Links between discriminating and identifying codes in the binary Hamming space

I. Charon, G. Cohen, O. Hudry, A. Lobstein

Links between Discriminating and Identifying Codes in the Binary Hamming Space

Irène Charon
GET - Télécom Paris & CNRS - LTCI UMR 5141
46, rue Barrault, 75634 Paris Cedex 13 - France
irene.charon@enst.fr

Gérard Cohen
GET - Télécom Paris & CNRS - LTCI UMR 5141
46, rue Barrault, 75634 Paris Cedex 13 - France
gerard.cohen@enst.fr

Olivier Hudry
GET - Télécom Paris & CNRS - LTCI UMR 5141
46, rue Barrault, 75634 Paris Cedex 13 - France
olivier.hudry@enst.fr

Antoine Lobstein
CNRS - LTCI UMR 5141 & GET - Télécom Paris
46, rue Barrault, 75634 Paris Cedex 13 - France
antoine.lobstein@enst.fr

Abstract

Let F^n be the binary n-cube, or binary Hamming space of dimension n, endowed with the Hamming distance, and $\mathcal{E}^n$ (respectively, $\mathcal{O}^n$) the set of vectors with even (respectively, odd) weight. For $r \geq 1$ and $x \in F^n$, we denote by $B_r(x)$ the ball of radius r and centre x. A code $C \subseteq F^n$ is said to be r-identifying if the sets $B_r(x) \cap C$, $x \in F^n$, are all nonempty and distinct. A code $C \subseteq \mathcal{E}^n$ is said to be r-discriminating if the sets $B_r(x) \cap C$, $x \in \mathcal{O}^n$, are all nonempty and distinct. We show that the two definitions, which were given for general graphs, are equivalent in the case of the Hamming space, in the following sense: for any odd r, there is a bijection between the set of r-identifying codes in F^n and the set of r-discriminating codes in F^{n+1}.

Key Words: Graph Theory, Coding Theory, Discriminating Codes, Identifying Codes, Hamming Space, Hypercube

1 Introduction

We define identifying and discriminating codes in a connected, undirected graph $G = (V, E)$, in which a *code* is simply a nonempty subset of vertices. These definitions can help, in various meanings, to unambiguously determine a vertex. The motivations may come from processor networks where we wish to locate a faulty vertex under certain conditions, or from the need to identify an individual, given its set of attributes.

In G we define the usual distance $d(v_1, v_2)$ between two vertices $v_1, v_2 \in V$ as the smallest possible number of edges in any path between them. For an integer $r \geq 0$ and a vertex $v \in V$, we define $B_r(v)$ the *ball* of radius r centred at v, as the set of vertices within distance r from v. Whenever two vertices v_1 and v_2 are such that $v_1 \in B_r(v_2)$ (or, equivalently, $v_2 \in B_r(v_1)$), we say that they *r-cover* each other. A set $X \subseteq V$ r-covers a set $Y \subseteq V$ if every vertex in Y is r-covered by at least one vertex in X.

The elements of a code $C \subseteq V$ are called *codewords*. For each vertex $v \in V$, we denote by

$$K_{C,r}(v) = C \cap B_r(v)$$

the set of codewords r-covering v. Two vertices v_1 and v_2 with $K_{C,r}(v_1) \neq K_{C,r}(v_2)$ are said to be *r-separated* by code C, and any codeword belonging to exactly one of the two sets $B_r(v_1)$ and $B_r(v_2)$ is said to *r-separate* v_1 and v_2.

A code $C \subseteq V$ is called *r-identifying* [10] if all the sets $K_{C,r}(v)$, $v \in V$, are nonempty and distinct. In other words, every vertex is r-covered by at least one codeword, and every pair of vertices is r-separated by at least one codeword. Such codes are also sometimes called *differentiating dominating sets* [8].

We now suppose that G is bipartite: $G = (V = I \cup A, E)$, with no edges inside I nor A — here, A stands for *attributes* and I for *individuals*. A code $C \subseteq A$ is said to be *r-discriminating* [4] if all the sets $K_{C,r}(i)$, $i \in I$, are nonempty and distinct. From the definition we see that we can consider only odd values of r.

In the following, we drop the general case and turn to the binary Hamming space of dimension n, also called the binary n-cube, which is a regular bipartite graph. First we need to give some specific definitions and notation.

