbf{F}_p^N$ — линейный k-мерный код в $\mathbf{F}_p^N$ с весовым спектром $\{v_s = v_s(L)\}_{s=0}^N$, а L^* — случайный равновероятный k^*-мерный подкод L, $1 \le k^* < k$, то при $s = 1, \dots, N$*

$$\mathbf{E}v_s(L^*) = v_s(L)\frac{p^{k^*}-1}{p^k-1},$$

$$\mathbf{D}v_s(L^*) = v_s(L)\frac{(p^{k^*}-1)(p^k-p^{k^*})(p-1)}{(p^k-1)^2}\left(1 - \frac{v_s(L)-(p-1)}{p^k-p}\right),$$

и при $s, t \in \{1, \dots, N\}$, $s \ne t$,

$$\operatorname{cov}(v_s(L^*), v_t(L^*)) = -v_s(L)v_t(L)\frac{(p^{k^*}-1)(p^k-p^{k^*})(p-1)}{(p^k-1)^2(p^k-p)}.$$

Доказательство. Из п. б) леммы 1 следует, что $\mathbf{P}\{X \in L^*\} = \frac{p^{k^*}-1}{p^k-1}$ для любого $X \in L$, $X \neq 0$. Поэтому при $s = 1, \ldots, N$

$$\mathbf{E}v_s(L^*) = \sum_{X \in L \cap (\mathbf{F}_p^N)_s} \mathbf{P}\{X \in L^*\} = v_s(L)\frac{p^{k^*}-1}{p^k-1}.$$

Аналогично, для любых двух линейно независимых векторов $X_1, X_2 \in L$

$$\begin{aligned}\operatorname{cov}(I\{X_1 \in L^*\}, I\{X_2 \in L^*\}) = \\ = \mathbf{E}I\{X_1, X_2 \in L^*\} - \mathbf{E}I\{X_1 \in L^*\}\mathbf{E}I\{X_2 \in L^*\} = \\ = \frac{\left(p^{k^*}-1\right)\left(p^{k^*}-p\right)}{\left(p^k-1\right)\left(p^k-p\right)} - \frac{\left(p^{k^*}-1\right)^2}{\left(p^k-1\right)^2} = -\frac{(p^{k^*}-1)(p^k-p^{k^*})(p-1)}{(p^k-1)^2(p^k-p)},\end{aligned}$$

а для любых двух ненулевых линейно зависимых векторов $X_1, X_2 \in L$ ($X_2 = aX_1 \neq 0$ при некотором $a \in \mathbf{F}_p^* = \mathbf{F}_p \backslash \{0\}$)

$$\begin{aligned}\operatorname{cov}(I\{X_1 \in L^*\}, I\{X_2 \in L^*\}) = \\ = \mathbf{E}I\{X_1, X_2 \in L^*\} - \mathbf{E}I\{X_1 \in L^*\}\mathbf{E}I\{X_2 \in L^*\} = \\ = \frac{p^{k^*}-1}{p^k-1} - \frac{\left(p^{k^*}-1\right)^2}{\left(p^k-1\right)^2} = \frac{(p^{k^*}-1)(p^k-p^{k^*})}{(p^k-1)^2}.\end{aligned}$$

Поэтому

$$\begin{aligned}\mathbf{D}v_s(L^*) = \mathbf{D}\left(\sum\nolimits_{X \in L \cap (\mathbf{F}_p^N)_s} I\{X \in L^*\}\right) = \\ = \sum\nolimits_{X_1 \in L \cap (\mathbf{F}_p^N)_s} \sum\nolimits_{a \in \mathbf{F}_p^*} \operatorname{cov}(I\{X_1 \in L^*\}, I\{aX_1 \in L^*\}) + \\ + \sum_{\substack{X_1, X_2 \in L \cap (\mathbf{F}_p^N)_s \\ X_2 \neq aX_1 \, \forall a \in \mathbf{F}_p}} \operatorname{cov}(I\{X_1 \in L^*\}, I\{X_2 \in L^*\}) = \\ = v_s(p-1)\frac{(p^{k^*}-1)(p^k-p^{k^*})}{(p^k-1)^2} - v_s\left(v_s-(p-1)\right)\frac{(p^{k^*}-1)(p^k-p^{k^*})(p-1)}{(p^k-1)^2(p^k-p)} = \\ = v_s(L)\frac{(p^{k^*}-1)(p^k-p^{k^*})(p-1)}{(p^k-1)^2}\left(1 - \frac{v_s(L)-(p-1)}{p^k-p}\right).\end{aligned}$$

