ԵՐԵՎԱՆԻ ՊԵՏԱԿԱՆ ՀԱՄԱԼՍԱՐԱՆԻ ԳԻՏԱԿԱՆ ՏԵՂԵԿԱԳԻՐ
УЧЕНЫЕ ЗАПИСКИ ЕРЕВАНСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА

Բնական գիտություններ 3, 1984 Естественные науки

Математика

УДК 621.391.15

Ж. Г. МАРГАРЯН

КОДЫ С ЗАДАННЫМ МНОЖЕСТВОМ РАССТОЯНИЙ

Подмножество (код) V единичного n-мерного куба характеризуется множеством расстояний R(V) и параметрами n, M, где R(V)—множество значений, принимаемых Хэмминговым расстоянием между различными кодовыми словами; n—длина; M—число наборов кода. Обозначим через M(n; R) максимум параметра M по всем кодам V длины n и с множеством расстояний $R(V) \subseteq R$. При использовании информации о строении множества R нам удалось в ряде случаев улучшить известную верхнюю оценку метрического функционала M(n; R).

Рассмотрим n-мерное векторное пространство B_n над [1] полем GF(2). Элементы этого пространства будем называть n-точками и обозначать их через $\vec{\alpha}$, $\vec{\beta}$ и т. д. или в координатной записи через

$$\vec{\alpha}=(\alpha_1, \alpha_2, \ldots, \alpha_n) \text{ или } \vec{\beta}=(\beta_1, \beta_2, \ldots \beta_n),$$

где α_i, β_i равны 0 или 1. Геометрически n-точки представляют собой вершины n-мерного куба в Евклидовом пространстве E_n. Операция сложения $\oplus$ двух n-точек $\vec{\alpha}$ и $\vec{\beta}$ естественно определяется соотношением

$$\vec{\gamma}=\vec{\alpha}\oplus\vec{\beta},$$

где $\gamma_i=\alpha_i\oplus\beta_i=\alpha_i+\beta_i \pmod 2$.

Расстояние по Хэммингу между n-точками $\vec{\alpha}$ и $\vec{\beta}$ определим по формуле

$$d(\vec{\alpha}, \vec{\beta})=\sum_{i=1}^{n}(\alpha_i \oplus \beta_i).$$

Вес n-точки $\vec{\alpha}$ обозначим через $\|\vec{\alpha}\|$ и определим его как положительное число

$$\|\vec{\alpha}\|=\sum_{i=1}^{n}\alpha_i.$$

Обозначим через B_n^k подмножество n-точек веса k.

Кодом длины n над полем GF(2) является любое непустое подмножество V множества B_n. Элементы подмножества (кода) V называются кодовыми словами. Код V характеризуется множеством R(V) и параметрами n, M, где R(V)—множество значений, принимаемых расстоянием между различными кодовыми словами, т. е. $a \in R(V)$ тогда и только тогда, когда существуют такие различные кодовые слова $\vec{\alpha}$, $\vec{\beta}$ из V,

что $d(\vec{\alpha}, \vec{\beta})=a$; n—длина кода; M—число кодовых слов кода V (мощность кода).

Точная постановка задачи, которая рассматривается в данной работе, состоит в следующем: заданы параметр n и множество R. Требуется узнать, какое максимальное значение может принимать параметр M и какой именно код является оптимальным в смысле максимальной мощности.

Предположим, что $R=(d_1, d_2, ..., d_s)$—некоторое подмножество множества $A=\{1, 2, ...n\}$. Обозначим через F(n; R) семейство V кодов таких, что $R(V)\subseteq R$. Пусть

$$M(n; R)=\max_{V\in F(n;R)} |V|,$$

где |V| обозначает мощность множества V.

Относительно метрического функционала M (n; R) известна [2] следующая верхняя оценка, имеющая место для всех R одинаковой мощности, а именно

$$M(n; R)\leqslant 1+C_n^1+C_n^2+ \dots C_n^s,$$

где $s=|R|$.

При использовании информации о строении множества R автору в ряде случаев удалось улучшить приведенную оценку. Основные результаты этой работы были опубликованы без доказательств в [3].

