ДИСКРЕТНЫЙ АНАЛИЗ И ИССЛЕДОВАНИЕ ОПЕРАЦИЙ
Июль—сентябрь 2007. Серия 1. Том 14, № 3. 90–97

УДК 517.919

О ВЕРХНЕЙ ОЦЕНКЕ ЧИСЛА РАВНОМЕРНО УПАКОВАННЫХ ДВОИЧНЫХ КОДОВ*)

Н. Н. Токарева

Рассматриваются равномерно упакованные (в широком смысле) двоичные коды длины n с кодовым расстоянием d и радиусом покрытия ρ. Показано, что любой такой код однозначно определяется множеством своих кодовых слов весов $\lceil n/2 \rceil - \rho, \ldots, \lfloor n/2 \rfloor + \rho$, и в случае нечётного d число различных таких кодов не превышает числа $2^{2^{n - \frac{d}{2}\log_2 n + o(\log_2 n)}}$.

Введение

Рассматривается метрическое пространство E^n на множестве двоичных векторов длины n с метрикой Хемминга $d(\cdot,\cdot)$ (расстояние между двумя векторами равно числу координат, в которых векторы различаются). *Вес Хемминга* $wt(\cdot)$ вектора из E^n определяется как число его ненулевых координат (т. е. как расстояние до нулевого вектора $\mathbf{0}$). Непустое подмножество C в пространстве E^n с минимальным расстоянием d между его различными элементами называется *двоичным* (n,d)*-кодом*, где n — *длина*, а d — *кодовое расстояние* кода. *Радиусом покрытия* ρ двоичного кода C длины n называется максимальное расстояние, на которое может быть удалён от кода C двоичный вектор длины n, т. е. $\rho = \max\limits_{x \in E^n} d(x, C)$.

Согласно работе Л. А. Бассалыго, Г. В. Зайцева и В. А. Зиновьева [2] двоичный (n,d)-код C с радиусом покрытия ρ называется *равномерно упакованным в широком смысле*, если существуют действительные числа $\alpha_0, \alpha_1, \ldots, \alpha_\rho$ такие, что для любого двоичного вектора x длины n выполняется равенство $\sum\limits_{i=0}^{\rho} \alpha_i f_i(x) = 1$, где $f_i(x)$ — число кодовых слов кода C, находящихся на расстоянии i от вектора x, $i = 0, 1, \ldots, \rho$. Пусть

*)Исследование выполнено при финансовой поддержке интеграционного проекта СО РАН № 35 «Древовидный каталог математических Интернет-ресурсов» и Российского фонда фундаментальных исследований (проект 07–01–00248).

$d = 2t + 1$. Известно другое определение равномерно упакованных двоичных кодов j-го порядка ($j = 1, \ldots, t$), введённое Дж. М. Гёталсом и Х. ван Тилборгом [7], которое при $j = \rho - t$ является частным случаем определения из [2]. При $j = \rho - t = 1$ определения из [2] и [7] совпадают; при этом соответствующие коды называются *строго равномерно упакованными* или *равномерно упакованными в узком смысле*. Такие коды впервые были рассмотрены Н. В. Семаковым, В. А. Зиновьевым и Г. В. Зайцевым [5]. Далее под термином «равномерно упакованный» будем понимать «равномерно упакованный в широком смысле».

С. В. Августинович [1] показал, что каждый двоичный совершенный код длины n с кодовым расстоянием $d = 3$ однозначно определяется множеством своих кодовых слов веса $(n-1)/2$. Используя это свойство, в [1] было показано, что число различных совершенных двоичных кодов не превосходит $2^{2^{n-\frac{3}{2}\log n + o(\log n)}}$ (здесь и далее log обозначает логарифм по основанию 2).

Рассмотрим произвольный класс $\mathbb{L}(n, d, \rho; \alpha_0, \ldots, \alpha_\rho)$ двоичных равномерно упакованных (в широком смысле) (n, d)-кодов с радиусом покрытия ρ и параметрами равномерной упаковки $\alpha_0, \ldots, \alpha_\rho$. Считаем, что d и ρ — константы. Число различных кодов в классе $\mathbb{L}(n, d, \rho; \alpha_0, \ldots, \alpha_\rho)$ обозначим через $L_{n,d}$. Используя границу сферической упаковки для мощности (n, d)-кода, несложно получить следующую тривиальную оценку: $L_{n,d} \leqslant 2^{2^{n-\frac{d-1}{2}\log n + o(\log n)}}$. Обобщая метод работы [1], покажем, что любой код из класса $\mathbb{L}(n, d, \rho; \alpha_0, \ldots, \alpha_\rho)$ однозначно определяется множеством своих кодовых слов весов $\lceil n/2 \rceil - \rho, \ldots, \lfloor n/2 \rfloor + \rho$, и в случае нечётного d имеет место оценка: $L_{n,d} < 2^{2^{n-\frac{d}{2}\log n + \log\log n + \delta}}$, где константа $\delta = d \log d + \log(\rho + 1)$.