Наконец, при $s, t \in \{1, \ldots, N\}$, $s \neq t$, пользуясь линейностью оператора ковариации и тем, что любые векторы $X_1, X_2 \in L$ линейно независимы при

$w(X_1) = s, w(X_2) = t$ и $s \neq t$, находим:

$$\operatorname{cov}(v_s(L^*), v_t(L^*)) = \operatorname{cov}\left(\sum_{X\in L\cap(\mathbf{F}_p^N)_s} I\{X \in L^*\}, \sum_{X\in L\cap(\mathbf{F}_p^N)_t} I\{X \in L^*\}\right) =$$

$$= \sum\nolimits_{X_1\in L\cap(\mathbf{F}_p^N)_s} \sum\nolimits_{X_2\in L\cap(\mathbf{F}_p^N)_t} \operatorname{cov}(I\{X_1 \in L^*\}, I\{X_2 \in L^*\}) =$$

$$= -v_s(L)v_t(L)\frac{(p^{k^*}-1)(p^k-p^{k^*})(p-1)}{(p^k-1)^2(p^k-p)}.$$

Если $L = \mathbf{F}_p^N$, т. е. $1 \le k^* < k = N$, то, согласно теореме 1,

$$\mathbf{E}v_s(L^*) = v_s(\mathbf{F}_p^N)\frac{p^{k^*}-1}{p^N-1}, \quad v_s(\mathbf{F}_p^N) = C_N^s(p-1)^s,$$

$$\mathbf{D}v_s(L^*) = v_s(\mathbf{F}_p^N)\frac{(p^{k^*}-1)(p^N-p^{k^*})(p-1)}{(p^N-1)^2}\left(1 - \frac{v_s(\mathbf{F}_p^N)-(p-1)}{p^N-p}\right).$$

Линейный код L^* содержит $p^{k^*}-1$ ненулевых векторов. Покажем, что дисперсии элементов его весового спектра заметно отличаются от дисперсий элементов весового спектра случайного множества M, выбранного равновероятно из совокупности $(p^{k^*}-1)$-элементных подмножеств множества $\mathbf{F}_p^N$.

Очевидно, в этом случае

$$\mathbf{P}\{X \in M\} = \frac{|M|}{p^N-1}, \quad \mathbf{P}\{X_1, X_2 \in M\} = \frac{C_{p^N-3}^{|M|-2}}{C_{p^N-1}^{|M|}} = \frac{|M|(|M|-1)}{(p^N-1)(p^N-2)}$$

при любых фиксированных $X, X_1, X_2 \in \mathbf{F}_p^N\backslash\{0\}$, $X_1 \neq X_2$. Поэтому при любом $s \in \{1,\dots,N\}$ аналогично доказательству теоремы 1 получаем

$$\mathbf{E}v_s(M) = \sum_{X\in(\mathbf{F}_p^N)_s} \mathbf{P}\{X \in M\} = v_s(\mathbf{F}_p^N)\frac{|M|}{p^N-1},$$

$$\mathbf{D}v_s(M) = \mathbf{D}\sum_{X\in(\mathbf{F}_p^N)_s} I\{X \in M\} = v_s(\mathbf{F}_p^N)\frac{|M|}{p^N-1}\left(1 - \frac{|M|}{p^N-1}\right) +$$

