國 立 交 通 大 學

應用數學系

碩 士 論 文

智財碼和非覆集合族的關連探討

Codes and Cover-Free Families for Copyright Protection

研 究 生：汪政緯

指導教授：翁志文　教授

中 華 民 國 九 十 七 年 六 月

智財碼和非覆集合族的關連探討

Codes and Cover-Free Families for Copyright Protection

研 究 生 ： 汪政緯　　　Student : Cheng-Wei Wang

指導教授 ： 翁志文　　　Advisor : Chih-Wen Weng

國 立 交 通 大 學

應 用 數 學 系

碩 士 論 文

A Thesis
Submitted to Department of Applied Mathematics
College of Science National Chiao Tung University
In partial Fulfillment of Requirement
For the Degree of Master In Applied Mathematics

June 2008

Hsinchu, Taiwan, Republic of China

中華民國九十七年六月

智財碼與非覆蓋集合族的關連探討

研究生 : 汪政緯　　　　　　　　　　　　指導教授 : 翁志文

國立交通大學應用數學系

摘要

TA 碼，IPP 碼， SFP 碼和 FP 碼的應用在數位資料的保護上有著重要的價值，目的在預防未授權產品的非法拷貝。在此論文中，我們造了些上述碼，並研究碼的基本性質和探討碼與 cover-free family 的關係。根據 cover-free family 的定義，我們構造了些新的關係矩陣，並証明上述矩陣為 disjunct matrix。用布林代數的語言，即我們允許某種程度上的容錯率。文末我們蒐集了前人關於 SFP 碼及 IPP 碼簡單且重要的構造法。

Codes and Cover-free Families for Copyright Protection

Student : Cheng-Wei Wang Advisor : Chih-Wen Weng

Department of Applied Mathematics
National Chiao Tung University

Abstract

The applications of TA codes, IPP codes, SFP codes and FP codes play an important role in the protection of digital data. The destination of these codes is to prevent an unauthorized copy. Some new and old examples of these codes are given. This thesis studies basic properties of the above codes and the relationships between theses codes and cover-free families. Therefore, we construct some new incidence matrices and prove these matrices are disjunct matrices. According to our constructions, in the language of pooling design, the construction allows some test errors. In the end, we collect some simple and important constructions of SFP codes and IPP codes.

誌謝

此論文的完成，要感謝許多人，尤其是翁志文老師，在老師的指導下，使我獲益良多，明白做研究必須一步一腳印，小心仔細地驗證。也感謝所上所有關心我的師長與同學。黃大原老師，傅恆霖老師，陳秋媛老師和喻培學長... 等，在課業上不斷地給予我指導。明耀學長，元勳學長，雅榕，若宇，偉慈，鎬文，敏筠，威雄，佩純...等，更豐富了我的研究所生活。最後感謝家父，家母及舍妹，一直在背後支持與鼓勵我，使我能順利完成學業，真的謝謝大家。

Contents

1 Introduction

To protect an electronic product, such as digital data, a distributor marks each copy with some codeword and then ships each user his data ”marked” with that codeword. This marking, a ”digital fingerprint”, permits the distributor to detect any unauthorized copy and trace it back to the user that created it. This will prevent users from releasing an unauthorized copy. A coalition of users, yet, may detect some of the marks where their copies differ. They can then change these marks arbitrarily. Boneh and Shaw (1995)[2] defined ”w-frame proof codes” as preventing users from ”framing” another user. A w-frame proof code possesses the property that no coalition of at most w can frame another registered user. In Stinson and Wei (1998)[15], combinatorial methods are used to further probe frame proof codes. Several constructions of w-frame proof codes are given in Boneh and Shaw (1995)[2], Chee (1996)[4] and Stinson and Wei (1998)[15].

In Chapter 2, we introduce five classes of codes w-TA codes, w-IPP codes, w-SFP codes, and w-FP codes from the most to the least restrictive. By above codes, we define the registered user, unregistered user, and guilty user in order to apply to copyright protection. We provide examples and counter examples for theses definition originally introduced by D.R. Stinson, Tran van Trung and R. Wei (2000)[13]. Define $\mathrm{desc}^{-1}(x)$ consisting of all the coalitions of size at most w that could framed x and suppose x is an unregistered user in 2-SFP code C ($x \notin C$). Since $\mathrm{desc}^{-1}(x)$ consists of a collection of 2-subsets of C, we can view it as the set of edges of a graph on vertex set C. That is, we can give the link from a 2-SFP code to a star graph (i.e. there exists a vertex that is incident to every edges) and K_3(the complete graph on three vertices).

In Chapter 3, we first introduce the set system $(P, \mathcal{B})$ and the $(w; \alpha)$-cover-free family. Lemma 3.2 give relationships between a cover-free family and a w-FP code. By above lemma, we generalize a w-FP code to a $(w; \alpha)$-FP code in our new Definition 3.3. Finally, we analyse minimum distance d and α of a $(w; \alpha)$-FP code and reprove Corollary 3.6.(Staddon, Stinson and Wei, 2001)[14].

In Chapter 4 and 5, in our language, we generalize a $(w; \alpha)$-cover-free family to an $(\ell, s; e)$-cover-free family in Definition 4.1. Our treatment simplifies the original definition of an (ℓ, s)-sandwich-free family in [13]. Theorem 4.2 which connects a w-SFP code with a cover-free family is similiar to lemma 3.2. We research the properties relating to w-SFP codes. In Theorem 4.5, we construct some new incidence matrices and prove these matrices are disjunct matrices. Recalling the definition of a $(w;\alpha)$-FP code, we construct a $(w;\alpha)$-CFF in Theorem 5.1 by means of the disjunct matrix. This tells us, in the language of pooling design, the construction allows some test errors.

In Chapter 6 and 7, we collect and introduce some simple constructions of SFP and IPP codes. In Chapter 7, let C_1 and C_2 be two different codes with the same length. Bush (1952)[3] proved the existence of combination of C_1 and C_2 in Theorem 7.6. Further, Tran and Sosina (2004) [16] constructed a similiar one, but more general with distinct length in Theorem 7.4. Based on above two theorems, Tran and Sosina (2005)[17] used concatenation technique to construct a new w-IPP code with the same parameter q_2 in Theorem 7.14.

2 Codes for copyright protection

Definition 2.1. Let Q denote a set of q elements. A subset $C \subseteq Q^n$ is called *a code of length n over Q*. The elements in C are called *codewords*. The number of codewords in C is called the *size* of C. C is called an (n, N, q)-*code* over Q if $|C| = N$ and Q is the set of *alphabets*. An $(n, N, 2)$-code is called an (n, N)-code for short.

To reveal the application for codes to copyright protection, an element in Q^n is also called a *user*, in C is a *registered user*, and in $Q^n - C$ is an *unregistered user*, or an *illegal copy*.

