Available online at www.sciencedirect.com
ScienceDirect
Discrete Applied Mathematics 156 (2008) 2250–2263
DISCRETE APPLIED MATHEMATICS
www.elsevier.com/locate/dam

New bounds on binary identifying codes[☆]

Geoffrey Exoo[a], Tero Laihonen[b], Sanna Ranto[b,*]

[a] *Department of Mathematics and Computer Science, Indiana State University, Terre Haute, IN 47809, USA*
[b] *Department of Mathematics and Turku Centre for Computer Science TUCS, University of Turku, FIN-20014 Turku, Finland*

Received 13 March 2007; received in revised form 13 September 2007; accepted 20 September 2007
Available online 20 February 2008

Abstract

The original motivation for identifying codes comes from fault diagnosis in multiprocessor systems. Currently, the subject forms a topic of its own with several possible applications, for example, to sensor networks.

In this paper, we concentrate on identification in binary Hamming spaces. We give a new lower bound on the cardinality of r-identifying codes when $r \geq 2$. Moreover, by a computational method, we show that $M_1(6) = 19$. It is also shown, using a non-constructive approach, that there exist asymptotically good $(r, \leq \ell)$-identifying codes for fixed $\ell \geq 2$. In order to construct $(r, \leq \ell)$-identifying codes, we prove that a direct sum of r codes that are $(1, \leq \ell)$-identifying is an $(r, \leq \ell)$-identifying code for $\ell \geq 2$.

Keywords: Identifying code; Hamming space; Lower bound; Asymptotic behaviour; Direct sum

1. Introduction

Let $\mathbb{F} = \{0, 1\}$ be the binary field and denote by $\mathbb{F}^n$ the n-fold Cartesian product of it, i.e. the Hamming space. We denote by $A \triangle B$ the *symmetric difference* $(A \setminus B) \cup (B \setminus A)$ of two sets A and B. The *(Hamming) distance* $d(\mathbf{x}, \mathbf{y})$ between words $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$ is the number of coordinate places in which they differ. We say that $\mathbf{x}$ r-*covers* (or *covers*) $\mathbf{y}$ if $d(\mathbf{x}, \mathbf{y}) \leq r$. The *(Hamming) ball* of radius r centered at $\mathbf{x} \in \mathbb{F}^n$ is

$$B_r(\mathbf{x}) = \{\mathbf{y} \in \mathbb{F}^n \mid d(\mathbf{x}, \mathbf{y}) \leq r\}$$

and its cardinality is denoted by $V(n, r)$. For $X \subseteq \mathbb{F}^n$, denote

$$B_r(X) = \bigcup_{\mathbf{x} \in X} B_r(\mathbf{x}).$$

We also use the notation

$$S_r(\mathbf{x}) = \{\mathbf{y} \in \mathbb{F}^n \mid d(\mathbf{x}, \mathbf{y}) = r\}.$$

[☆] Some of the results of this paper have been presented at the International Workshop on Coding and Cryptography, WCC 2007.
* Corresponding author. Fax: +358 23336595.
E-mail addresses: ge@ginger.indstate.edu (G. Exoo), terolai@utu.fi (T. Laihonen), samano@utu.fi (S. Ranto).

0166-218X/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.dam.2007.09.017

Let C be a code of length n (i.e., a non-empty subset of $\mathbb{F}^n$) and $X \subseteq \mathbb{F}^n$. An *I-set* of the set X (with respect to the code C) is

$$I_r(C; X) = I_r(X) = B_r(X) \cap C.$$

We write for short $I_r(C; \{\mathbf{x}_1, \ldots, \mathbf{x}_k\}) = I_r(C; \mathbf{x}_1, \ldots, \mathbf{x}_k) = I_r(\mathbf{x}_1, \ldots, \mathbf{x}_k)$. If $r = 1$, we omit it from the notation whenever convenient.

Definition 1. Let r and ℓ be non-negative integers. A code $C \subseteq \mathbb{F}^n$ is said to be $(r, \leq \ell)$-*identifying* if for all $X, Y \subseteq \mathbb{F}^n$ such that $|X| \leq \ell$, $|Y| \leq \ell$ and $X \neq Y$ we have

$$I_r(C; X) \neq I_r(C; Y).$$

If $\ell = 1$, we say, for short, that C is r-*identifying*.

Note that a code $C \subseteq \mathbb{F}^n$ is $(r, \leq \ell)$-identifying if and only if

$$I_r(C; X) \triangle I_r(C; Y) \neq \emptyset \tag{1}$$

for any subsets $X, Y \subseteq \mathbb{F}^n$, $X \neq Y$ and $|X| \leq \ell$ and $|Y| \leq \ell$.

A set $X \subseteq \mathbb{F}^n$ that we try to identify (knowing only the set $I_r(X)$) is called a *fault pattern*. Clearly, $I_r(C; \emptyset) = \emptyset$ for any code C, and if C is $(r \leq \ell)$-identifying, then $I_r(C; X) = \emptyset$ implies that there is unique such a set X, namely $X = \emptyset$.

The seminal paper [10] by Karpovsky, Chakrabarty and Levitin initiated research in identifying codes, and it is nowadays a topic of its own; for various papers dealing with identification, see [14]. Originally, identifying codes were designed for finding malfunctioning processors in multiprocessor systems (such as binary hypercubes, i.e., binary Hamming spaces); in this application we want to determine the set of malfunctioning processors X (the fault pattern) of size at most ℓ when the only information available is the set $I_r(C; X)$ provided by the code C. A natural goal there is to use identifying codes which are as small as possible. The theory of identification can also be applied to sensor networks, see [16]. Small identifying codes are needed for energy conservation in [11]. For other applications like environmental monitoring, we refer to [12] and the references therein.

The smallest possible cardinality of an $(r, \leq \ell)$-identifying code of length n is denoted by $M_r^{(\leq \ell)}(n)$ (whenever such a code exists). If $\ell = 1$, we denote $M_r^{(\leq 1)}(n) = M_r(n)$. Moreover, if $r = 1$, we denote $M_1(n) = M(n)$.

This paper is organized as follows. In Section 2 we improve on the known lower bounds on the cardinalities of r-identifying codes by combining a counting argument with partial constructions. On the other hand, by computational methods, we are able to show that $M_1(6) = 19$; thus closing the gap of $18 \leq M_1(6) \leq 19$ in [2]. New 1- and 2-identifying codes are given as well. An averaging method of Section 3 guarantees that good $(r, \leq \ell)$-identifying codes exist. Since the approach is non-constructive, we focus in the last section on constructing $(r, \leq \ell)$-identifying codes for $r \geq 2$ and $\ell \geq 2$. Although $(r, \leq \ell)$-identifying codes are studied in natural grids, see for instance [6,7], in $\mathbb{F}^n$ the problem has not been addressed before when $r \geq 2$ and $\ell \geq 2$.

2. On *r*-identifying codes

2.1. A lower bound

The following theorem improves the lower bound from [10, Theorem 1 (iii) and Theorem 2] for $r \geq 2$.

Theorem 2. *Let $C \subseteq \mathbb{F}^n$ be r-identifying and $m = \max\{|I_r(\mathbf{x})| : \mathbf{x} \in \mathbb{F}^n\}$. Denote*

$$f_r(x) = \frac{(x-2)\left(\binom{2r}{r} - 1\right)}{\binom{2r}{r} + \binom{x}{2} - 1}.$$

We have

$$|C| \geq \frac{2^n(2 + f_r(v))}{V(n, r) + f_r(v) + 1}$$

where $v = m$, if $m \geq 2 + 2\binom{2r}{r}$, and $v = 3$ otherwise.