$$+ v_s(\mathbf{F}_p^N)(v_s(\mathbf{F}_p^N)-1)\left(\frac{|M|(|M|-1)}{(p^N-1)(p^N-2)} - \frac{|M|^2}{(p^N-1)^2}\right) =$$

$$= v_s(\mathbf{F}_p^N)\frac{|M|}{p^N-1}\left(1 - \frac{|M|-1}{p^N-2} - v_s(\mathbf{F}_p^N)\frac{p^N-|M|-1}{(p^N-1)(p^N-2)}\right).$$

Заменяя $|M|$ на $p^{k^*}-1$, находим, что отношение дисперсий

$$\frac{\mathbf{D}v_s(M)}{\mathbf{D}v_s(L^*)} = \frac{1}{p-1}\left(1+\frac{p^{k^*}}{p^N-p^{k^*}}\right)\frac{1-\frac{p^{k^*}-2}{p^N-2}-v_s(\mathbf{F}_p^N)\frac{p^N-p^{k^*}}{(p^N-1)(p^N-2)}}{1-\frac{v_s(\mathbf{F}_p^N)-(p-1)}{p^N-p}}$$

заметно меньше 1 (более того, меньше $\frac{1}{p-1}$) при $p \geq 3$.

Теорема 2. *Если выполнены условия теоремы* 1, *то*

$$\mathbf{E}v_{\leq s}(L^*) = \frac{p^{k^*}-1}{p^k-1}\,v_{\leq s}(L),$$

$$\mathbf{D}v_{\leq s}(L^*) = \frac{(p^{k^*}-1)(p^k-p^{k^*})(p-1)}{(p^k-1)^2}\,\frac{p^k-1-v_{\leq s}(L)}{p^k-p}\,v_{\leq s}(L) \leq \qquad (1)$$

$$\leq (p-1)\,\frac{p^k-p^{k^*}}{p^k-1}\,\mathbf{E}v_{\leq s}(L^*)$$

и для минимального веса $\mu(L^*) = \min\{w(X)\colon X \in L^*\backslash\{0\}\}$ *ненулевых векторов в* L^* *справедлива оценка*

$$\frac{1}{1+\frac{p^k-p^{k^*}}{p^k-1}(p-1)(\mathbf{E}v_{\leq s}(L^*))^{-1}} \leq \mathbf{P}\{\mu(L^*)\leq s\} \leq \mathbf{E}v_{\leq s}(L^*). \qquad (2)$$

Доказательство. Используя теорему 1 и равенство $\mathbf{P}\{v_0(L^*)=1\}=1$, находим

$$\mathbf{E}v_{\leq s}(L^*) = \sum_{r=1}^{s}\mathbf{E}v_r(L^*) = \frac{p^{k^*}-1}{p^k-1}\sum_{r=1}^{s}v_r(L) = \frac{p^{k^*}-1}{p^k-1}\,v_{\leq s}(L).$$

Далее,

$$\mathbf{D}v_{\leq s}(L^*) = \mathbf{D}\sum\nolimits_{r=1}^{s}v_r(L^*) = \sum\nolimits_{r_1,r_2=1}^{s}\operatorname{cov}(v_{r_1}(L^*),v_{r_2}(L^*)) =$$

$$= \sum\nolimits_{r=1}^{s}v_r(L)\frac{(p^{k^*}-1)(p^k-p^{k^*})(p-1)}{(p^k-1)^2}\left(1-\frac{v_r(L)-(p-1)}{p^k-p}\right) -$$

$$-\sum_{\substack{r_1,r_2=1\\ r_1\neq r_2}}^{s}v_{r_1}v_{r_2}\frac{(p^{k^*}-1)(p^k-p^{k^*})(p-1)}{(p^k-1)^2(p^k-p)} = \frac{(p^{k^*}-1)(p^k-p^{k^*})(p-1)}{(p^k-1)^2}\times$$

$$\times\left(\left(1+\frac{p-1}{p^k-p}\right)\sum\nolimits_{r=1}^{s}v_r(L) - \frac{1}{p^k-p}\left(\sum\nolimits_{r=1}^{s}v_r(L)\right)^2\right).$$

Чтобы получить указанную в формулировке теоремы формулу для дисперсии, остается вынести за скобки множитель $\sum_{r=1}^{s} v_r(L) = v_{\le s}(L)$.

