INFORMATION AND CONTROL 1, 289-295 (1958)

A Note on Upper Bounds for Minimum Distance Codes

D. D. Joshi

Indian Statistical Institute, Calcutta, India

Some upper bounds are given on the number of sequences of n binary symbols which can be found such that every pair of sequences differs in at least d positions.

Let C_n denote the set of all sequences of length n formed out of the symbols 0 and 1. Any element α of C_n can be written as

$$\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n); \qquad \alpha_i = 0 \text{ or } 1.$$

The norm $\| \alpha \|$ of an element α is defined to be the positive number

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

The set C_n is an Abelian group if "addition" (noted $\oplus$) is defined as

$$\alpha \oplus \beta = (\alpha_1 \oplus \beta_1, \alpha_2 \oplus \beta_2, \cdots, \alpha_n \oplus \beta_n)$$

the addition of the numbers α_i and β_i being addition modulo 2. If the identity or the zero element of the group is noted as ϕ.

$$\phi = (0, 0, \cdots, 0),$$

then

$$\alpha \oplus \alpha = \phi$$

for each α in C_n.

Let $\delta(\alpha, \beta)$ denote the Hamming distance (Hamming, 1950) between two elements α and β of C_n. Then

$$\delta(\alpha, \beta) = \| \alpha \oplus \beta \|$$

A subset of C_n such that for any two elements $\alpha, \beta (\alpha \neq \beta)$ belonging to it

$$\| \alpha \oplus \beta \| \geqq d$$

will be called a code with minimum distance d. If the subset is a subgroup of C_n it shall be called a group code. The term group code used here should be distinguished from the group alphabet of Slepian (1956). The latter does not depend on the idea of a minimum distance. Minimum distance codes shall be denoted by $M(n, d)$, and group codes by $G(n, d)$.

We shall denote by $C_{n,r}$ (r being an integer less than n) the subgroup of C_n obtained by adding $n - r$ zeros to each element of C_r. $C_{n,r}$ contains 2^r elements. The number of elements of any subset E of C_n shall be denoted by $[E]$. Thus

$$[C_r] = 2^r.$$

Given a fixed element α and a subset E of C_n, the set of elements

$$\alpha \oplus \beta, \qquad \beta \varepsilon E$$

shall be denoted by $E \oplus \alpha$. If E is a code with minimum distance d, so is the set $E \oplus \alpha$. Similarly, the set E' obtained by any permutation of the columns of a minimum distance code E (E being written as a matrix with n columns) is also a minimum distance code.

THEOREM 1. $$[M(n, d)] \leqq 2^{n-d+1}.$$

Consider the subgroup $C_{n,d-1}$ of C_n. For any two elements $\alpha, \beta (\alpha \neq \beta)$ of $C_{n,d-1}$

$$\| \alpha \oplus \beta \| \leqq d - 1.$$

The sets $M \oplus \alpha$, $M \oplus \beta$ are therefore disjoint. Hence

$$[M(n, d)]2^{d-1} \leqq 2^n$$

which proves the theorem. This result had been obtained earlier by Komamiya (1954) by using a more complicated method of proof.

THEOREM 2. For $n > d > 2$

$$[G(n, d)] \leqq 2^{n-d}.$$

Consider any two elements α, β $(\alpha \neq \beta)$ of $C_{n,d}$. Then

$$\| \alpha \oplus \beta \| \leqq d.$$

If the inequality strictly holds, the sets $G \oplus \alpha$, $G \oplus \beta$ are disjoint. Consider next the case when

$$\| \alpha \oplus \beta \| = d.$$

If $\alpha \oplus \beta$ does not belong to G the sets $G \oplus \alpha$ and $G \oplus \beta$ are still disjoint. Hence if the element

$$\xi = (\underbrace{1, 1, \cdots, 1}_{d}, \underbrace{0, 0, \cdots, 0}_{n-d})$$

does not belong to G, all the sets $G \oplus \alpha$ where α belongs to $C_{n,d}$ are disjoint and we have

$$[G(n, d)]2^d \leqq 2^n$$

If ξ belongs to G, then by interchanging the first and the last columns of G we obtain a code G' with the same minimum distance and having the same number of elements as G. If ξ does not belong to G' the theorem is true. If ξ belongs to G', then G must have an element

$$\eta = (0, \underbrace{1, 1, \cdots, 1}_{d-1}, \underbrace{0, 0, \cdots, 0}_{n-d-1}, 1)$$