We consider the n-cube as the set of binary row-vectors of length n, and as so, we denote it by $G = (F^n, E)$ with $F = \{0, 1\}$ and $E = \{\{x, y\} : d(x, y) = 1\}$, the usual graph distance $d(x, y)$ between two vectors x and y

being called here the *Hamming distance* — it simply consists of the number of coordinates where x and y differ. The *Hamming weight* of a vector x is its distance to the all-zero vector, i.e., the number of its nonzero coordinates. A vector is said to be *even* (respectively, *odd*) if its weight is even (respectively, odd), and we denote by $\mathcal{E}^n$ (respectively, $\mathcal{O}^n$) the set of the 2^{n-1} even (respectively, odd) vectors in F^n. Without loss of generality, for the definition of an r-discriminating code, we choose the set A to be $\mathcal{E}^n$, and the set I to be $\mathcal{O}^n$. Additions are carried coordinatewise and modulo two.

Given a vector $x \in F^n$, we denote by $\pi(x)$ its parity-check bit: $\pi(x) = 0$ if x is even, $\pi(x) = 1$ if x is odd. Therefore, if $|$ stands for concatenation of vectors, $x|\pi(x)$ is an even vector. Finally, we denote by $M_r(n)$ (respectively, $D_r(n)$) the smallest possible cardinality of an r-identifying (respectively, r-discriminating) code in F^n.

In Section 2, we show that in the particular case of Hamming space, the two notions of r-identifying and r-discriminating codes actually coincide for all odd values of r and all $n \geq 2$, in the sense that there is a bijection between the set of r-identifying codes in F^n and the set of r-discriminating codes in F^{n+1}.

2 Identifying is discriminating

As we now show with the following two theorems, for any odd $r \geq 1$, any r-identifying code in F^n can be extended into an r-discriminating code in F^{n+1}, and any r-discriminating code in F^n can be shortened into an r-identifying code in F^{n-1}. First, observe that r-identifying codes exist in F^n if and only if $r < n$.

Theorem 1 *Let $n \geq 2, p \geq 0$ be such that $2p+1 < n$, let $C \subseteq F^n$ be a $(2p+1)$-identifying code and let*

$$C' = \{c|\pi(c) : c \in C\}.$$

Then C' is $(2p+1)$-discriminating in F^{n+1}. Therefore,

$$D_{2p+1}(n+1) \leq M_{2p+1}(n). \tag{1}$$

Proof. Let $r = 2p+1$. By construction, C' contains only even vectors. We shall prove that (a) any odd vector $x \in \mathcal{O}^{n+1}$ is r-covered by at least one codeword of C'; (b) given any two distinct odd vectors $x, y \in \mathcal{O}^{n+1}$, there is at least one codeword in C' which r-separates them.

(a) We write $x = x_1|x_2$ with $x_1 \in F^n$ and $x_2 \in F$. Because C is r-identifying in F^n, there is a codeword $c \in C$ with $d(x_1, c) \leq r$. Let $c' = c|\pi(c)$.

If $d(x_1, c) \leq r-1$, then whatever the values of x_2 and $\pi(c)$ are, we have $d(x, c') \leq r$; we assume therefore that $d(x_1, c) = r = 2p+1$, which implies that x_1 and c have different parities. Since $x_1|x_2$ and $c|\pi(c)$ also have different parities, we have $x_2 = \pi(c)$ and $d(x, c') = r$. So the codeword $c' \in C'$ r-covers x.

(b) We write $x = x_1|x_2$, $y = y_1|y_2$, with $x_1, y_1 \in F^n$, $x_2, y_2 \in F$. Since C is r-identifying in F^n, there is a codeword $c \in C$ which is, say, within distance r from x_1 and not from y_1: $d(x_1, c) \leq r$, $d(y_1, c) > r$. Let $c' = c|\pi(c)$.

For the same reasons as above, x is within distance r from c', whereas obviously, $d(y, c') \geq d(y_1, c) > r$. So $c' \in C'$ r-separates x and y.

Inequality (1) follows. $\square$

Theorem 2 *Let $n \geq 3, p \geq 0$ be such that $2p + 2 < n$, let $C \subseteq \mathcal{E}^n$ be a $(2p + 1)$-discriminating code and let $C' \subseteq F^{n-1}$ be any code obtained by the deletion of one coordinate in C. Then C' is $(2p + 1)$-identifying in F^{n-1}. Therefore,*

$$M_{2p+1}(n - 1) \leq D_{2p+1}(n). \tag{2}$$

Proof. Let $r = 2p + 1$. Let $C \subseteq \mathcal{E}^n$ be an r-discriminating code and $C' \subseteq F^{n-1}$ be the code obtained by deleting, say, the last coordinate in C. We shall prove that (a) any vector $x \in F^{n-1}$ is r-covered by at least one codeword of C'; (b) given any two distinct vectors $x, y \in F^{n-1}$, there is at least one codeword in C' which r-separates them.

