Discrete Mathematics 340 (2017) 713–721

Contents lists available at ScienceDirect

Discrete Mathematics

journal homepage: www.elsevier.com/locate/disc

On traceability property of equidistant codes

Anu Kathuria [a], S.K. Arora [b], Sudhir Batra [c,*]

[a] The Technological Institute of Textiles and Sciences, Bhiwani-127021, India
[b] Department of Mathematics, M.D. University, Rohtak-124001, India
[c] Department of Mathematics, DCR University of Science and Technology Murthal, (Sonepat)-131039, India

ARTICLE INFO

Article history:
Received 27 October 2015
Received in revised form 10 October 2016
Accepted 13 December 2016
Available online 6 January 2017

Keywords:
Maximum Distance Separable (MDS) code
Equidistant code
Traceable (TA) code
Jabobsthal matrix
Hadamard matrix
RS code

ABSTRACT

Necessary and Sufficient conditions for an equidistant code to be a 2-TA code are obtained. An explicit construction method is proposed to obtain linear MDS $[p+1, 2, p]$ codes over the finite field F_p, where p is a prime. These codes can be used as 2-TA codes for $p > 2$. In particular, for $p = 3$, it is observed that the linear [4, 2, 3] MDS code contradicts a result of Jin and Blaum (2007). The correct version of this result and its proof is given. Existence of some infinite families of equidistant 2-TA codes is shown by using Jacobsthal and Hadamard matrices. Some of these codes are also observed to be good equidistant code (Sinha et al., 2008).

1. Introduction

Chor, Fiat and Naor [3] introduced the concept of traitor tracing as a means to limit piracy. Traitor tracing schemes may prove quite useful in protecting copyrighted digital data. When a pirated copy created by a group of authorized users of the copyrighted data is traced, traitor tracing schemes allow to trace it back to at least one producer of it. In recent years several codes have been studied for the purpose of their usefulness in traceability schemes. In general, these codes are called fingerprinting codes. The weak forms of these codes, called frameproof codes were introduced by Boneh and Shaw [2]. Strong form of codes, called identifiable parent property (IPP) codes have been introduced by Hollman, Van Lint, Linnartz and Tolhuizen [6]. Other form of codes, called traceability codes were introduced by Chor, Fiat and Naor in [3]. TA codes are stronger than IPP codes and is a subclass of IPP codes and generally have efficient traitor tracer algorithms. IPP codes on the other hand are capable of identifying traitors requiring less restrictive conditions than TA codes at the expense of not having efficient traitor tracing algorithm. Combinatorial properties of traceability schemes and frameproof codes have been studied by Staddon, Stinson and Wei [10]. Sufficient conditions for an equidistant code to be an IPP code have been derived in [6]. Some constructions of equidistant frameproof codes have been suggested in [5]. In [7], Jin and Blaum have obtained the necessary and sufficient conditions for a linear MDS $[n, k, d]$ code over F_q with $n \leq q+1$ to be a c-TA code. In the present paper, we obtain necessary and sufficient conditions under which an equidistant code can be used as a 2-TA code (see Theorems 3.4 and 3.5). We also derive a condition under which an equidistant code cannot be a m-TA code, where $m \geq 3$ (see Theorem 3.6).

In Section 4, we propose a method to construct linear MDS $[p+1, 2, p]$ codes over the finite field F_p, where p is an odd prime. These codes are extended RS codes and can be used as 2-TA codes. In particular, for $p = 3$, the ternary Hamming code

[*] Corresponding author.
E-mail addresses: anu_sept24@rediffmail.com (A. Kathuria), batrasudhir@rediffmail.com (S. Batra).

http://dx.doi.org/10.1016/j.disc.2016.12.009
0012-365X/© 2016 Elsevier B.V. All rights reserved.

[4, 2, 3] turns out to be a 2-TA code, contradicting Theorem 2.2 of [7] (see Example 2 in Section 4). At the end of this section we provide the correct version of this theorem and its proof (see Theorem 4.11).

In Section 5, we show that Jacobsthal matrices [4,11] for primes of the form $4m + 3$ generate equidistant 2-TA codes of length n with distance $d < 2n/3$. Further, using the results of [1,9] the existence of two infinite classes of non-linear equidistant 2-TA codes of length n and distance d such that $2n/3 < d \leq 3n/4$ is shown. Here it is also observed that some of these are good equidistant codes.

The following section is devoted to some preliminaries required for the discussion in the subsequent sections.