Definition 2.2. Let C denote an (n, N, q)-code over Q. For $X \subseteq C$, the set of *descendants* of X is the subset

$$\text{desc}(X) := X_1 \times X_2 \times \cdots \times X_n$$

of Q^n, where $X_i := \{c_i \mid c \in X\}$ is the set of alphabets used in the ith coordinate of X.

An element in $desc(X)$ is referred to as a *user framed by the coalition* X. For $x \in desc(X)$, X is called the *set of parents* of x. The set $X \subseteq C$ is intercepted as a family of registered users and $x \in \text{desc}(X) - C$ is an illegal copy produced by X.

It is clear that $C \subseteq \text{desc}(C)$.

We see an example before going to our new definition.

Example 2.3. Set $Q = \{0, 1\}$, and

$$C = \{(0, 0, 0), (1, 0, 0)(0, 1, 0)(0, 0, 1)\} \subseteq Q^3.$$

Then C is an (3,4,2)-code. Observe $\text{desc}(C) = Q^3$.

Throughout the remaining of the section, C is an (n, N, q)-code over $Q := \{1, 2, \ldots, q\}$ and $w \leq N$ is a positive integer.

Definition 2.4. For $x, y \in Q^n$, define the Hamming distance $\partial(x, y)$ to be the number of different positions in x, y. That is

$$\partial(x, y) := |\{i \mid x_i \neq y_i\}|$$

for $x, y \in Q^n$. An $(n, N, q; d)$-code C is an (n, N, q)-code with

$$d = \min\{\partial(x, y) \mid x, y \in C, x \neq y\}.$$

Now we are ready to introduce the first class of codes.

Definition 2.5. C is a *w-traceability code* (*w-TA code*) whenever for any $X \subseteq C$ with $|X| \leq w$ and for any $x \in \mathrm{desc}(X)$,

$$\partial(x, X) < \partial(x, C - X), \tag{2.1}$$

where $\partial(x, X) := \min\{\partial(x, y) \mid y \in X\}$.

Note that every code is 1-TA code. In a w-TA code, $\mathrm{desc}(X) \cap C = X$ for any $X \subseteq C$ with $|X| \leq w$.

A code is w-TA if, for any n-tuple x framed by a set X of w parents, the nearest codeword to the x is taken from the set of parents. In particular, the register users with minimum Hamming distance to x are all in X. Hence we can trace some register users in X from an illegal copy x. Hence TA codes are designed to be used in schemes that protect copyrighted digital data against piracy.

Example 2.6. Set

$$C = \{(1, 1, \ldots, 1, i) \mid i \in Q\} \subseteq Q^n,$$

observe $\mathrm{desc}(X) = X$ for any $X \subseteq C$. Then C is a q-TA code.

The following property of w-TA codes will give link to our next definition.

Lemma 2.7. *Suppose C is a w-TA code. Then for any $X, Y \subseteq C$ with $|X|, |Y| \leq w$ and for any $x \in \mathrm{desc}(X) \cap \mathrm{desc}(Y)$,*

$$\{y \in X \mid \partial(y, x) = \partial(X, x)\} \subseteq Y. \tag{2.2}$$

Proof. Assume that there exists $y \in X$ with $\partial(y, x) = \partial(X, x)$ and there exists $Y \subseteq C$ with $|Y| \le w$, $x \in \mathrm{desc}(X) \cap \mathrm{desc}(Y)$ and $y \notin Y$. Then

$$\begin{aligned} \partial(x, y) &< \partial(x, C - X) \\ &< \partial(x, Y - X) \end{aligned}$$

and

$$\begin{aligned} \partial(x, y) &= \partial(x, X) \\ &\le \partial(x, Y \cap X). \end{aligned}$$

Hence $\partial(x, y) \le \partial(x, Y)$, a contradiction. □

Now we give the second class of codes.

Definition 2.8. C is a *w-identifiable parent property code* (*w-IPP code*) whenever for all $x \in \mathrm{desc}(C)$,

$$\bigcap Y \neq \emptyset, \tag{2.3}$$

where $Y \in \mathrm{desc}^{-1}(x)$.

An registered user $y \in \cap Y$ in (2.3) is called a *guilty user* for x. An w-IPP code is also called a code with traceability. If there is no $Y \subseteq C$ with $|Y| \le w$ and $x \in \mathrm{desc}(Y)$ in the above definition then in convention we realize $\bigcap Y$ as Q^n.

A code is w-IPP if for all $x \in \mathrm{desc}(C)$, then there exists a quilty user for x. Hence IPP codes are introduced to provide protection against illegal producing of copyrighted digital material.

Observe that if $x \in C$ then the set in (2.3) is $\{x\}$ since we can choose one of the Y to be $\{x\}$. By Lemma 2.7, we have

Corollary 2.9. *A w-TA code is a w-IPP code.* $\square$

We see two examples.

Example 2.10. Set

$$C = \{1212, 2121, 4343, 3434, 1144\}.$$

It is easy to see that C is a 2-IPP $(4, 5, 4)$-code. If we set

$$X = \{1212, 2121\} \subseteq C,$$

$x = 1111 \in \mathrm{desc}(X)$, then $d(x, X) = 2 \not< 2 = \mathrm{desc}(x, C - X)$. Hence C is not a 2-TA code.

Example 2.11. Set

$$C := \{(i, i, \ldots, i) \mid i \in Q\} \subseteq Q^n.$$

Then C is an (n, q, q)-code. Observe $\mathrm{desc}(C) = Q^n$, and for any $x = (x_1, x_2, \ldots, x_n) \in Q^n$,

$$\begin{aligned} \bigcap Y &= \{(i, i, \ldots, i) \mid i \in \{x_1, x_2, \ldots, x_n\}\} \\ &\neq \emptyset. \end{aligned}$$

where the intersection is taking for all $Y \subseteq C$ and $x \in \mathrm{desc}(Y)$. Hence C is a w-IPP code for any w.

Now we are ready to give the 3th and 4th class of codes.

Definition 2.12. C is a *w-secure frame proof code* (*w-SFP code*) whenever for any $X, Y \subseteq C$ with $|X|, |Y| \leq w$,

$$\mathrm{desc}(X) \cap \mathrm{desc}(Y) \neq \emptyset \implies X \cap Y \neq \emptyset.$$

Note that $\text{desc}(X) \cap \text{desc}(Y) = \phi$ iff $X_i \cap Y_i = \phi$ for some i.

A code is w-SFP if no two disjoint coalitions of size at most w can frame a common user.

Definition 2.13. Suppose that C is a (n, N)-code and for any $x \in \{0,1\}^n$, define

$$\text{desc}^{-1}(x) = \{X \subseteq C \mid |X| \leq w \text{ and } x \in \text{desc}(X)\}.$$

Evidently, $\text{desc}^{-1}(x)$ consists of all the coalitions of size at most w that could have framed x.

A w-SFP (n, N)-code does not permit traceability, but it does afford some security, as follows:

(i) It is impossible for a coalition C_1 of size at most w to implicate a disjoint coalition C_2 of size at most w by constructing an unregistered user $x \in \text{desc}(C_1)$.

