

УДК 519.72

О НАИБОЛЬШЕМ РАЗМЕРЕ АНТИКОДОВ

А. Я. ван Зантен

Приводится краткое доказательство известного выражения для наибольшего размера двоичного антикода длины n с максимальным расстоянием m . Это выражение впервые было изучено Д. Клейтманом в контексте экстремальных задач о подмножествах конечного множества, а несколько ранее Д. Катоней.

Введение

В алгебраической теории кодирования обычно изучаются двоичные (n, M, d) -коды. Двоичный (n, M, d) -код состоит из $|M|$ двоичных слов длины n с минимальным расстоянием Хемминга, равным d . При фиксированных двух параметрах кода ставится вопрос об экстремальном значении третьего параметра. Аналогичные оптимизационные задачи рассматриваются и для q -значных кодов в алфавите $\{0, 1, \dots, q-1\}$, когда расстояние Хемминга между двумя словами определяется как число позиций, в которых эти слова различаются. Если вместо минимального расстояния рассматривается максимальное расстояние между двумя словами из M , то говорят об *антикоде* [3, 6, 7]. В частности, можно ставить вопрос о наибольшем размере $N(n, m)$ антикода с наибольшим расстоянием m между словами длины n . В [7] было высказано предположение, что

$$N(n, m) = \sum_{i=0}^k \binom{n}{i} \quad \text{при чётном } m,$$
$$N(n, m) = \sum_{i=0}^k \binom{n}{i} + \binom{n-1}{k} \quad \text{при нечётном } m,$$

где $k = \left\lfloor \frac{m}{2} \right\rfloor$ и $1 \leq m < n$. Правые части этих соотношений равны размеру шара радиуса $m/2$ с дополнительным слагаемым для нечётных значений m . Хотя этот результат представляется естественным, но его доказательство было получено не сразу. Фактически это сделали П. Эрдёш

(для чётного m), затем Д. Клейтман [5] и Д. Катона [4] с использованием различных подходов в терминах экстремальной теории множеств. Здесь представлено относительно короткое индуктивное доказательство в терминах двоичных слов, основанное на известной теореме о сложении биномиальных коэффициентов. В разделе 2 кратко показано, как этот подход может быть использован для получения аналогичного результата в случае q -значных кодов (см. также [1, 2]).

1. Доказательство теоремы

Пусть C — множество мощности $N := |C|$ двоичных слов длины n . Назовём множество C *двоичным кодом* длины n и размера N . Как обычно, *расстояние Хемминга* $d(\mathbf{v}, \mathbf{w})$ между словами \mathbf{v} и \mathbf{w} из C определяется как число позиций, в которых \mathbf{v} и \mathbf{w} различаются. Наибольшее расстояние в C называется *диаметром* $d(C)$ кода:

$$d(C) := \max\{d(\mathbf{v}, \mathbf{w}) \mid \mathbf{v}, \mathbf{w} \in C\}. \quad (1)$$

Пусть $N(n, m)$ — наибольший размер двоичного кода длины n и диаметра m , $0 \leq m \leq n$. Определим \mathcal{F} как семейство кодов длины n , диаметра m и размера $N(n, m)$. Для произвольного кода $C \in \mathcal{F}$ определим подкоды

$$C_i := \{\mathbf{v} \mid \mathbf{v} \in C, v_n = i\} \quad (2)$$

при $i \in \{0, 1\}$.

Утверждение 1. Для произвольных n и m , $0 < m < n$, существует такой код $C \in \mathcal{F}$, что $d(\mathbf{v}, \mathbf{w}) \leq m - 2$ для всех \mathbf{v}, \mathbf{w} из C_1 .

Доказательство. Возьмём произвольный код B из \mathcal{F} . Если B удовлетворяет вышеуказанному неравенству, то полагаем $C := B$. Если B не удовлетворяет этому неравенству, преобразуем B в другой код из \mathcal{F} , который удовлетворяет этому неравенству.

А. Сначала преобразуем код B в код D из \mathcal{F} такой, что D_1 не содержит пары слов, расстояние между которыми равно m . Пусть $S \subseteq B_1$ определено следующим способом:

$$S = \{\mathbf{x} \mid \mathbf{x} \in B_1, \exists \mathbf{y} \in B_1 \text{ такое, что } d(\mathbf{x}, \mathbf{y}) = m\}. \quad (3)$$

Ясно, что либо $S = \emptyset$, либо $|S| \geq 2$. Если $S \neq \emptyset$, для каждого слова $\mathbf{x} \in S$ определим слово \mathbf{x}' с $x'_j = x_j$, $1 \leq j < n$, и $x'_n = 0$. Ни одно из этих слов \mathbf{x}' не принадлежит B , так как условие $d(\mathbf{x}, \mathbf{y}) = m$ в (3) означает, что $d(\mathbf{x}, \mathbf{y}) = m + 1$. Кроме того, для любого \mathbf{x}' имеем

$$d(\mathbf{x}', \mathbf{v}) = d(\mathbf{x}, \mathbf{v}) - 1 \leq m - 1, \text{ if } \mathbf{v} \in B_0, \quad (4)$$

$$d(\mathbf{x}', \mathbf{v}) = d(\mathbf{x}, \mathbf{v}) + 1 \leq (m - 1) + 1 = m, \text{ if } \mathbf{v} \in B_1 \setminus S, \quad (5)$$

$$d(\mathbf{x}', \mathbf{v}') = d(\mathbf{x}, \mathbf{v}) \leq m, \text{ if } \mathbf{v} \in S. \quad (6)$$

Следовательно, заменив все $\mathbf{x} \in S$ на соответствующие \mathbf{x}' , получим код D , который обладает свойством, указанным в начале доказательства.