Верхняя оценка дисперсии для любого $s \in \{1, \dots, N\}$ следует из того, что

$$\frac{(p^{k^*}-1)(p^k-p^{k^*})(p-1)}{(p^k-1)^2}\, v_{\le s}(L) = (p-1)\frac{p^k-p^{k^*}}{p^k-1}\, \mathbf{E} v_{\le s}(L^*),$$

и, если хотя бы одна из величин $v_1(L), \dots, v_s(L)$ отлична от нуля, то $v_{\le s}(L) \ge$ $\ge p-1$ и имеет место неравенство

$$\frac{p^k - 1 - v_{\le s}(L)}{p^k - p} \le 1,$$

если же $v_1(L) = \dots = v_s(L) = 0$, то $\mathbf{D} v_{\le s}(L^*) = \mathbf{E} v_{\le s}(L^*) = 0$.

Доказательство (2), как и в [1], основано на равенстве $\{\mu(L^*) \le s\} =$ $= \{v_{\le s}(L^*) \ge 1\}$ и двойном неравенстве для целочисленной неотрицательной случайной величины ξ

$$\frac{(\mathbf{E}\xi)^2}{\mathbf{D}\xi + (\mathbf{E}\xi)^2} = \frac{(\mathbf{E}\xi)^2}{\mathbf{E}\xi^2} \le \mathbf{P}\{\xi \ge 1\} \le \mathbf{E}\xi,$$

правая часть которого следует из формулы полной вероятности и определения математического ожидания, а левая — из неравенства Иенсена $\mathbf{E}\{\xi^2 \mid \xi \ge$ $\ge 1\} \ge (\mathbf{E}\{\xi \mid \xi \ge 1\})^2$, которое с учетом равенств $\mathbf{E}\xi I\{\xi \ge 1\} = \mathbf{E}\xi$, $\mathbf{E}\xi^2 I\{\xi \ge 1\} = \mathbf{E}\xi^2$ принимает вид $\frac{\mathbf{E}\xi^2}{\mathbf{P}\{\xi \ge 1\}} \ge \left(\frac{\mathbf{E}\xi}{\mathbf{P}\{\xi \ge 1\}}\right)^2$.

Тогда правое неравенство в соотношении (2) теоремы 2 становится очевидным, а левое вытекает из верхней оценки дисперсии в (1):

$$\frac{(\mathbf{E} v_{\le s}(L^*))^2}{\mathbf{D} v_{\le s}(L^*) + (\mathbf{E} v_{\le s}(L^*))^2} \ge \frac{(\mathbf{E} v_{\le s}(L^*))^2}{(p-1)\,\frac{p^k-p^{k^*}}{p^k-1}\,\mathbf{E} v_{\le s}(L^*) + (\mathbf{E} v_{\le s}(L^*))^2} =$$

$$= \frac{1}{1 + \frac{p^k-p^{k^*}}{p^k-1}\,(p-1)(\mathbf{E} v_{\le s}(L^*))^{-1}}.$$

3. Характеристики распределения суммы векторов с заданными весами

Теорема 3. *Если X и Y — независимые случайные векторы, причем вектор X имеет равномерное распределение на множестве $(\mathbf{F}_p^N)_s$ всех векторов*