(ii) If x is an unregistered user that has been constructed by a coalition of size at most w, then any $X \in \text{desc}^{-1}(x)$ contains at least one guilty user.

From (2.3) we have

Corollary 2.14. *A w-IPP code is w-SFP code.* □

Example 2.15. Set

$$C = \{(1,0,1), (1,1,0), (0,1,1)\}.$$

Then C is a 2-SFP code over $\{0,1\}$. Note that C is not a 2-IPP code because for

$$Y = \{(1,0,1),(1,1,0)\}$$

,

$$Z = \{(1,1,0),(0,1,1)\},$$

and

$$W = \{(1,0,1),(0,1,1)\},$$

we have $(1,1,1) \in \mathrm{desc}(Y) \cap \mathrm{desc}(Z) \cap \mathrm{desc}(W)$ and $Y \cap Z \cap W = \emptyset$.

Definition 2.16. C is a *w-frame proof code* (*w-FP code*) whenever for any $X \subseteq C$ with $|X| \leq w$, we have

$$\mathrm{desc}(X) \cap C = X.$$

A code is w-FP if no coalition of size at most w can frame another registered user.

FP codes were introduced by Boneh and Shaw[2] as a method of " digital fingerprinting" which prevents a coalition of a special size w from framing a user not in the coalition. Stinson and Wei [15] then gave a combinatorial formulation of the problem in terms of certain types of extremal set systems. We study FP codes that provide a certain (weak) form of traceability.

Lemma 2.17. *A w-SFP code is w-FP code.*

Proof. $X \subseteq \mathrm{desc}(X) \cap C$ is clear. Suppose $y \in (\mathrm{desc}(X) \cap C) - X$. Then by setting $Y = \{y\}$ in Definition 2.12 we find $X \cap \{y\} = \emptyset$, a contradiction. $\square$

We see an example.

Example 2.18. Set

$$C = \{111, 123, 132, 222, 213, 231, 333, 312, 321\}.$$

It is easy to see that C is a 2-FP $(3, 9, 3)$-code. If we set

$$X = \{111, 123\}, Y = \{132, 321\},$$

then $X \cap Y = \phi$, but $\mathrm{desc}(X) \cap \mathrm{desc}(Y) = \{121\} \neq \phi$. Hence C is not a 2-SFP $(3, 9, 3)$-code.

Related questions, including generalizations of frame proof codes to the setting of public-key, cryptography, have been studied in Biehl and Meyer (1997) [1], Chor et al. (1994)[5], Pfitzmann (1996)[10], and Pfitzmann and Waidner (1997a,b) [11], [12].

Suppose that C is a w-FP (n, N)-code and $x \in \{0,1\}^n \setminus C$ (i.e., x is an unregistered user). If it happened that $|\mathrm{desc}^{-1}(x)| = 1$, say $\mathrm{desc}^{-1}(x) = \{X\}$, then we could conclude that X was the coalition that constructed x (assuming, of course, that all coalitions have size at most w). More generally, if $\mathrm{desc}^{-1}(x) \neq \emptyset$ and there exists a codeword $c^{(j)}$ such that $c^{(j)} \in X$ for all $X \in \mathrm{desc}^{-1}(x)$, then we would at least be able to identify user j as being guilty. Unfortunately, as shown in Boneh and Shaw (1995)[2], this is hoping for too much. The following theorem is a simple generalization of (Boneh and Shaw, 1995 [2], Theorem 11), which concerned the case $w = 2$.

A w-FP (n,N)-code is not necessary to permit traceability. D.R. Stinson, Tran van Trung and R. Wei (2000) [13] claimed why in following.

Theorem 2.19. *(D.R. Stinson , Tran van Trung and R. Wei, 2000)[13]. Suppose C is a w-FP (n,N)-code with $N \geq 2w-1$. Suppose $D \subseteq C$ with $|D| = 2w-1$. Let* $\mathrm{maj}(D) \in \{0,1\}^n$ *be defined as*

$$\mathbf{maj}(D)_i = \begin{cases} 1, & \textit{if } |\{\ c \in D \mid c_i = 1\ \}| \geq w, \\ 0, & \textit{if } |\{\ c \in D \mid c_i = 0\ \}| \geq w. \end{cases}$$

Then maj(D) is an unregistered user and $\mathbf{maj}(D) \in \mathrm{desc}(X)$ *for all $X \subseteq D$ with $|X| = w$. That is, C does not permit traceability.*

Proof. It is easy to see that $\mathbf{maj}(D) \in \mathrm{desc}(X)$ for all $X \subseteq D$ with $|X| = w$. It remains to show that $\mathbf{maj}(D)$ is an unregistered user. Suppose not; then $\mathbf{maj}(D) = c^{(u)}$ for some u. Let

$$X \subseteq D \setminus \{c^{(u)}\} \text{ with } |X| = w.$$

Then $c^{(u)} \in \mathrm{desc}(X) \cap C = X$, which contradicts the fact that C is a w-FP code. □

The above theorem says that we cannot be guaranteed of identifying a guilty user in a w-FP (n,N)-code. For, if $x = \mathbf{maj}(D)$ for some D where $|D| = 2w-1$, then

$$\bigcap_{X \in \mathrm{desc}^{-1}(x)} X = \emptyset.$$

Corollary 2.20. *Any w-IPP (n,N)-codes have $N < 2w-1$.* □

We now consider 2-SFP (n, N)-code in more detail. Suppose that C is a 2-SFP (n, N)-code, suppose that x is an unregistered user, and suppose that $X \in \mathrm{desc}^{-1}(x)$ with $|X| \leq 2$. Since x is an unregistered user, $|X| \neq 1$. Therefore, $|X| = 2$.

Since $\mathrm{desc}^{-1}(x)$ consists of a collection of 2-subsets of C, we can view it as the set of edges of a graph on vertex set C. Since C is a 2-SFP code, it must be the case that any two distinct edges in $\mathrm{desc}^{-1}(x)$ are incident. From this it is easily seen that one of two possibilities must occur:

(i) $\mathrm{desc}^{-1}(x)$ is a *star graph* (i.e., there exists a vertex that is incident to every edge of $\mathrm{desc}^{-1}(x)$).

(ii) $\mathrm{desc}^{-1}(x)$ is isomorphic to K_3 (the complete graph on three vertices).

As a consequence of this characterization of $\mathrm{desc}^{-1}(x)$ in the case $w = 2$, we obtain the following result.

Theorem 2.21. *(D.R. Stinson, Tran van Trung and R. Wei, 2000)[13]. Suppose that C is a 2-SFP (n, N)-code and suppose that x is an unregistered user that is produced by a coalition of size at most two. Then one of the following two possibilities must occur:*

(i) at least one guilty user can be identified; or

(ii) a set of three user can be identified, two of which must be guilty.