В. Теперь преобразуем код D в код $C \in \mathcal{F}$ такой, что C_1 не содержит пары кодовых слов на расстоянии, большем $m - 2$.

Пусть $T \subseteq D_1$ определяется следующим образом:

$$T = \{\mathbf{x} | \mathbf{x} \in D_1, \exists \mathbf{y} \in D_1 \text{ with } d(\mathbf{x}, \mathbf{y}) = m - 1, x_1 = y_1\}. \quad (7)$$

Очевидно, что T является объединением двух непересекающихся множеств T_0 и T_1 таких, что в T_i имеем $x_1 = y_1 = i$, а $|T_i| = \emptyset$, либо $|T_i| \geq 2$ при $i = 0, 1$. Если $T \neq \emptyset$, для каждого $\mathbf{x} \in T$ определим слово \mathbf{x}' с $x'_1 = x_1 + 1 \pmod{2}$, $x'_j = x_j$ при $1 < j < n$ и $x'_n = 0$. Ни одно из этих слов \mathbf{x}' не принадлежит D , поскольку $d(\mathbf{x}, \mathbf{y}) = m - 1$ и $x_1 = y_1$ в (7) означает, что $d(\mathbf{x}', \mathbf{y}) = (m - 1) + 2 = m + 1$. Кроме того, для каждого \mathbf{x}' выполнены соотношения:

$$d(\mathbf{x}', \mathbf{v}) \leq d(\mathbf{x}, \mathbf{v}) + 1 - 1 \leq m, \quad \text{если } \mathbf{v} \in D_0, \quad (8)$$

$$d(\mathbf{x}', \mathbf{v}) \leq d(\mathbf{x}, \mathbf{v}) + 1 \leq (m - 1) + 1 = m, \quad \text{если } \mathbf{v} \in D_1 \setminus T, \quad (9)$$

$$d(\mathbf{x}', \mathbf{v}') \leq d(\mathbf{x}, \mathbf{v}) \leq m - 1, \quad \text{если } \mathbf{v} \in T. \quad (10)$$

Следовательно, заменив все $\mathbf{x} \in T$ на соответствующие \mathbf{x}' , получим код $E^{(1)} \in \mathcal{F}$ такой, что $d(E^{(1)}) \leq m - 1$, и если $E^{(1)}_1$ содержит пару (\mathbf{x}, \mathbf{y}) с $d(\mathbf{x}, \mathbf{y}) = m - 1$, то $x_1 \neq y_1$. Аналогично, удалив из $E^{(1)}_1$ пары (\mathbf{x}, \mathbf{y}) с $d(\mathbf{x}, \mathbf{y}) = m - 1$ и $x_2 = y_2$, получим код $E^{(2)} \in \mathcal{F}$ такой, что $d(E^{(2)}) \leq m - 1$, и если $E^{(2)}_1$ содержит пару (\mathbf{x}, \mathbf{y}) с $d(\mathbf{x}, \mathbf{y}) = m - 1$, то $x_1 \neq y_1$ и $x_2 \neq y_2$. Продолжая этот процесс, получаем код $E^{(n-1)} \in \mathcal{F}$ с $d(E^{(n-1)}) \leq m - 1$, обладающий тем свойством, что если $E^{(n-1)}_1$ содержит пару слов \mathbf{x} и \mathbf{y} с $d(\mathbf{x}, \mathbf{y}) = m - 1$, то $x_1 \neq y_1, x_2 \neq y_2, \dots, x_{n-1} \neq y_{n-1}$. Так как $x_n = y_n = 1$, то $d(\mathbf{x}, \mathbf{y}) = n - 1$. Следовательно, $m = n$. Последнее равенство противоречит условию $m < n$. Поэтому $E^{(n-1)}_1$ не содержит пары слов \mathbf{x} и \mathbf{y} с $d(\mathbf{x}, \mathbf{y}) > m - 2$. Следовательно, $C := E^{(n-1)}$ является кодом, который удовлетворяет утверждению. Утверждение 1 доказано.

Замечание. Использованный выше метод аналогичен методу “pushing-techniques” из [1, 2].

Следующее свойство является непосредственным следствием утверждения 1. Его доказательство проводится удалением последней координаты во всех словах кода C .

Следствие Максимальный размер $N(n, m)$ двоичного кода длины n и диаметра m , $1 < m < n$, удовлетворяет рекуррентному соотношению

$$N(n, m) \leq N(n-1, m) + N(n-1, m-2).$$

Теперь всё готово к доказательству известного результата, полученного Д. Клейтманом [5] и Д. Катоней [4].