1. Вспомогательные утверждения

Пусть x, y — любые двоичные векторы длины n, и пусть $d(x, y) = k$. Известно (см., например, [4, гл. 21]), что число векторов $z \in E^n$ таких, что $d(x, z) = i$ и $d(y, z) = j$, не зависит от выбора векторов x и y, а зависит лишь от чисел i, j, k, n. Обозначим это число через p_{ijk} (подразумевая также зависимость этого параметра от n). Ясно, что

$$p_{ijk} = \binom{k}{(i-j+k)/2}\binom{n-k}{(i+j-k)/2},$$

если $i + j - k$ чётно; $p_{ijk} = 0$, если $i + j - k$ нечётно. Будем считать, что параметр p_{ijk} определён при любых значениях i, j и k, $0 \leqslant i, j, k \leqslant n$, и равен нулю, если соответствующее множество векторов z пусто.

Пусть C — произвольный код из класса равномерно упакованных кодов $\mathbb{L}(n, d, \rho; \alpha_0, \ldots, \alpha_\rho)$. Обозначим через C_i и E_i множества векторов веса i кода C и пространства E^n соответственно, где $i = 0, 1, \ldots, n$. Пусть μ_C^i — мощность множества C_i. Набор $\mu(C) = \{\mu_C^0, \mu_C^1, \ldots, \mu_C^n\}$ называется *весовым спектром* кода C, а числа μ_C^i, $i = 0, 1, \ldots, n$, — *спектральными значениями* кода. В работе [2] приведена формула для вычисления весового спектра (более точно: весовой функции) произвольного равномерно упакованного кода, содержащая ρ неизвестных констант. Для определения этих констант требуется знать любые ρ спектральных значений кода, при которых возможно решение соответствующей системы линейных уравнений (см. подробнее [2]).

Убедимся в справедливости следующего утверждения.

Лемма 1. *Весовой спектр произвольного кода C из класса равномерно упакованных кодов $\mathbb{L}(n, d, \rho; \alpha_0, \ldots, \alpha_\rho)$ однозначно определяется значениями $\mu_C^0, \ldots \mu_C^{\rho-1}$.*

Доказательство. Покажем, как с помощью известных значений $\mu_C^0, \ldots \mu_C^{j+\rho-1}$ при любом $j = 0, 1, \ldots, n-\rho$ восстановить значение $\mu_C^{j+\rho}$. При каждом $i = 0, 1, \ldots, \rho$ имеет место следующее равенство

$$\sum_{x \in E_j} f_i(x) = \sum_{k=\max\{0, j-i\}}^{j+i} p_{ijk}\, \mu_C^k. \tag{1}$$

Действительно, каждый кодовый вектор веса k находится на расстоянии i в точности от p_{ijk} двоичных векторов веса j. Заметим, что в каждом соотношении (1) при $i = 0, 1, \ldots, \rho - 1$ участвуют лишь известные спектральные значения $\mu_C^{\max\{0, j-\rho+1\}}, \ldots, \mu_C^{j+\rho-1}$, а при $i = \rho$ единственным неизвестным спектральным значением является $\mu_C^{j+\rho}$, причём оно входит в это равенство с ненулевым коэффициентом. В силу равномерной упакованности кода C справедливо равенство

$$\sum_{x \in E_j} \sum_{i=0}^{\rho} \alpha_i f_i(x) = \binom{n}{j}.$$

Меняя порядок суммирования в этом равенстве и пользуясь (1), получаем

$$\sum_{i=0}^{\rho} \alpha_i \sum_{k=\max\{0, j-i\}}^{j+i} p_{ijk}\, \mu_C^k = \binom{n}{j}.$$

Отсюда однозначно определяется значение $\mu_C^{j+\rho}$. Таким образом последовательно восстанавливаются значения $\mu_C^\rho, \ldots, \mu_C^n$. Лемма 1 доказана.

Следующая лемма является обобщением одного свойства совершенных двоичных кодов, приведённого в работе [1].