$\square$

Since its inception in the early 1980's, the field of copyright and distribution rights protection of multimedia documents has become an essential concern to companies that distribute digital documents. This is the case of Networked University for e-Learning. Independently of the use of the documents and the type of organization (public or private) the authors of educational documents have to be protected against dishonest users. The possibility of making copies of these documents without a quality degradation constitutes a severe threat to authors rights.

The security mechanism in this environment must be more strict than in the e-commerce market with physical goods delivered to the user using traditional networks. Cryptographic techniques are insufficient because the lack of confidence about the receiver behavior. The most acceptable techniques to solve this situation are watermarking and fingerprinting. Both techniques are based on embedding an imperceptible mark in the document. In the case of fingerprinting, analogously to the human fingerprint, the mark is unique for every legally distributed copy with the aim of discovering fraudulent redistributors.

3 Cover-Free Families

We first define some terminologies concerning set systems. A set system is a pair $(P, \mathcal{B})$ where P is a set of elements called *points*, and $\mathcal{B}$ is a set consisting of subsets of P, the members of $\mathcal{B}$ which are called *blocks*.

Let $(P,\mathcal{B})$ be a set system with $|\mathcal{B}|=N$. Fix $w\le N$.

Definition 3.1. A set system $(P,\mathcal{B})$ is a $(w;\alpha)$-*cover-free family* ($(w;\alpha)$-*CFF*) whenever for any $\mathcal{X}\subseteq\mathcal{B}$ with $|\mathcal{X}|\le w$ and any $A\in\mathcal{B}-\mathcal{X}$,

$$|A-\bigcup_{X\in\mathcal{X}}X|\ge\alpha+1.$$

We refer a $(w;0)$-CFF to w-*CFF* for short. $(P,\mathcal{B})$ is k-*uniform* whenever $|B|=k$ for any $B\in\mathcal{B}$.

Let C denote an (n,N,q)-code over Q. For each $c\in C$, set

$$B_c:=\{(i,c_i)\mid 1\le i\le n\}\subseteq[n]\times Q.$$

Then $([n]\times Q,\{B_c\}_{c\in C})$ is an n-uniform family. Observe for any $x,y\in C$,

$$B_x=B_y\quad\text{iff}\quad x=y,$$

and for $X\subseteq C$, $x\in Q^n$, we have

$$B_x\subseteq\bigcup_{c\in X}B_c\quad\text{iff}\quad x\in\operatorname{desc}(X).$$

Then we immediately have

Lemma 3.2. *Let C be an (n,N,q)-code over Q. Then the set system $([n]\times Q,\{B_c\}_{c\in C})$ is a w-CFF if and only if C is a w-FP code.*

Proof. ($\Longrightarrow$) Suppose a set system $([n]\times Q,\{B_c\}_{c\in C})$is a w-CFF. Fix $X\subseteq C$ with $|X|\le w$, and given any codeword $x\in\operatorname{desc}(X)\cap C$. Hence

$$B_x\subseteq\bigcup_{c\in X}B_c$$

and $x \in C$. Since $([n] \times Q, \{B_c\}_{c\in C})$ is a w-CFF, we know $x \in X$.

($\Longleftarrow$) Suppose C is a w-FP code. Given any $X \subseteq C$ with $|X| \leq w$, and pick any $y \in C - X$. Since C is a w-FP code, we know $\mathrm{desc}(X) \cap C = X$. Thus

$$y \notin \mathrm{desc}(X) \text{ implies } B_y \not\subseteq \bigcup_{x\in X} B_x.$$

Hence $|B_y - \bigcup_{x\in X} B_x| \geqslant 1$ □

It is natural to generalize the definition of a w-FP code to

Definition 3.3. An (n, N, q)-code C is a $(w;\alpha)$-*frame proof code* ($(w;\alpha)$-*FP code*) whenever $([n] \times Q, \{B_c\}_{c\in C})$ is a $(w;\alpha)$-CFF.

Hence a $(w;0)$-FP code is a w-FP code.

Proposition 3.4. *Suppose C is an $(n, N, q; d)$-code , where $d > n(1 - \frac{1}{w^2})$. Then C is a $(w;\alpha)$-FP code where*

$$\alpha = \left\lfloor\ n(1 - \tfrac{1}{w})\ \right\rfloor.$$

Proof. Fix $\mathcal{X} \subseteq \{B_c\}_{c\in C}$ with $|\mathcal{X}| \leq w$ and $B \in \{B_c\}_{c\in C} - \mathcal{X}$. Observe $|B \cap B'| \leq n - d$ for any $B' \in \mathcal{X}$. Hence

$$\begin{aligned} |B - \bigcup_{B'\in\mathcal{X}} B'| &\geq n - w(n-d) \\ &> n(1 - \frac{1}{w}). \end{aligned}$$

Since $|B - \bigcup_{B'\in\mathcal{X}} B'|$ is an integer, we have

$$|B - \bigcup_{B'\in\mathcal{X}} B'| \geq \left\lfloor\ n(1 - \tfrac{1}{w})\ \right\rfloor + 1.$$

□

Proposition 3.5. *Suppose that an (n, N, q)-code C is a $(w; \alpha)$-FP code, where*

$$\alpha = \left\lfloor\, n(1 - \tfrac{1}{w})\, \right\rfloor.$$

Then C is a w-TA code.

Proof. Fix $X \subseteq C$ with $|X| \leq w$ and $x \in \mathrm{desc}(X)$. Since $x \in \mathrm{desc}(X)$, there exists $y \in X$ such that $|B_x \cap B_y| \geq n/w$. Hence $\partial(x, X) \leq \alpha$. Since C is a $(w; \alpha)$-FP code,

$$\begin{aligned} |B_z - B_x| &\geq |B_z - \bigcup_{y \in X} B_y| \\ &\geq \alpha + 1 \end{aligned}$$

for any $z \in C - X$. Hence $\partial(x, C - X) \geq \alpha + 1 > \partial(x, X)$. □

From the above two Propositions, we reprove the following results.

Corollary 3.6. *(Staddon, Stinson and Wei, 2001)[14] Suppose C is an $(n, N, q; d)$-code with $d > n(1 - \frac{1}{w^2})$. Then C is an w-TA (n, N, q)-code.* □

4 Complexes

Definition 4.1. A set system $(P, \mathcal{B})$ is an $(\ell, s; e)$-*cover-free family* ($(\ell, s; e)$-CFF) whenever for any ℓ members $A_1, A_2, \ldots, A_\ell \in \mathcal{B}$ and any other s members $B_1, B_2, \ldots, B_s \in \mathcal{B}$,

$$|\bigcap_{i=1}^{\ell} A_i - \bigcup_{j=1}^{s} B_j| \geq e + 1.$$

By an $(\ell, s; e)$-*disjunct matrix* M we mean an incidence matrix of some $(\ell, s; e)$-cover-free family $(P, \mathcal{B})$, i.e. M is a binary matrix with rows and columns indexed by $\mathcal{B}$ and P respectively such that

$$M_{ij} = \begin{cases} 1, & \text{if } j \in i; \\ 0, & \text{if } j \notin i. \end{cases}$$

Our matrix is the transpose of the one studied in pooling designs [6].