веса s, а вектор Y имеет равномерное распределение на множестве $(\mathbf{F}_p^N)_t$ всех векторов веса t, то при $|s-t| \le m \le \min\{s+t, N\}$

$$\mathbf{P}\{w(X+Y)=m\} = \sum_{j=\max\{0,s+t-N\}}^{s} \frac{C_s^j C_{N-s}^{t-j}}{C_N^t} C_j^{m-(s+t-2j)} \frac{(p-2)^{m-(s+t-2j)}}{(p-1)^j}$$

и

$$\mathbf{E}w(X+Y) = s + t - \frac{p}{p-1}\frac{st}{N},$$

$$\mathbf{D}w(X+Y) = \frac{st}{N}\left(\frac{p^2}{(p-1)^2}\frac{N}{N-1}\left(1-\frac{t}{N}\right)\left(1-\frac{s}{N}\right) + \frac{p-2}{(p-1)^2}\right).$$

Доказательство. Пусть $X = (x_1, \ldots, x_N) \in (\mathbf{F}_p^N)_s$, $Y = (y_1, \ldots, y_N) \in (\mathbf{F}_p^N)_t$, $0 \le s \le t \le N$. Тогда

$$t - s \le w(X+Y) \le \min\{s+t, N\}$$

и

$$\max\{0, s+t-N\} \le |\{k \in \{1,\ldots,N\} \colon x_k \ne 0 \ne y_k\}| \le \min\{s,t\} = s.$$

При $m \in \{t-s, t-s+1, \ldots, \min\{s+t, N\}\}$ по формуле полной вероятности по значениям $j = |\{k \in \{1,\ldots,N\} \colon x_k \ne 0 \ne y_k\}|$ получаем указанную в условии теоремы формулу:

$$\mathbf{P}\{w(X+Y)=m\} = \sum_{j=\max\{0,s+t-N\}}^{s} \frac{C_s^j C_{N-s}^{t-j}}{C_N^t} C_j^{m-(s+t-2j)} \frac{(p-2)^{m-(s+t-2j)}}{(p-1)^j}.$$

Для вычисления $\mathbf{E}w(X+Y)$ удобно использовать аддитивность математического ожидания и инвариантность распределений X и Y относительно перенумераций координат:

$$\mathbf{E}w(X+Y) = \sum_{k=1}^{N} \mathbf{P}\{x_k + y_k \ne 0\} = N\mathbf{P}\{x_1 + y_1 \ne 0\}.$$

При каждом $k = 1, \ldots, N$ три события $A_k^0 = \{x_k = 0, y_k \ne 0\}$, $A_k^1 = \{x_k \ne 0, y_k = 0\}$, $A_k^2 = \{x_k \ne 0, y_k \ne 0, x_k + y_k \ne 0\}$ попарно не пересекаются и

$$\{x_k + y_k \ne 0\} = A_k^0 \cup A_k^1 \cup A_k^2$$

(события A_k^i и A_m^j при $k \neq m$ не являются ни несовместными, ни независимыми). Используя это разложение и независимость X и Y, находим, что

$$\mathbf{P}\{x_1+y_1 \neq 0\} = \frac{s}{N}\left(1-\frac{t}{N}\right)+\left(1-\frac{s}{N}\right)\frac{t}{N}+\frac{s}{N}\frac{t}{N}\frac{p-2}{p-1} = \frac{s+t}{N}-\frac{p}{p-1}\frac{st}{N^2}.$$

Следовательно, $\mathbf{E}w(X+Y) = s+t-\frac{p}{p-1}\frac{st}{N}$.