Лемма 2. *Множество* $X = C_{\lceil n/2\rceil-\rho} \cup \cdots \cup C_{\lfloor n/2\rfloor+\rho}$ *однозначно определяет код* C *из класса равномерно упакованных кодов* $\mathbb{L}(n, d, \rho; \alpha_0, \ldots, \alpha_\rho)$.

Доказательство. Для кода C обозначим через A и B следующие множества $A = C_0 \cup \cdots \cup C_{\lceil n/2\rceil-\rho-1}$ и $B = C_{\lfloor n/2\rfloor+\rho+1} \cup \cdots \cup C_n$. Имеем $C = A \cup X \cup B$. Несложно заметить, что расстояние между множествами A и B не меньше $2\rho+1$ и, следовательно, не меньше d. Предположим, что существует другой код $C' = A' \cup X \cup B'$ из класса $\mathbb{L}(n, d, \rho; \alpha_0, \ldots, \alpha_\rho)$, и пусть $B \neq B'$. Тогда код C'', полученный из C заменой множества B на B', также имеет кодовое расстояние d. Покажем, что C'' принадлежит классу $\mathbb{L}(n, d, \rho; \alpha_0, \ldots, \alpha_\rho)$, т. е. является равномерно упакованным кодом с параметрами $\alpha_0, \ldots, \alpha_\rho$. Для произвольного вектора $x \in E^n$ рассмотрим сумму

$$\sum_{i=0}^{\rho} \alpha_i f_i(x), \tag{2}$$

где $f_i(x)$ — число кодовых слов кода C'', находящихся на расстоянии i от вектора x. Обозначим через $T_\rho^D(x)$ множество всех кодовых слов произвольного кода D длины n, содержащихся в шаре радиуса ρ с центром в вершине x, т. е. $T_\rho^D(x) = \{\ y \in D \mid d(x, y) \leqslant \rho\ \}$. По построению кода C'' имеем

$$T_\rho^{C''}(x) = \begin{cases} T_\rho^C(x) & \text{при } wt(x) \leqslant \lfloor n/2\rfloor, \\ \\ T_\rho^{C'}(x) & \text{при } wt(x) \geqslant \lceil n/2\rceil. \end{cases}$$

Так как коды C и C' являются равномерно упакованными с параметрами $\alpha_0, \ldots, \alpha_\rho$, то для любого вектора $x \in E^n$ сумма (2) равна 1. Таким образом, код C'' принадлежит классу равномерно упакованных кодов $\mathbb{L}(n, d, \rho; \alpha_0, \ldots, \alpha_\rho)$.

Поскольку $B \neq B'$, без ограничения общности можно считать, что найдётся вектор $y \in E^n$ такой, что $y \in B$ и $y \notin B'$. Пусть $z = y \oplus \mathbf{1}$, где $\mathbf{1}$ — вектор со всеми координатами, равными 1, и $\oplus$ обозначает покоординатное сложение векторов по модулю 2. Тогда, как нетрудно заметить, выполняется неравенство $wt(z) \leqslant \lceil n/2\rceil - \rho - 1$. Поэтому $T_\rho^C(z) = T_\rho^{C''}(z)$. Отсюда следует, что для равномерно упакованных кодов $z \oplus C$ и $z \oplus C''$ (сдвигов кодов C и C'' соответственно на вектор z) первые $\rho + 1$ спектральных значений одинаковы, т. е.

$$\mu^0_{z\oplus C} = \mu^0_{z\oplus C''}, \ldots, \mu^\rho_{z\oplus C} = \mu^\rho_{z\oplus C''}.$$

Тогда согласно лемме 1 коды $z \oplus C$ и $z \oplus C''$ имеют одинаковые весовые спектры. Но поскольку $\mathbf{1} \in z \oplus C$ и $\mathbf{1} \notin z \oplus C''$, имеем $\mu^n_{z\oplus C} \neq \mu^n_{z\oplus C''}$. Полученное противоречие доказывает лемму 2.

Далее докажем одну простую оценку для числа кодовых слов веса i произвольного двоичного кода C. Этой оценки достаточно для доказательства основного результата работы, хотя следует отметить, что известны существенно более сильные оценки для числа $|C_i|$ (см., например, [4, гл. 17]).

Лемма 3. *Для любого двоичного кода C длины n с кодовым расстоянием $d = 2t + 1$ при любом $i = t, \dots, n - t$ справедливо неравенство* $|C_i| \leqslant \frac{2^t t!}{n^t}\binom{n}{i}$.