In the language of pooling designs, the above ℓ is refer to the size of *complexes*, s to the number of *positive complexes*, e to the number of allowed *test errors*, $|P|$ to the number of *tests*, and $|\mathcal{B}|$ to the number of *items* respectively.

Theorem 4.2. *Let C be an (n, N)-code. Then the set system $([n]\times Q, \{B_c\}_{c\in C})$ is an $(w, w; 0)$-CFF if and only if C is an w-SFP code for $1 \le w \le n-1$.*

Proof. ($\Longrightarrow$) Pick any $X, Y \subseteq C$ with $|X|, |Y| \le w$ and $X \cap Y = \emptyset$. Then $\bigcap_{x\in X} B_x - \bigcup_{y\in Y} B_y = \emptyset$ by assumption. Choose

$$(i, c_i) \in \bigcap_{x\in X} B_x - \bigcup_{y\in Y} B_y.$$

Then with refering to the Definition 2.2, $X_i = \{c_i\}$ and $c_i \notin Y_i$. Hence $X_i \cap Y_i = \emptyset$. Thus $\mathrm{desc}(X) \cap \mathrm{desc}(Y) = \emptyset$.

($\Longleftarrow$) Pick any $X, Y \subseteq C$ with $|X|, |Y| \le w$ and $X \cap Y = \emptyset$. Then $\mathrm{desc}(X) \cap \mathrm{desc}(Y) = \emptyset$. That is

$$X_i \cap Y_i = \emptyset \text{ for some } i.$$

Note that $X_i \neq \{0, 1\}$, $X_i \neq \emptyset$, and similarly for Y_i. Hence we can assume $X_i = \{0\}$ and $Y_i = \{1\}$. Then $(i, 0) \in \bigcap_{x\in X} B_x - \bigcup_{y\in Y} B_y$. $\square$

Unlike Lemma 3.2, here we only can consider the binary code in Theorem 4.2.

Example 4.3. Set

$$C = \{100, 010, 001, 111\}.$$

It is easy to see that C is a 2-SFP $(3, 4)$-code by computing $\mathrm{desc}(X) \cap \mathrm{desc}(Y) = \emptyset$ for all $X, Y \subseteq C$ with $|X| = |Y| = 2$. The following $(2, 2; 0)$-CFF is equivalent to the 2-SFP $(3, 4)$-code presented

$$\begin{aligned} P &= \{(1,0),(1,1),(2,0),(2,1),(3,0)(3,1)\}, \\ \mathcal{B} &= \{\{(1,1),(2,0),(3,0)\},\{(1,0),(2,1),(3,0)\}, \\ &\quad \{(1,0),(2,0),(3,1)\},\{(1,1),(2,1),(3,1)\}\}. \end{aligned}$$

Lemma 4.4. *Set* $P = [n] = \{1, 2, \ldots, n\}$ *and* $\mathcal{B} = \begin{pmatrix} [n] \\ n-1 \end{pmatrix}$, *the set of* $(n-1)$*-subsets of* P. *Then* $(P, \mathcal{B})$ *is an* $(\ell, 1; 0)$*-CFF.*

Proof. For any ℓ members $A_1, A_2, \cdots, A_\ell \in \begin{pmatrix} [n] \\ n-1 \end{pmatrix}$, and other $B \in \begin{pmatrix} [n] \\ n-1 \end{pmatrix}$, note that

$$\bigcap_{i=1}^{\ell} A_i \in \begin{pmatrix} [n] \\ n-\ell \end{pmatrix},$$

and $|\bigcap_{i=1}^{\ell} A_i - B| = 1$. □

Motivated by the above fact $B \not\subseteq \bigcap_{i=1}^{\ell} A_i$ in the proof of Lemma 4.4 , we immediately have the following theorem.

Theorem 4.5. *Fix $n-\ell \le n-1$. Let M denote the incidence matrix of $\binom{[n]}{n-1}$ and $\binom{[n]}{n-\ell}$ i.e. M is a binary matrix with rows and columns indexed by $\binom{[n]}{n-1}$ and $\binom{[n]}{n-\ell}$ respectively such that $M_{ij} = \begin{cases} 1, & \text{if } j \subseteq i; \\ 0, & \text{if } j \nsubseteq i. \end{cases}$; Then M is an $(\ell, s; 0)$-disjunct matrix of size $n \times \binom{n}{\ell}$, where $\ell + s \le n$.* □

Note that when $\ell = 1$ the above M is an identity matrix, hence we refer this construction as a *trivial construction.*

5 Allowing Test Errors

Recalling the definition of $(w;\alpha)$-FP code, we want to construct $(w;\alpha)$-CFF by means of the disjunct matrix. In the study of pooling design, this α is related to the error correcting ability [8]. The following theorem give a construction of disjunct matrices with some error correcting ability.

Theorem 5.1. *Fix $s < n-\ell \le n-1$. Let M denote the incidence matrix of $\binom{[n]}{n-1}$ and $\binom{[n]}{n-\ell-1}$. Then M is an $(\ell, s; n-\ell-s-2)$-disjunct matrix of size $n \times \binom{n}{\ell+1}$.*

Proof. Pick any distinct $A_1, A_2, \cdots, A_\ell, B_1, B_2, \cdots, B_s \in \begin{pmatrix} [n] \\ n-1 \end{pmatrix}$. Then

$$(\bigcap_i^\ell A_i) \cap B_j \in \begin{pmatrix} [n] \\ n-\ell-1 \end{pmatrix}$$

for any $1 \leq j \leq s$. Note that there are $n-\ell$ $(n-\ell-1)$-subsets contained in $\bigcap_i^\ell A_i$ and s of then are contained in some B_j for $1 \leq j \leq s$. Hence we still can pick $e+1 = n-\ell-1-s$ $(n-\ell-1)$-subsets which are contained in each of A_i, but none of B_j. $\square$

We believe the existence of a $(\ell, s; e)$-disjunct matrix is applicable to the study of codes for copyright protection with error correcting ability. Further study is necessary.

6 A Simple Construction of SFP codes

An (n, N)-code C can be depicted as an $N \times n$ binary matrix M, where each row of the matrix corresponds to one of the codewords.