Рассматривая случайную величину $w(x+y)$ как сумму индикаторов, найдем ее дисперсию:

$$\mathbf{D}w(x+y) = \mathbf{D}\sum_{k=1}^{N} I\{x_k+y_k \neq 0\} = \\ = \sum_{k,m=1}^{N} \operatorname{cov}(I\{x_k+y_k \neq 0\}, I\{x_m+y_m \neq 0\}).$$

Слагаемые с $k = m$ одинаковы и равны

$$\mathbf{P}\{x_1+y_1 \neq 0\}(1-\mathbf{P}\{x_1+y_1 \neq 0\}),$$

следовательно, их вклад в дисперсию составляет

$$N\left(\frac{s+t}{N}-\frac{p}{p-1}\frac{st}{N^2}\right)\left(1-\frac{s+t}{N}+\frac{p}{p-1}\frac{st}{N^2}\right). \tag{3}$$

Все $N(N-1)$ слагаемых с $k \neq m$ тоже одинаковы и равны:

$$\operatorname{cov}(I\{x_1+y_1 \neq 0\}, I\{x_2+y_2 \neq 0\}) = \sum_{i,j=0}^{2} \operatorname{cov}(I\{A_1^i\}, I\{A_2^j\}).$$

Вычислим отдельно каждое из девяти слагаемых в последней сумме. Если $i = j = 0$ или $i = j = 1$, то

$$\operatorname{cov}(I\{A_1^0\}, I\{A_2^0\}) = \mathbf{P}\{x_1 = 0 \neq y_1, x_2 = 0 \neq y_2\} - \mathbf{P}\{x_1 = 0, y_1 \neq 0\}^2 = \\ = \frac{(N-s)(N-s-1)}{N(N-1)}\frac{t(t-1)}{N(N-1)} - \frac{(N-s)^2}{N^2}\frac{t^2}{N^2} = \\ = \left(1-\frac{s}{N}\right)\frac{t}{N}\left(\frac{1+s+t-N}{(N-1)^2}-\frac{st(2N-1)+Nt}{N^2(N-1)^2}\right), \\ \operatorname{cov}(I\{A_1^1\}, I\{A_2^1\}) = \mathbf{P}\{x_1 \neq 0 = y_1, x_2 \neq 0 = y_2\} - \mathbf{P}\{x_1 \neq 0 = y_1\}^2 = \\ = \left(1-\frac{t}{N}\right)\frac{s}{N}\left(\frac{1+s+t-N}{(N-1)^2}-\frac{st(2N-1)+Ns}{N^2(N-1)^2}\right). \tag{4}$$

Если $i + j = 1$, то

$$\begin{gathered}\operatorname{cov}(I\{A_1^0\}, I\{A_2^1\}) = \operatorname{cov}(I\{A_1^1\}, I\{A_2^0\}) = \\ = \mathbf{P}\{x_1 = 0 \neq y_1, x_2 \neq 0 = y_2\} - \mathbf{P}\{x_1 = 0, y_1 \neq 0\}\mathbf{P}\{x_2 \neq 0, y_2 = 0\} = \\ = \frac{N-s}{N}\frac{s}{N-1}\frac{t}{N}\frac{N-t}{N-1} - \frac{(N-s)ts(N-t)}{N^4} = \\ = \left(1 - \frac{s}{N}\right)\left(1 - \frac{t}{N}\right)\frac{st(2N-1)}{N^2(N-1)^2}.\end{gathered} \tag{5}$$

Если $i = 0$, $j = 2$ или $i = 2$, $j = 0$, то

$$\begin{gathered}\operatorname{cov}(I\{A_1^0\}, I\{A_2^2\}) = \operatorname{cov}(I\{A_1^2\}, I\{A_2^0\}) = \\ = \mathbf{P}\{x_1 = 0 \neq y_1, x_2 \neq 0 \neq y_2, x_2 + y_2 \neq 0\} - \\ -\mathbf{P}\{x_1 = 0 \neq y_1\}\mathbf{P}\{x_2 \neq 0 \neq y_2, x_2 + y_2 \neq 0\} = \\ = \frac{N-s}{N}\frac{s}{N-1}\frac{t}{N}\frac{t-1}{N-1}\frac{p-2}{p-1} - \frac{N-s}{N}\frac{t}{N}\frac{st}{N^2}\frac{p-2}{p-1} = \\ = \left(1 - \frac{s}{N}\right)\frac{t}{N}\frac{p-2}{p-1}\left(\frac{st(2N-1)}{N^2(N-1)^2} - \frac{s}{(N-1)^2}\right).\end{gathered} \tag{6}$$