Теорема 1. Если любой элемент множества R является числом вида 4p+2, то имеет место следующее неравенство:

$$M(n; R)\leqslant |R|(n+1)+1.$$

Доказательство. Сначала докажем следующее утверждение: если $(x_1, x_2, ...x_t)$ такое подмножество из E_n, что скалярное произведение любых двух векторов—четное число, а норма каждого вектора—нечетное, то $t\leqslant n$. Для этого достаточно доказать, что эти векторы линейно независимы в n-мерном эвклидовом пространстве E_n. Рассмотрим детерминант Грама [4]. Элементы главной диагонали этой матрицы являются нечетными числами, а все остальные элементы—четными. Это значит, что детерминант этой матрицы невырожден. Отсюда следует, что $t\leqslant n$. Не нарушая общность, предположим, что код $V\in F(n; R)$ содержит нулевую n-точку. Обозначим через T_{d_i} множество кодовых слов, веса которых равны d_i; так как $d_i\equiv 2 (\mathrm{mod}\ 4)$, то скалярное произведение любых двух кодовых слов из T_{d_i} является нечетным числом. Если в конце каждого кодового слова в коде T_{d_i} добавим единицу, то получается код из $B_{n+1}^{d_i+1}$ такой, что скалярное произведение любых двух кодовых слов—четное число. Это значит, что $|T_{d_i}|\leqslant n+1$. Так как $|V|=\sum_{i=1}^{s}|T_{d_i}|+1$, то

$$M(n; R)\leqslant |R|(n+1)+1.$$

Следствие. Если код $V\subseteq B_n$ такой, что $|V|>2\cdot\left\lfloor\frac{n+2}{4}\right\rfloor(n+1)+3$, то множество R(V) содержит элемент, кратный 4.

Линейным (n, k) кодом называется любое k-мерное подпространство n-мерного векторного пространства B_n. Число k называется размерностью линейного (n, k) кода. Линейный (n, k) код можно задавать при помощи k линейно независимых n-точек, и код состоит из всех возможных линейных комбинаций этих n-точек над полем GF(2). Следо-

вательно, линейный код содержит всего 2^k кодовых слов. Порождающей матрицей линейного (n, k) кода называется любая матрица G размерности $(k \times n)$, содержащая в качестве строк k линейно независимых кодовых слов. Поскольку порождающая матрица имеет k строк и чисто нулевой столбец может быть исключен из рассмотрения как бесполезный, то существует всего 2^k-1 различных типов возможных столбцов. Если какой-то столбец матрицы G является двоичной записью числа j $(1 \leqslant j \leqslant 2^k-1)$, то он называется столбцом типа j. Если **не обращать** внимания на порядок расположения столбцов, то линейный код можно однозначно задать набором неотрицательных чисел

$$\bar{n}=(n^1, n^2, \ldots, n^{2^k-1}),$$

где n^j —число столбцов типа j в матрице G, называемом модулярным представлением линейного (n, k) кода.

Следующие два результата относятся к таким кодам, для которых расстояние между различными кодовыми словами равно d_1 или d_2 $(d_1 > d_2)$.

Любое натуральное число N можно представить в виде $N=P \cdot 2^q$, где P—нечетное число. Когда N—нечетное число, $P=N$, $q=0$.

Предположим, что $d_i = P_{d_i} \cdot 2^{q_{d_i}}$ при $i=1, 2$.

Теорема 2. Если существует линейный (n, k) код, такой, что расстояние между различными кодовыми словами равно d_1 или d_2, то имеет место неравенство

$$k \leqslant \log_2 \frac{2d_1 \cdot d_2}{P_{d_1} \cdot P_{d_2}}.$$

Доказательство. Рассмотрим матрицу C порядка $((2^k-1) \times n)$ из нулей и единиц, строки которой являются все ненулевые кодовые слова линейного (n, k) кода. В каждой строке матрицы C число единиц равно d_1 или d_2, а число единиц в каждом столбце равно 2^{k-1}. Это значит, что число строк в матрице C, имеющих веса d_1 (d_2), равно

$$\frac{(n-2d_2)2^{k-1}+d_2}{d_1-d_2} \cdot \left(\frac{(2d_1-n)2^{k-1}-d_1}{d_1-d_2} \right).$$

Хэммингово расстояние между i-той и j-той столбцами обозначим через $d(i; j)$. Вычислим сумму $\sum_{i \neq j} d(i; j)$ двумя способами.