Example 6.1. Let $C = \{c^{(1)} = 111, c^{(2)} = 100, c^{(3)} = 010, c^{(4)} = 001\}$, and C can be depicted as

$$M = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

We will show that C is a 2-SFP $(3, 4)$-code by computing $\mathrm{desc}(X)$ for all X with $|X| = 2$:

$$\begin{aligned}
\mathrm{desc}(\{c^{(1)}, c^{(2)}\}) &= \{100, 111, 101, 110\}, \\
\mathrm{desc}(\{c^{(1)}, c^{(3)}\}) &= \{010, 111, 011, 110\}, \\
\mathrm{desc}(\{c^{(1)}, c^{(4)}\}) &= \{001, 111, 011, 101\}, \\
\mathrm{desc}(\{c^{(2)}, c^{(3)}\}) &= \{100, 010, 110, 000\}, \\
\mathrm{desc}(\{c^{(2)}, c^{(4)}\}) &= \{100, 001, 101, 000\},
\end{aligned}$$

and

$$\mathrm{desc}(\{c^{(3)}, c^{(4)}\}) = \{010, 001, 000, 011\}.$$

From this, it can easily be checked that

$$\begin{aligned}
\mathrm{desc}(\{c^{(1)}, c^{(2)}\}) \cap \mathrm{desc}(\{c^{(3)}, c^{(4)}\}) &= \emptyset, \\
\mathrm{desc}(\{c^{(1)}, c^{(3)}\}) \cap \mathrm{desc}(\{c^{(2)}, c^{(4)}\}) &= \emptyset,
\end{aligned}$$

and

$$\mathrm{desc}(\{c^{(1)}, c^{(4)}\}) \cap \mathrm{desc}(\{c^{(2)}, c^{(3)}\}) = \emptyset.$$

Next, we collect some direct and explicit constructions for secure frame proof codes.

Theorem 6.2. *(D.R. Stinson, Tran van Trung and R. Wei, 2000)[13]. For any integer $w \geq 2$, there is a w-SFP $\left(\binom{2w-1}{w-1}, 2w \right)$-code.*

Proof. We define a binary matrix M and the rows of M will be a w-SFP $(\binom{2w-1}{w-1}, 2w)$-code. The rows of M are indexed by the elements in the set $\{1, \ldots, 2w\}$, and the columns are indexed by the w-subsets $S \subseteq \{1, \ldots, 2w\}$ such that $1 \in S$. Denote these subsets as $S_1, \ldots, S_n$, where $n = \binom{2w-1}{w-1}$. Now, the entry in row i and column j of M is defined to be

$$M_{ij} = \begin{cases} 1 & \text{if } i \in j, \\ 0 & \text{if } i \notin j. \end{cases}$$

We show that $C = \{c^{(1)}, \ldots, c^{(2w)}\}$ is a w-SFP $(\binom{2w-1}{w-1}, 2w)$-code. It suffices to verify that Definition is satisfied for all $X, Y \subseteq C$ such that $|X| = |Y| = w$ and $X \cap Y = \emptyset$. Since $N = 2w$, it follows that $Y = C \setminus X$. Without loss of generality, suppose that $c^{(1)} \in X$. Now, there is a unique bit position i such that $X_i = \{1\}$ and $Y_i = \{0\}$ which implies $X_i \cap Y_i = \emptyset$. Hence, $\mathrm{desc}(X) \cap \mathrm{desc}(Y) = \emptyset$, as desired. $\square$

Example 6.3. The 2-SFP $(3, 4)$-code given in Example 6.1 is constructed by the method of Theorem 6.2.

Example 6.4. We present a 3-SFP $(10, 6)$-code constructed using the method described in Theorem 6.2. The binary matrix M is as follows:

$$M = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}$$

The following result can be proved in a similar way.

Theorem 6.5. *(D.R. Stinson, Tran van Trung and R. Wei, 2000)[13]. For any integer $w \geq 2$, there is a w-SFP $(2\begin{pmatrix} 2w-1 \\ w-1 \end{pmatrix}, 2w+1)$-code.*

Proof. Let the $2w \times \begin{pmatrix} 2w-1 \\ w-1 \end{pmatrix}$ matrix M be defined as in Theorem 6.2. Then construct a $(2w+1) \times 2\begin{pmatrix} 2w-1 \\ w-1 \end{pmatrix}$ matrix M' as follows:

$$M' = \begin{bmatrix} M & M \\ 0 \cdots 0 & 1 \cdots 1 \end{bmatrix}.$$

It is not hard to show that the set of rows in M' is the incidence matrix of a w-SFP $(2\begin{pmatrix} 2w-1 \\ w-1 \end{pmatrix}, 2w+1)$-code. □

7 A Simple Construction of IPP Codes

We depict an $(n, N, q; d)$-code C as an $N \times n$ matrix $M(C)$ on q symbols, where each row of the matrix corresponds to one of the codewords of C. For any $a \in Q$, define

$$m_j(a) = |\{i \mid M(C)_{ij} = a\}|,$$

i.e.,$m_j(a)$ is the frequency of a on the j-th column of $M(C)$. Define

$$m(C) = \max_{1 \leq j \leq n, a \in Q}(m_j(a)).$$

Example 7.1. Set

$$M(C) = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix},$$

then $m_1(0) = 2, m_2(0) = 2, m_3(0) = 0$ and $m_1(1) = 1, m_2(1) = 1, m_3(1) = 3$. So $m(C) = 3$.

Definition 7.2. Let C be an $(n, N, q; d)$-code.We say that C has an σ-*resolution* if the codewords of C can be partitioned into s subsets $A_1,\ldots,A_s$, where $|A_i| = \sigma$, for $i = 1, \ldots, s$, in such a way that each A_i is a code of minimum distance equal to n, i.e., any two codewords of A_i agree in no position.

We see an example.

Example 7.3. Set

$$C = \{123, 132, 213, 231, 312, 321\}$$

be a $(3, 6, 3; 2)$-code. Set

$$A_1 = \{123, 231, 312\}, A_2 = \{132, 321, 213\}.$$

Since C can be partitioned into 2-subsets A_1, A_2, and the minimum distance of A_1 and A_2 are equal to $n = 3$, we say C has a 3-resolution.

Theorem 7.4. *(Tran and Sosina, 2004)[16]. Let C_1 be an $(n_1, N_1, q_1; d_1)$-code over Q_1 and let C_2 be an $(n_2, N_2, q_2; d_2)$-code over Q_2 with a σ-resolution $A_1, \ldots, A_s$ such that $s \geq m(C_1)$. Then the following hold.*

(i) there exist an $(n_1 n_2, \sigma N_1, q_1 q_2; n_1 n_2 - (n_1 - d_1)(n_2 - d_2))$ code C.

(ii) Further, if $q_1 q_2 \geq N_1$, then C can be extended to a code C^* having parameters $(n_1 n_2 + 1, \sigma N_1, q_1 q_2; d)$, where $d = \min\{n_1 n_2; n_1 n_2 + 1 - (n_1 - d_1)(n_2 - d_2)\}$.

Proof. Let C_1 be an $(n_1, N_1, q_1; d_1)$-code over Q_1. Let C_2 be an $(n_2, N_2, q_2; d_2)$ code over Q_2 with a σ-resolution $A_1, \ldots, A_s$. Suppose $s \geq m(C_1)$. For each $a \in Q_1$ denote by $C_2(a)$ a copy of C_2 defined over $Q(a)$ such that

$$Q(a_1) \cap Q(a_2) = \emptyset \text{ if } a_1, a_2 \in Q_1 \text{ and } a_1 \neq a_2.$$

Denote by $A_1(a), \ldots, A_s(a)$ a σ-resolution of $C_2(a)$.