Аналогично, если $i = 1$, $j = 2$ или $i = 2$, $j = 1$, то

$$\begin{gathered}\operatorname{cov}(I\{A_1^1\}, I\{A_2^2\}) = \operatorname{cov}(I\{A_1^2\}, I\{A_2^1\}) = \\ = \mathbf{P}\{x_1 \neq 0 = y_1, x_2 \neq 0 \neq y_2, x_2 + y_2 \neq 0\} - \\ -\mathbf{P}\{x_1 \neq 0 = y_1\}\mathbf{P}\{x_2 \neq 0 \neq y_2, x_2 + y_2 \neq 0\} = \\ = \left(1 - \frac{s}{N}\right)\frac{t}{N}\frac{p-2}{p-1}\left(\frac{st(2N-1)}{N^2(N-1)^2} - \frac{t}{(N-1)^2}\right).\end{gathered} \tag{7}$$

Наконец, при $i = j = 2$

$$\begin{gathered}\operatorname{cov}(I\{A_1^2\}, I\{A_2^2\}) = \\ = \mathbf{P}\{x_1 \neq 0 \neq y_1, x_1 + y_1 \neq 0, x_2 \neq 0 \neq y_2, x_2 + y_2 \neq 0\} - \\ -\mathbf{P}\{x_1 \neq 0 \neq y_1, x_1 + y_1 \neq 0\}\mathbf{P}\{x_2 \neq 0 \neq y_2, x_2 + y_2 \neq 0\} = \\ = \frac{s(s-1)}{N(N-1)}\frac{t(t-1)}{N(N-1)}\left(\frac{p-2}{p-1}\right)^2 - \frac{s^2}{N^2}\frac{t^2}{N^2}\left(\frac{p-2}{p-1}\right)^2 = \\ = \frac{st}{N^2}\left(\frac{p-2}{p-1}\right)^2\left(\frac{st(2N-1)}{N^2(N-1)^2} - \frac{s+t-1}{(N-1)^2}\right).\end{gathered} \tag{8}$$

Складывая (3) с правыми частями (4)–(8), умноженными на $N(N-1)$, завершаем доказательство теоремы.

Рассмотренная в теореме 3 ситуация возникает, например, при анализе предложенной в [10] системы шифрования Мак-Элис, которая при шифровании k-битового блока открытого сообщения строит n-битовый блок шифртекста, $n > k$.

При построении ключей Алиса выбирает двоичный линейный (n, k)-код C, исправляющий t ошибок, и порождающую его $k \times n$-матрицу G. Для кода C должен существовать эффективный алгоритм декодирования. Далее Алиса выбирает случайную двоичную невырожденную $k \times k$-матрицу S и случайную перестановочную $n \times n$-матрицу P, после чего вычисляет $k \times n$-матрицу $\widehat{G} = SGP$. Открытым ключом Алисы является $(\widehat{G}, t)$, а секретным — разложение (S, G, P) матрицы $\widehat{G}$.

Если Боб хочет послать сообщение m Алисе, то он записывает сообщение m в виде двоичной строки длины k и вычисляет вектор $c^* = m\widehat{G}$. Затем Боб выбирает случайный n-битовый вектор z, содержащий ровно t единиц и называемый вектором ошибки, и строит шифртекст по формуле $c = c^* + z$.

При расшифровании сообщения c, Алиса с помощью обратной к P матрицы P^{-1} находит $\widehat{c} = cP^{-1}$, с помощью алгоритма декодирования кода C строит по $\widehat{c}$ вектор $\widehat{m}$ и находит сообщение m по формуле $m = \widehat{m}S^{-1}$.