But this is impossible if $d > 2$, for

$$\| \xi \oplus \eta \| = 2.$$

We now state without proof a lemma which will be used to obtain a more general form of Theorem 2.

LEMMA 1.

$$\| \alpha \oplus \beta \| \leqq \min (\| \alpha \| + \| \beta \|, \quad 2n - \| \alpha \| - \| \beta \|)$$

THEOREM 3. If $n > d + r$, $d > 2r + 2$ (r being an integer), then

$$[G(n, d)] \leqq 2^{n-(d+r)}.$$

For any two elements α, $\beta(\alpha \neq \beta)$ of $C_{n,d+r}$ for which $\alpha \oplus \beta$ does not belong to G, the sets $G \oplus \alpha$, $G \oplus \beta$ are disjoint. Consider the case where $\alpha \oplus \beta$ belongs to G. We must have then $\| \alpha \oplus \beta \| \geqq d$. Let E be the set of elements of $C_{n,d+r}$ of norm greater than or equal to d. Suppose G contains an element ξ of E. We can, without loss of generality, suppose that ξ is of the form

$$\xi = (\underbrace{1, \text{at least } d-1 \text{ one's}}_{d+r-1}, \underbrace{0, 0, \cdots, 0}_{n-d-r})$$

Let G' be the code obtained from G by interchanging the first and the last columns. If G' has no element of E our result is proved. If on the other hand G' contains an element of E then G must have either an element η of the form

$$\eta = (0, \underbrace{\text{at least } d \text{ one's}}_{d+r-1}, \underbrace{0, 0, \cdots, 0}_{n-d-r})$$

or an element ζ of the form

$$\zeta = (\underbrace{0, \text{at least } d - 1 \text{ one's}}_{d+r-1}, \underbrace{0, 0, \cdots, 0, 1}_{n-d-r}).$$

But this is impossible, for we have (using Lemma 1)

$$\| \xi \oplus \eta \| \leqq 1 + 2(d + r - 1) - (d + d - 1) = 2r$$

and

$$\| \xi \oplus \zeta \| \leqq 2 + 2(d + r - 1) - (d - 1 + d - 1) = 2 + 2r.$$

The theorem now follows without difficulty.

Theorems 2 and 3 are valid only for group codes. The following theorem holds for all minimum distance codes.

THEOREM 4. If d is an odd number and if $2d + 1 > n$, then

$$[M(n, d)] \leqq (2d + 2)/(2d + 1 - n).$$

Let the elements of M

$$\alpha_1, \alpha_2, \cdots, \alpha_m; \qquad m = [M(n, d)]$$

be written in the form of an $m \times n$ matrix and let k_i denote the sum of the ith column ($i = 1, 2, \cdots, n$). We have then (cf. Schützenberger, 1953)

$$n \cdot \text{variance } (k_i) = mA - (A^2/n) - \sum_{\alpha_j \neq \alpha_k} \delta(\alpha_j, \alpha_k)$$

where

$$A = \sum_i k_i = \sum_j \| \alpha_j \|$$

This gives

$$\sum_{\alpha_j \neq \alpha_k} \delta(\alpha_j, \alpha_k) \leqq mA - (A^2/n)$$

Since A lies between 0 and mn the quantity on the right is bounded by $(m^2n)/4$. Also

$$\delta(\alpha_j, \alpha_k) \geqq d$$

Hence

$$\frac{m(m-1)}{2} d \leqq \frac{m^2 n}{4}$$

or,

$$m \leqq 2d/(2d - n)$$

if $2d > n$.