Во-первых,

$$\sum_{i \neq j} d(i; j) = 2^{k-1} \left[\frac{n(n-1)}{2} - \sum_{i=1}^{2^k-1} \frac{n^i(n^i-1)}{2} \right].$$

С другой стороны, если i-ая строка матрицы C содержит $d_1(d_2)$ единиц, то вклад ее в нашу сумму равен $d_1(n-d_1)$, $(d_2(n-d_2))$, и поэтому эта сумма равна

$$\sum_{i \neq j} d(i; j) = d_1(n-d_1) \left[\frac{(n-2d_2)2^{k-1}+d_2}{d_1-d_2} \right] + d_2(n-d_2) \left[\frac{(2d_1-n)2^{k-1}-d_1}{d_1-d_2} \right],$$

Следовательно,

$$\frac{4d_1d_2}{2^k} = (n-2d_1)(n-2d_2) + \sum_{i=1}^{2^k-1} (n^i)^2.$$

Так как $\sum_{i=1}^{2^k-1} n^i = n$, то из этого равенства получаем

$$\frac{2d_1d_2}{2^k}=\sum_{i=1}^{2^k-1}(n^i)^2+\sum_{i\neq j}n^i\cdot n^j-n(d_1+d_2)+2d_1d_2.$$

Это значит, что $\frac{2d_1d_2}{2^k}$—целое число, т. е. имеет место неравенство

$$k\leqslant\log_2\frac{2d_1\cdot d_2}{p_{d_1}\ p_{d_2}}.$$

Замечание. Для некоторых d_1 и d_2 существует семейство линейных кодов (напр., $d_1=2^{\kappa-1}$, $d_2=2^\kappa-1$ и линейные коды с модулярным представлением $\vec{n}=(0,\underbrace{1,1,1,\ldots1,}_{2^k-2})$), для которых достигается граница, задаваемая в теореме 2.

Теорема 3. Имеет место неравенство

$$M(n;\{d_1;n-d_1\})\leqslant1+\frac{n(n-1)}{2}.$$

Если имеет место равенство, то существует матрица Адамара порядка $1+\frac{(n-1)}{2}$.

Доказательство. Обозначим через G_2 матрицу, столбцами которой являются все n-точки из B_n^2. Предположим, что $\vec{\alpha}\in B_n$. Тогда вес $\frac{n(n-1)}{2}$ точки $\vec{\alpha}G_2$ (все операции в произведении αG_2 выполняются над полем GF(2)) равен $d(n-d)$, где $d=\|\vec{\alpha}\|$. А это значит, что $d(\vec{\alpha}G_2,\ \vec{\beta}G_2)=d(\vec{\alpha},\ \vec{\beta})\ (n-d(\vec{\alpha},\ \vec{\beta}))$. Если код $V=\{\vec{\alpha},\ \vec{\beta},\ \vec{\gamma},\ \ldots\}$ такой, что расстояние между различными кодовыми словами равно d_1 или $n-d_1$, то код $VG_2=\{\vec{\alpha}G_2,\ \vec{\beta}G_2,\ \vec{\gamma}G_2\ldots\}$ является эквидистантным кодом [2] с расстоянием $d_1(n-d_1)$.

Из теорем 4,2 и 4,3 работы [2] следует

$$M(n;\{d_1;n-d_1\})\leqslant1+\frac{n(n-1)}{2}$$

притом если имеет место равенство, то существует матрица Адамора порядка $1+\frac{n(n-1)}{2}$, что и требовалось доказать.

Лаборатория теоретической кибернетики ЕГУ *Поступила 4.11.1983*

ЛИТЕРАТУРА

1. **Мак-Вильямс Ф. Дж., Слоэн Н. Дж. А.** Теория кодов, исправляющих ошибки. М.: Связь. 1979.
2. **Delsarte P.** Four fundamental parameters of a Code and their combinatorial significance.—Info. and Control, 1973, № 23, p. 407—438.
3. **Маргарян Ж. Г., Мовсисян Г. Л.** Метод построения линейных кодов.—ДАН Арм. ССР, 1984, LXXVIII, № 3.
4. **Воеводин В. В.** Линейная алгебра. М.: Наука, 1980.

Ժ. Գ. ՄԱՐԳԱՐՅԱՆ

ՏՎԱԾ ՀԵՌԱՎՈՐՈՒԹՅՈՒՆՆԵՐԻ ԲԱԶՄՈՒԹՅՈՒՆՈՎ ԿՈԴԵՐ

Ամփոփում

n-չափանի միավոր խորանարդի ցանկացած ներքաբազմություն կոչվում է կոդ, իսկ նրա տարրերը կոդային բառեր։ Երկու իրարից տարբեր կոդային բառերի հեմմինգյան հեռավորության արժեքների բազմությունը կոչվում է կոդի հեռավորությունների բազմություն։ Նշանակենք M(n;R)-ով այն կոդի հզորությունը, որը ունի ամենամեծ հզորությունը և որի հեռավորությունների բազմությունը հանդիսանում է R-ի ներքաբազմություն։ R բազմության վրա որոշակի սահմանափակումների դեպքում աշխատանքում հաջողվել է լավացնել M(n;R) մետրիկական ֆունկցիոնալի հայտնի [2] վերին գնահատականը։