Теорема. Пусть C — двоичный код длины $n \geq 1$ с диаметром m , $1 \leq m \leq n$, который имеет максимальный размер $N(n, m)$. Тогда $N(n, n) = 2^n$, а при $m < n$

$$N(n, m) = \begin{cases} \sum_{i=0}^k \binom{n}{i}, & \text{если } m \text{ чётно,} \\ \sum_{i=0}^k \binom{n}{i} + \binom{n-1}{k}, & \text{если } m \text{ нечётно,} \end{cases}$$

где $k = \left\lfloor \frac{m}{2} \right\rfloor$.

Доказательство. Равенство $N(n, n) = 2^n$ тривиально. Для $m < n$ мы сначала докажем более слабые соотношения:

$$N(n, m) \leq \begin{cases} \sum_{i=0}^k \binom{n}{i}, & \text{если } m \text{ чётно,} \\ \sum_{i=0}^k \binom{n}{i} + \binom{n-1}{k}, & \text{если } m \text{ нечётно.} \end{cases} \quad (11)$$

Предположим, что неравенства (11) верны для всех значений, меньших чем некоторое фиксированное $n \geq 1$. Пусть m чётно. По следствию и индуктивному предположению при $m < n-1$ имеем

$$\begin{aligned} N(n, m) &\leq \sum_{i=0}^k \binom{n-1}{i} + \sum_{i=0}^{k-1} \binom{n-1}{i} \\ &= 1 + \sum_{i=1}^k \binom{n-1}{i} + \sum_{i=1}^k \binom{n-1}{i-1} = \sum_{i=0}^k \binom{n}{i}. \end{aligned}$$

Поскольку код, содержащий все слова из шара радиуса k с некоторым фиксированным центром (скажем $\mathbf{0}$) имеет точно такое число слов, то правая часть приведённого неравенства справедлива. Таким образом, теорема доказана для n в случае, когда m чётно, и $0 < m < n - 1$. Чтобы завершить доказательство теоремы в случае, когда m чётно, мы докажем теорему для $m = n - 1$ и нечётного n . Единственное ограничение на код C в этом случае: если слово принадлежит C , то его дополнение (отрицание в двоичном случае) не принадлежит C . По принципу Дирихле имеем $N(n, n - 1) = 2^{n-1} = \sum_{i=0}^k \binom{n}{i}$ при $k = (n - 1)/2$.

Доказательство для нечётных m аналогично. Теорема верна также для значения n . Тем самым, поскольку при $n = 1$ теорема тривиальна, мы доказали ее для всех $n \geq 1$ по индукции. Теорема доказана.

2. Обобщение на q -значные антикоды

Наш метод может быть легко обобщён для получения верхней оценки для наибольшего размера q -значного антикода длины n с максимальным расстоянием m . Обозначив эту величину через $N_q(n, m)$, аналогично неравенствам (11) получаем соотношения

$$N_q(n, m) \leq \begin{cases} \sum_{i=0}^k \binom{n}{i} (q-1)^i, & \text{если } m \text{ чётно,} \\ \sum_{i=0}^k \binom{n}{i} (q-1)^i + \binom{n-1}{k} (q-1)^{k+1}, & \text{если } m \text{ нечётно.} \end{cases} \quad (12)$$

Однако верхняя оценка (12) не является точной для всех целых чисел m из интервала $[1, n - 1]$. Для того чтобы получить равенства в (12) при $1 \leq m \leq m_0$, где m_0 — целое число, введённое в [1, 2], мы должны обобщить аргументы, использованные в последней части доказательства теоремы для такого m_0 . Это находится на стадии исследования.

ЛИТЕРАТУРА

1. Ahlswede R., Cai N., Zhang Z. Diametric theorems in sequence spaces // Combinatorica. 1992. V. 12, N 1. P. 1–17.
2. Ahlswede R. F., Khachatrian L. H. The diametric theorem in Hamming spaces—optimal anticodes // Adv. in Appl. Math. 1998. V. 20, N 4. P. 429–449.

3. **Farrell P. G.** Linear binary anticodes // Electronic Letter. 1970. V. 6, N 13. P. 419–421.
4. **Katona G. O. H.** Intersection theorem for systems of finites sets // Acta Math. Acad. Sci. Hungary. 1964. V. 15, N 3–4. P. 329–337.
5. **Kleitman D. J.** On a combinatorial conjecture of Erdös // J. Comb. Theory. 1966. V. 1, N 1. P. 209–214.
6. **Maki G. K., Tracey J. H.** Maximum distance linear codes // IEEE Trans. Inform. Theory. 1971. V. 17, N 5. P. 632.
7. **Reddy S. M.** On block codes with specified maximum distance // IEEE Trans. Inform. Theory. 1973. V. 18, N 6. P. 823–824.

Адрес автора:
Department of Mathematics
Faculty of Electrical Engineering,
Mathematics and Computer Sciences
Delft University of Technology
P.O. BOX 5031, 2600 GA Delft,
The Netherlands

Статья поступила
30 июня 2004 г.