2. Some preliminaries

Throughout this paper, the following basic definitions and terminology from [5–8,11] will be used and F_q denotes a finite field with q elements.

2.1. Here we recall some basic definitions related to error correcting codes.

(i) Let Q be a finite set of alphabets. Then a subset $C \subseteq Q^n$ is called a code of length n over Q. The elements of Q^n are called words and the elements of C are called codewords of length n.

(ii) Let a and b be two words, then the hamming distance $d(a, b)$ between a and b is the number of coordinates in which they differ and the number of non-zero coordinates of a word c is called the weight of c. The minimum distance d of C is $d = min\{d(a, b)|a, b \in C\}$.

(iii) $I(x, y) = \{i : x_i = y_i\}$ for $x = (x_1, x_2, \dots, x_n)$, $y = (y_1, y_2, \dots, y_n) \in Q^n$. Similarly, we can define $I(x, y, z, \dots)$ for any finite number of words $x, y, z, \dots$.

(iv) A subspace C of F_q^n is called a linear code over F_q. The dimension of the code is defined as the dimension of the subspace. A linear code with length n, dimension k and minimum distance d is denoted as $[n, k, d]$ code. A linear code $C[n, k, d]$ is an MDS code if $d = n - k + 1$.

(v) A code C with same distance between every pair of its codewords is called an equidistant code. If all the codewords of a code C carry same weight then code C is called constant weight code. A code C with both of these properties is known as equidistant constant weight code.

(vi) Let $\alpha = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$, where the α_i are distinct elements of F_q, and let $v = \{v_1, v_2, \dots, v_n\}$, where the v_i are non-zero (but not necessarily distinct) elements of F_q. Then the Reed–Solomon(RS) code of length n and dimension k, consists of all vectors $(v_1 f(\alpha_1), v_2 f(\alpha_2), \dots, v_n f(\alpha_n))$, where $f(x)$ ranges over all polynomials of degree less than k with coefficients from F_q.

Choose the particular polynomial basis $1, x, x^2, \dots, x^{k-1}$. In this basis the generator matrix of the above defined code is given by

$$\begin{bmatrix} v_1 & v_2 & \cdots & v_n \\ v_1\alpha_1 & v_2\alpha_2 & \cdots & v_n\alpha_n \\ \vdots & \vdots & \cdots & \vdots \\ v_1\alpha_1^{k-1} & v_2\alpha_2^{k-1} & \cdots & v_n\alpha_n^{k-1} \end{bmatrix}.$$

A code whose generator matrix is given by adding a new column $(0, 0, \dots, 1)^T$ to the above matrix is called extended RS code of length $n + 1$.

2.2. Now let us define some terms related to fingerprinting codes.

(i) Detectable and Undetectable Positions. Let $X \subseteq Q^n$. Then we say that the position $i \in \{1, 2, \dots, n\}$ is undetectable for X if ith position of each word $x \in X$ is occupied with the same alphabet, otherwise the position i is detectable.

(ii) Coalition. It means two or more users meet for the purpose of creating an illegal copy of a digital object (see marking assumption (iv) also) by comparing their copies. A member of the coalition is called a pirate.

(iii) Descendant Set. For any two words $a = (a_1, a_2, \dots, a_n)$ and $b = (b_1, b_2, \dots, b_n)$ in Q^n, the set of descendants $Desc(a, b) = D(a, b)$ is defined as $D(a, b) = \{x \in Q^n|x_i \in \{a_i, b_i\}, i = 1, 2, \dots, n\}$.
The above definition of Descendant set can be naturally extended to any finite number of words $a, b, c, \dots$.

(iv) Marking Assumption. In the static form of finger printing scheme, each digital content is divided into multiple segments, among which n segments are chosen for marking them with symbols which correspond to alphabets in Q. Each user receives a copy of the content with differently marked symbols. If a code C over Q of length n is used to assign the symbols for each segment to each user. Then each copy can be denoted as a codeword of C and each coordinate x_i of a codeword $(x_1, x_2, \dots, x_n)$ can be termed as a symbol. Further, assume that any coalition of c users is only capable of creating a pirated copy whose marked symbols correspond to a word of Q^n that lie in the Descendant set of these c users.

(v) Identifiable Parent Property Code. A code C is c-IPP if for any $x \in Desc(C)$ it holds that if $C_i \subseteq C$ with $|C_i| \leq c$, then $\cap_{\{i:x \in Desc(C_i)\}} C_i \neq \phi$.