Как показано в [9], если одно и то же сообщение m, представляющее собой двоичный вектор размерности k, было зашифровано и отправлено дважды с использованием шифрующей $k \times n$-матрицы SGP и различных векторов ошибок e_1 и e_2 размерности n, то лицо, перехватившее шифртексты $c_1 = mSGP + e_1$ и $c_2 = mSGP + e_2$, может эффективно восстановить сообщение m (но не узнать ключ). Аналогичная возможность возникает, если получены зашифрованные блоки двух сообщений m_1 и m_2, для которых известна их поразрядная сумма $m_1 + m_2$.

При реализации подобной атаки прежде всего нужно определить, являются ли два наблюдаемых вектора c_1 и c_2 независимыми случайными векторами, имеющими равномерное распределение на множестве всех двоичных векторов размерности N (гипотеза H_0), или они были получены путем зашифровки некоторого сообщения m с использованием независимых векторов ошибок e_1 и e_2, имеющих одинаковый вес t и размерность N (гипотеза H_1).

Если справедлива гипотеза H_0, то сумма векторов $c_1 + c_2$ имеет равновероятное распределение на множестве двоичных векторов размерности N и

$$\mathbf{E}w(c_1 + c_2) = N/2, \quad \mathbf{D}w(c_1 + c_2) = N/4, \quad \mathbf{P}\{w(c_1 + c_2) = s\} = C_N^s/2^N.$$

Если же верна гипотеза H_1, то $c_1 + c_2 = mSGP + e_1 + mSGP + e_2 = e_1 + e_2$ и векторы e_1, e_2 независимы и равномерно распределены на множестве

векторов веса t и размерности N. В этом случае аналогичные характеристики распределения величины $w(c_1 + c_2)$ можно найти с помощью теоремы 3, утверждение которой при $p = 2$ совпадает с утверждением теоремы 3 работы [1].

Список литературы

1. *Зубков А. М., Круглов В. И.* Моментные характеристики весов векторов в случайных двоичных линейных кодах // Математические вопросы криптографии. — 2012. — Т. 3. Вып. 4. — С. 55–70.
2. *Копытцев В. А.* О числе решений системы случайных линейных уравнений в множестве векторов специального вида // Дискретная математика. — 2006. — Т. 18. Вып. 1. — С. 40–62.
3. *Копытцев В. А., Михайлов В. Г.* Теоремы пуассоновского типа для числа специальных решений случайного линейного включения // Дискретная математика. — 2010. — Т. 22. Вып. 2. — С. 3–21.
4. *Мак-Вильямс Ф. Дж., Слоэн Н. Дж. А.* Теория кодов, исправляющих ошибки. — М.: Связь, 1979.
5. *Михайлов В. Г.* Предельные теоремы для числа точек случайного линейного подпространства, попавших в заданное множество // Дискретная математика. — 2003. — Т. 15. Вып. 2. — С. 128–137.
6. *Михайлов В. Г.* Предельные теоремы для числа решений системы случайных линейных уравнений, попавших в заданное множество // Дискретная математика. — 2007. — Т. 19. Вып. 1. — С. 17–26.
7. *Сачков В. Н.* Введение в комбинаторные методы дискретной математики. — М. : Наука, 1982.
8. *Холл М.* Комбинаторика. — М.: Мир, 1970.
9. *Berson T.* Failure of the McEliece public-key cryptosystem under message-resend and related-message attack // CRYPTO 1997. Lect. Notes Comput. Sci. — 1997. — V. 1294. — P. 213–220.
10. *McEliece R. J.* A public-key cryptosystem based on algebraic coding theory // Jet Propulsion Lab. DSN Progress Report 42-44. — 1978. — URL: http://ipnpr.jpl.nasa.gov/progress_report2/42-44/44N.PDF
11. *Özen İ., Tekin E.* Moments of the support weight distribution of linear codes // Des., Codes & Cryptogr., DOI 10.1077/s10623-011-9597-7