Let $col_j = (a_{1,j}, a_{2,j}, \ldots, a_{b_1,j})^T$ be the j-th column of $M(C_1)$, $1 \leq j \leq n_1$. Let $a(1), \ldots, a(t)$, say, be t positions of col_j at which symbol $a \in Q_1$ appears. Note that $t \leq m(C_1)$. Now replace a at position $a(1)$ by $A_1(a)$, a at position $a(2)$ by $A_1(a)$, etc., and a at position $a(t)$ by $A_t(a)$. Perform this process for every symbol of Q_1 and for every column of $M(C_1)$. The resulting code C obtained by this replacement has parameters $(n_1 n_2, \sigma N_1, q_1 q_2; n_1 n_2 - (n_1 - d_1)(n_2 - d_2))$.

Obviously, the length and the number of codewords of C is $n_1 n_2$ and σN_1 respectively. Further, any two codewords $c_1, c_2 \in C_1$ agree in at most

$(n_1 - d_2)$ positions. After replacement c_1 and c_2 correspond to two subsets R_1 and R_2 of σ codewords each. Any two codewords in R_1 (resp. R_2) agree in no position, whereas a codeword from R_1 and a codeword from R_2 agree in at most $(n_1 - d_1)(n_2 - d_2)$ positions. Hence the minimum distance of C is $n_1 n_2 - (n_1 - d_1)(n_2 - d_2)$, as stated.

Further, if $q_1 q_2 \geq N_1$ then C can be extended to a code C^* having parameters $(n_1 n_2 + 1, \sigma N_1, q_1 q_2; d)$, where $d = \min\{n_1 n_2, n_1 n_2 + 1 - (n_1 - d_1)(n_2 - d_2)\}$. Let $Q = \{a_1, a_2, \ldots, a_{q_1 q_2}\}$ be the alphabet of C and let $C_1 = \{c_1, c_2, \ldots, c_{N_1}\}$. By construction, any codeword $c_i \in C_1$ corresponds to a subset R_i of σ codewords. For any $i = 1, \ldots, N_1$, we add symbol a_i to the $(n_1 n_2 + 1)$-th column of each codeword of R_i. This forms a set R_i^*. The collection of all R_i^* forms an $(n_1 n_2 + 1, \sigma N_1, q_1 q_2; d)$ code C^* with $d = \min\{n_1 n_2, n_1 n_2 + 1 - (n_1 - d_1)(n_2 - d_2)\}$. This can be seen as follows. Any two codewords x^* and y^* of C^* belong either to some R_i^* or to two different R_i^* and R_j^*. In the first case their distance is $n_1 n_2$ because their components agree only at the $(n_1 n_2 + 1)$-th column, and in the second case their distance is at least $n_1 n_2 + 1 - (n_1 - d_1)(n_2 - d_2)$ because their components at the $(n_1 n_2 + 1)$-th column are distinct. □

We illustrate the construction in Theorem 7.4 by the following example.

Example 7.5. Let C_1 be a $(3, 4, 2; 2)$-code over $Q_1 = \{0, 1\}$ given by

$$M(C_1) = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Let $C_2(0)$ be a $(3, 6, 3; 2)$-code over $\{1, 2, 3\}$ with a 3-resolution $A_1(0)$ and

$A_2(0)$:

$$\mathrm{A}_1(0) = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix}, \mathrm{A}_2(0) = \begin{bmatrix} 1 & 3 & 2 \\ 3 & 2 & 1 \\ 2 & 1 & 3 \end{bmatrix}.$$

Let $C_2(1)$ be a copy of $C_2(0)$ over $\{4,5,6\}$ with a 3-resolution

$$\mathrm{A}_1(1) = \begin{bmatrix} 4 & 5 & 6 \\ 5 & 6 & 4 \\ 6 & 4 & 5 \end{bmatrix}, \mathrm{A}_2(0) = \begin{bmatrix} 4 & 6 & 5 \\ 6 & 5 & 4 \\ 5 & 4 & 6 \end{bmatrix}.$$

Replacing entries of $M(C_1)$ by $A_i(j)$ gives

$$\begin{bmatrix} A_1(0) & A_1(1) & A_1(1) \\ A_1(1) & A_1(0) & A_2(1) \\ A_2(1) & A_2(1) & A_1(0) \\ A_2(0) & A_2(0) & A_2(0) \end{bmatrix}.$$

Thus, we obtain a $(9, 12, 6; 8)$-code C. Now, since the condition $q_1q_2 > N_1$ is satisfied, C can be extended to a $(10, 12, 6; 9)$-code C^*.

$$
M(C) = \begin{bmatrix}
1 & 2 & 3 & 4 & 5 & 6 & 4 & 5 & 6 \\
2 & 3 & 1 & 5 & 6 & 4 & 5 & 6 & 4 \\
3 & 1 & 2 & 6 & 4 & 5 & 6 & 4 & 5 \\
4 & 5 & 6 & 2 & 3 & 1 & 6 & 5 & 4 \\
5 & 6 & 4 & 2 & 3 & 1 & 6 & 5 & 4 \\
6 & 4 & 5 & 3 & 1 & 2 & 5 & 4 & 6 \\
4 & 6 & 5 & 4 & 6 & 5 & 1 & 2 & 3 \\
6 & 5 & 4 & 6 & 5 & 4 & 2 & 3 & 1 \\
5 & 4 & 6 & 5 & 4 & 6 & 3 & 1 & 2 \\
1 & 3 & 2 & 1 & 3 & 2 & 1 & 3 & 2 \\
3 & 2 & 1 & 3 & 2 & 1 & 3 & 2 & 1 \\
2 & 1 & 3 & 2 & 1 & 3 & 2 & 1 & 3
\end{bmatrix},
$$

$$M(C^*) = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 4 & 5 & 6 & 1 \\ 2 & 3 & 1 & 5 & 6 & 4 & 5 & 6 & 4 & 1 \\ 3 & 1 & 2 & 6 & 4 & 5 & 6 & 4 & 5 & 1 \\ 4 & 5 & 6 & 2 & 3 & 1 & 6 & 5 & 4 & 2 \\ 5 & 6 & 4 & 2 & 3 & 1 & 6 & 5 & 4 & 2 \\ 6 & 4 & 5 & 3 & 1 & 2 & 5 & 4 & 6 & 2 \\ 4 & 6 & 5 & 4 & 6 & 5 & 1 & 2 & 3 & 3 \\ 6 & 5 & 4 & 6 & 5 & 4 & 2 & 3 & 1 & 3 \\ 5 & 4 & 6 & 5 & 4 & 6 & 3 & 1 & 2 & 3 \\ 1 & 3 & 2 & 1 & 3 & 2 & 1 & 3 & 2 & 4 \\ 3 & 2 & 1 & 3 & 2 & 1 & 3 & 2 & 1 & 4 \\ 2 & 1 & 3 & 2 & 1 & 3 & 2 & 1 & 3 & 4 \end{bmatrix}$$

We describe a simple construction for q-ary codes which has been presented by Bush (1952)[3] for orthogonal arrays.