Доказательство. Пусть $i \leqslant \lfloor n/2 \rfloor$. Для каждого вектора x веса i определим множество V_x, состоящее из векторов веса $i - t$, все ненулевые координаты которых лежат среди ненулевых координат вектора x. Заметим, что $|V_x| = \binom{i}{t}$. Поскольку для любых кодовых векторов x и y из C_i множества V_x и V_y не пересекаются (иначе $d(x, y) < d$), имеем

$$|C_i| \leqslant \frac{|E_{i-t}|}{|V_x|} = \frac{\binom{n}{i-t}}{\binom{i}{t}},$$

откуда следует искомая оценка. Рассуждения легко переносятся на случай $i \geqslant \lceil n/2 \rceil$. Лемма 3 доказана.

2. Верхняя оценка

Основным результатом работы является

Теорема 1. *Для числа $L_{n,d}$ различных кодов из класса равномерно упакованных кодов $\mathbb{L}(n, d, \rho; \alpha_0, \dots, \alpha_\rho)$ при $n \geqslant 3$ и нечётном $d \geqslant 3$ справедлива оценка $L_{n,d} < 2^{2^{n - \frac{d}{2}\log n + \log\log n + \delta}}$, где константа $\delta = d \log d + \log(\rho + 1)$.*

Доказательство. Из леммы 2 следует, что

$$L_{n,d} \leqslant \begin{pmatrix} |E_{\lceil n/2 \rceil - \rho}| + \dots + |E_{\lfloor n/2 \rfloor + \rho}| \\ |C_{\lceil n/2 \rceil - \rho}| + \dots + |C_{\lfloor n/2 \rfloor + \rho}| \end{pmatrix}. \tag{3}$$

Имеем $|E_{\lceil n/2 \rceil - \rho}| + \dots + |E_{\lfloor n/2 \rfloor + \rho}| \leqslant (2\rho + 1)\begin{pmatrix} n \\ \lfloor n/2 \rfloor \end{pmatrix}$. По лемме 3 для произвольного двоичного кода C длины n с кодовым расстоянием d выполняется неравенство $|C_{\lceil n/2 \rceil - \rho}| + \dots + |C_{\lfloor n/2 \rfloor + \rho}| \leqslant \frac{\lambda}{n^t}\begin{pmatrix} n \\ \lfloor n/2 \rfloor \end{pmatrix}$, где

$\lambda = (2\rho+1)\cdot 2^t \cdot t!$ и $t = (d-1)/2$. Применяя формулу Стирлинга

$$n^n e^{-n}\sqrt{2\pi n} \leqslant n! \leqslant n^n e^{1-n}\sqrt{2\pi n},$$

получаем $\left(\begin{array}{c} n \\ \lfloor n/2 \rfloor \end{array} \right) \leqslant 2^{n-\frac{1}{2}\log n+2}$. Тогда в силу (3) имеем

$$L_{n,d} < \left(\begin{array}{c} 2^{n-\frac{1}{2}\log n+(2+\log(2\rho+1))} \\ 2^{n-\frac{d}{2}\log n+(2+\log\lambda)} \end{array} \right).$$

Отсюда и из неравенства $\left(\begin{array}{c} a \\ b \end{array} \right) < \left(\frac{3a}{b}\right)^b$ для любых $a > b > 1$ вытекает

$$L_{n,d} < 2^{2^{n-\frac{d}{2}\log n+\log\log n+(\log\lambda+\log d+1)}}.$$

Тогда, используя неравенство $c! \geqslant \left(\frac{c+1}{2}\right)^c$ для любого $c \geqslant 1$, несложно получить $\log\lambda + \log d + 1 \leqslant d\log d + \log(\rho+1)$, и следовательно,

$$L_{n,d} < 2^{2^{n-\frac{d}{2}\log n+\log\log n+\delta}}.$$

Теорема 1 доказана.

Приведём примеры классов двоичных кодов, к которым применима теорема 1.

1) Двоичные *совершенные коды* длины $n = 2^m - 1$ ($m \geqslant 2$), мощности $2^{n-\log(n+1)}$ с кодовым расстоянием $d = 3$ и параметрами равномерной упаковки $\alpha_0 = \alpha_1 = 1$ (см. [5]). Этот частный случай теоремы 1 был доказан в [1].

Другие примеры равномерно упакованных кодов с $d = 3$ можно найти в [7] (см. также [2, 8]).