If d is an odd number, let r and s be the numbers of elements of M of odd and even norms, respectively. The distance between two elements is an even number if both are of odd or even norm and is an odd number otherwise. We have thus

$$\{[r(r-1)/2] + [s(s-1)/2]\}(d+1) + rsd \leqq m^2n/4$$

which finally gives

$$m \leqq (2d+2)/(2d+1-n)$$

if $2d + 1 > n$.

I am grateful to the referee for pointing out that this result had already been obtained by Plotkin (1951) and for making available to me a copy of Plotkin's article. Since this article is not easily available it has been considered useful to summarize his results below.

Plotkin proves that for ($n > 2d$,) $2d > n$

$$[M(n, d)] \leqq 2d/(2d - n)$$

This result is equivalent to that proved above in the sense that if it is used for even values of d and the upper bound for odd values of d then obtained by using the well-known result

$$[M(n, 2k-1)] = [M(n+1, 2k)]$$

we get the same value as by applying Theorem 4 directly to odd values of d. Plotkin also proves a slightly stronger version, namely, that

$[M(n, d)] \leqq$ largest even integer contained in $2d/(2d - n)$.

Apart from this theorem he gives the following inequalities

$$[M(n, d)] \leqq 2[M(n - 1, d)]$$

$$[M(2n, 2d) \geqq [M(n, 2d)][M(n, d)]$$

He uses these to prove what is, in the author's opinion, the most important result available on the number of elements in a binary minimum distance code. His theorem is stated below in a slightly different form.

THEOREM 5 (Plotkin). If m is an integer such that $4m - 1$ is a prime, then

$$[M(4m, 2m)] \leqq 8m,$$

and a minimum distance code exists for which the equality is true.

The first part of the theorem follows from the earlier results. The second part is proved by constructing a code as follows.

It is known from elementary number theory that exactly half of the integers $1, 2, \cdots, p - 1$ are quadratic residues and half quadratic nonresidues of a prime p, and that -1 is a quadratic nonresidue of all primes of the form $4m - 1$. The numbers z_i are now defined as

$$\begin{aligned} z_i &= 1 \text{ if } i \text{ is a quadratic residue of } 4m - 1 \\ z_i &= 0 \text{ if } i \text{ is a quadratic nonresidue of } 4m - 1 \\ z_i &= 1 \text{ if } i \equiv 0 \end{aligned}$$

The z_i's are used to construct a code $M(4m - 1, 2m)$ of $4m - 1$ elements as follows. Code element a_1 is define as

$$a_1 = (z_1, z_2, \cdots, z_{4m-1}),$$

$a_2, a_3, \cdots, a_{4m-1}$ are then obtained by cyclic permutations of a_1. It is shown that this is a $M(4m - 1, 2m)$ code.

Adding a zero to each of these we get the set of elements, say,

$$b_1, b_2, \cdots, b_{4m-1}.$$

The required code of $8m$ elements is then given by

$$\phi, b_1, b_2, \cdots, b_{4m-1}$$

$$I, I \oplus b_1, I \oplus b_2, \cdots, I \oplus b_{4m-1},$$

where

$$I = (1, 1, \cdots, 1).$$

RECEIVED: March 3, 1958.

References

Hamming, R. W. (1950). Error detecting and error correcting codes. *Bell System Tech. J.*, **29**, 147–160.

Komamiya, Y. (1954). Application of logical mathematics to information theory. *Proc. 3rd Japan Natl. Congr. Appl. Math. 1953* p. 437.

Plotkin, M. (1951). Binary codes with specified minimum distance. Research Div. Rept. 51–20. The University of Pennsylvania Moore School of Electrical Engineering, Philadelphia, Pennsylvania.

Schützenberger, M. P. (1953). Sur un problème du codage binaire. *Publ. Inst. statist. Univ. Paris* **2**, 125–127.

Slepian, D. (1956). A class of binary signalling alphabets. *Bell System Tech. J.* **35**, 203–234.