(a) The vector $x|(\pi(x) + 1) \in F^n$ is odd. As such, it is r-covered by a codeword $c = c'|u \in C \subseteq \mathcal{E}^n$: $c' \in C'$, $u = \pi(c')$, and $d(x|(\pi(x) + 1), c) \leq r$. This proves that x is within distance r from a codeword of C'.

(b) Both $x|(\pi(x) + 1)$ and $y|(\pi(y) + 1)$ are odd vectors in F^n, and there is a codeword $c = c'|u \in C \subseteq \mathcal{E}^n$, with $c' \in C'$, $u = \pi(c')$, which r-separates them: without loss of generality, $d(x|(\pi(x) + 1), c) \leq r$ whereas $d(y|(\pi(y) + 1), c)$, which is an odd integer, is at least $r + 2$. Then obviously, $d(x, c') \leq r$ and $d(y, c') \geq r + 1$, i.e., there is a codeword in C' which r-separates x and y.

Inequality (2) follows. $\square$

Corollary 3 *For all $n \geq 2$ and $p \geq 0$ such that $2p + 1 < n$, we have:*

$$D_{2p+1}(n + 1) = M_{2p+1}(n).$$

$\square$

3 Conclusion

We have shown the equivalence between discriminating and identifying codes; the latter being already well studied, this entails a few consequences on discriminating codes.

For example, the complexity of problems on discriminating codes is the same as that for identifying codes; in particular, it is known [9] that deciding whether a given code $C \subseteq F^n$ is r-identifying is co-NP-complete.

For yet another issue, constructions, we refer to, e.g., [1]–[3], [6], [9], [10] or [11]; visit also [12]. In the recent [7], tables for exact values or bounds on $M_1(n)$, $2 \leq n \leq 19$, and $M_2(n)$, $3 \leq n \leq 21$, are given.

Discriminating codes have not been thoroughly studied so far; let us simply mention [4] for a general introduction and [5] in the case of planar graphs.

References

[1] U. Blass, I. Honkala, S. Litsyn: On the size of identifying codes, *Lecture Notes in Computer Science*, No. 1719, pp. 142–147, Springer-Verlag, 1999.

[2] U. Blass, I. Honkala, S. Litsyn: On binary codes for identification, *Journal of Combinatorial Designs*, Vol. 8, pp. 151–156, 2000.

[3] U. Blass, I. Honkala, S. Litsyn: Bounds on identifying codes, *Discrete Mathematics*, Vol. 241, pp. 119–128, 2001.

[4] E. Charbit, I. Charon, G. Cohen, O. Hudry: Discriminating codes in bipartite graphs, *Electronic Notes in Discrete Mathematics*, Vol. 26, pp. 29–35, 2006.

[5] I. Charon, G. Cohen, O. Hudry, A. Lobstein: Discriminating codes in (bipartite) planar graphs, *European Journal of Combinatorics*, to appear.

[6] G. Exoo: Computational results on identifying t-codes, Preprint, 1999.

[7] G. Exoo, T. Laihonen, S. Ranto: Improved upper bounds on binary identifying codes, *IEEE Transactions on Information Theory*, to appear.

[8] J. Gimbel, B. D. Van Gorden, M. Nicolescu, C. Umstead, N. Vaiana: Location with dominating sets, *Congressus Numerantium*, Vol. 151, pp. 129–144, 2001.

[9] I. Honkala, A. Lobstein: On the complexity of the identification problem in Hamming spaces, *Acta Informatica*, Vol. 38, pp. 839–845, 2002.

[10] M.G. Karpovsky, K. Chakrabarty, L.B. Levitin: On a new class of codes for identifying vertices in graphs, *IEEE Transactions on Information Theory*, Vol. 44(2), pp. 599–611, 1998.

[11] S. Ranto: Identifying and locating-dominating codes in binary Hamming spaces, Ph. D Thesis, University of Turku, 95 pp., 2007.

[12] http://www.infres.enst.fr/ lobstein/bibLOCDOMetID.html.