2) Двоичные *коды Препараты* длины $n = 2^m - 1$ ($m \geqslant 4$ чётно), мощности $2^{n-2\log(n+1)+1}$ с кодовым расстоянием $d = 5$ и параметрами упаковки $\alpha_0 = \alpha_1 = 1, \alpha_2 = \alpha_3 = 3/n$ [5, 2].

Следствие 1. *Число различных двоичных кодов Препараты длины n с кодовым расстоянием* 5 *не превосходит величины* $2^{2^{n-\frac{5}{2}\log n+o(\log n)}}$.

Замечание. Отметим, что для числа кодов одного специального подкласса кодов Препараты имеет место более точная оценка. А именно, согласно [6, следствие 2], число неэквивалентных четверичных линейных кодов Препараты длины n с кодовым расстоянием 6 не превосходит величины $2^{n\log n}$.

3) Двоичные примитивные *коды типа* БЧХ длины $n = 2^m - 1$ ($m \geqslant 5$ нечётно), мощности $2^{n-2\log(n+1)}$ с кодовым расстоянием $d = 5$, радиусом покрытия $\rho = 3$ и параметрами $\alpha_0 = \alpha_1 = 1, \alpha_2 = \alpha_3 = \frac{6}{n-1}$ (см. [2]).

4) Двоичные *коды Гётамса* (или *коды типа Гётамса*) длины $n = 2^m - 1$ ($m \geqslant 4$ чётно), мощности $2^{n-3\log(n+1)+2}$ с кодовым расстоянием $d = 7$, радиусом покрытия $\rho = 5$ и параметрами упаковки $\alpha_0 = \alpha_1 = 1, \alpha_2 = \alpha_3 = \frac{15}{2n}, \alpha_4 = \alpha_5 = \frac{30}{n(n-3)}$ (см. [7, 3]).

Следствие 2. *Число различных двоичных кодов Гётамса длины n с кодовым расстоянием 7 не превосходит величины* $2^{2^{n-\frac{7}{2}\log n + o(\log n)}}$.

5) Двоичные примитивные *коды типа* БЧХ длины $n = 2^m - 1$ ($m \geqslant 5$ нечётно) мощности $2^{n-3\log(n+1)}$ с кодовым расстоянием $d = 7$, радиусом покрытия $\rho = 5$ и параметрами упаковки $\alpha_0 = \alpha_1 = 1, -\alpha_2 = -\alpha_3 = \alpha_4 = \alpha_5 = \frac{120}{(n-1)(n-7)}$ (см. [7]).

Автор благодарен Д. С. Кротову за ценные замечания, позволившие существенно расширить множество кодов, для которых справедлива теорема 1.

ЛИТЕРАТУРА

1. **Августинович С. В.** Об одном свойстве совершенных двоичных кодов // Дискрет. анализ и исслед. операций. 1995. Т. 2, № 1. С. 4–6.
2. **Бассалыго Л. А., Зиновьев В. А., Зайцев Г. В.** О равномерно упакованных кодах // Проблемы передачи информации. 1974. Т. 10, вып. 1. С. 9–14.
3. **Зиновьев В. А., Хеллесет Т.** О весовых спектрах сдвигов кодов типа Геталса // Проблемы передачи информации. 2004. Т. 40, вып. 2. С. 19–36.
4. **Мак-Вильямс Ф. Дж., Слоэн Н. Дж. А.** Теория кодов, исправляющих ошибки. М: Связь, 1979.
5. **Семаков Н. В., Зиновьев В. А., Зайцев Г. В.** Равномерно упакованные коды // Проблемы передачи информации. 1971. Т. 7, вып. 1. С. 38–50.
6. **Токарева Н. Н.** Представление $\mathbb{Z}_4$-линейных кодов Препараты с помощью векторных полей // Проблемы передачи информации. 2005. Т. 41, вып. 2. С. 50–62.
7. **Goethals J. M., van Tilborg H. C. A.** Uniformly packed codes // Philips Res. Repts. 1975. V. 30. P. 9–36.

8. Rifa J., Zinoviev V. A. On completely regular codes from perfect codes // Proc. Tenth Int. Workshop «Algebraic and Combinatorial Coding Theory» (Zvenigorod, September 3–9, 2006). P. 225–229.

Адрес автора:

Институт математики
им. С. Л. Соболева СО РАН,
пр. Академика Коптюга, 4,
Новосибирский гос. университет,
ул. Пирогова, 2,
630090 Новосибирск,
Россия.
E-mail: tokareva@math.nsc.ru

Статья поступила
14 марта 2006 г.