Theorem 7.6. *(Bush,1952)[3]. Let C_1 be an $(n, N_1, q_1; d_1)$-code over Q_1 and C_2 be an $(n, N_2, q_2; d_2)$-code. Then there exists an $(n, N_1N_2, q_1q_2; d)$-code, where $d = \min\{d_1, d_2\}$.*

Proof. Let C_2 be an $(n, N_1, q_1; d_1)$-code over Q_1 and let C_2 be an $(n, N_2, q_2; d_2)$-code over Q_2. Let $Q = Q_1 \times Q_2$. We define a code C over Q as follows. For any pair of codewords $\mathbf{a} = (a_1, ..., a_n) \in C_1$ and $\mathbf{b} = (b_1, ..., b_n) \in C_2$ we construct a vector

$$\mathbf{c(a,b)} = ((a_1, b_1), ..., (a_n, b_n)) \in Q^n.$$

Then it is easy to verify that

$$C = \{\mathbf{c}(\mathbf{a}, \mathbf{b}) \mid \mathbf{a} \in C_1, \mathbf{b} \in C_2\} \subseteq Q^n$$

is an $(n, N_1N_2, q_1q_2; d)$-code, where $d = \min\{d_1, d_2\}$. □

Definition 7.7. A code $C \subseteq F_q^n$ is a $[n, k, d]$-linear code if C is a subspace of F_q^n with dimension k and minimum distance d.

Definition 7.8. A $[n, k, d]$-linear code with $d = n-k+1$ is called a maximum distance separable code, denoted MDS codes.

Theorem 7.6 can be used to construct q-ary codes achieving MDS codes, for which q is not a prime power, in the language of orthogonal arrays an $(n, N, q; d)$ MDS code is an $OA_1(n-d+1, n, q)$; here we have $N = q^{n-d+1}$.

We record this special case of the Bush construction in the following theorem.

Theorem 7.9. *(Bush, 1952)[3] The existence of* $(n, q_1^k, q_1; d)$ *and* $(n, q_2^k, q_2; d)$ *MDS codes having the same* $d = n - k + 1$ *implies the existence of an* $(n, (q_1q_2)^k, q_1q_2; d)$ *MDS code.*

As a consequence of Theorem 7.9 , we have the following corollary.

Corollary 7.10. *For any integer* $n \geq 2$ *and* s *with a prime factorization* $s = p_1^{e_1}...p_r^{e_r}$ *such that* $n \leq p_i^{e_i}$, $i = 1, 2, ..., r$, *there is an* (n, s^k, s) *MDS codes, for all* $2 \leq k \leq n$.

Proof. The corollary follows from the existence of $(n, (p_i^{e_i})^k, (p_i^{e_i}))$ MDS codes for $i = 1, ..., r$. $\square$

By combining Corollary 7.10 and Corollary 3.6 we obtain the following theorem.

Theorem 7.11. *Let $w \geq 2$ be any given integer. For any integer $n > w^2$ and s having $s = p_1^{e_1}...p_k^{e_k}$ as its prime factorization with $n \leq p_i^{e_i}$ for all $i = 1, ..., k$ there exists an w-IPP (n, N, s)-code, where $N = s^{\lceil \frac{n}{w^2} \rceil}$.* $\square$

Definition 7.12. Let C_1 be an (n_2, N_2, q_2)-code over Q_2 and let C_2 be an (n_1, q_2, q_1)-code over Q_1. We define the concatenated code of C_1 and C_2 as following: Let $Q_2 = \{a_1, ..., a_{q_2}\}$ and let $C_2 = \{\mathbf{b_1}, ..., \mathbf{b_{q_2}}\}$. Let $\theta : Q_2 \to C_2$ be the one-to-one mapping defined by

$$\theta(a_i) = \mathbf{b_i}$$

for $1 \leq i \leq q_2$. For any codeword $\mathbf{a} = (a_1, ..., a_{n_2}) \in C_1$ we denote by

$$\tilde{\mathbf{a}} = (\theta(a_1), ..., \theta(a_{n_2})) = (\mathbf{b_1}, ..., \mathbf{b_{n_2}})$$

the q_1-ary sequence of length $n_1 n_2$ obtained from $\mathbf{a}$ by using θ. The set

$$C = \{\tilde{\mathbf{a}} = (\mathbf{b_1}, ..., \mathbf{b_{n_2}}) \mid \mathbf{a} = (a_1, ..., a_{n_2}) \in C_1\}$$

is an $(n_1 n_2, N_2, q_1)$-code, called the concatenated code of C_1 and C_2.

Example 7.13. Set

$$C_1 = \{12, 13, 23\}$$

be a (2,3,3)-code over $\{a_1 = 1, a_2 = 2, a_3 = 3\}$. Set

$$C_2 = \{b_1 = 445, b_2 = 455, b_3 = 555\}$$

be a $(3,3,2)$-code over $\{4,5\}$. Define θ be the one to one mapping by $\theta(a_i) = b_i$ for $i = 1,2,3$. Then the concatenated code C of C_1 and C_2 presented

$$C = \{(445, 455), (445, 555), (455, 555)\}$$

be a $(6,3,2)$-code.

Next important theorem shows that the concatenation technique works for IPP codes.

Theorem 7.14. *(Tran and Sosina, 2005)[17]. Let C_1 be an w-IPP (n_2, N_2, q_2)-code over Q_2 and let C_2 be an w-IPP (n_1, q_2, q_1)-code over Q_1. Then the concatenated code C of C_1 and C_2 is an w-IPP (n_1n_2, N_2, q_1)-code.*

Proof. Let $\mathbf{x} = (x_1, ..., x_{n_1n_2}) \in Q_1^{n_1n_2}$. We partition $\mathbf{x}$ into n_2 blocks $\mathbf{x_1}, ..., \mathbf{x}_{n_2}$ with $\mathbf{x}_i = (x_{(i-1)n_1+1}, ..., x_{in_1}) \in Q_1^{n_1}$, $1 \leq i \leq n_2$. We will write $\mathbf{x} = (\mathbf{x}_1, ..., \mathbf{x}_{n_2})$. Specially, if $\mathbf{x} = \mathbf{c} = (\mathbf{b_1}, ..., \mathbf{b_{n_2}}) \in C$, then $\mathbf{b}_i's$ are themselves blocks of the partition of $\mathbf{c}$.

Suppose $\mathbf{x} \in \operatorname{desc}(C_i)$, $1 \leq i \leq r$, where $C_i \subseteq C$ with $|C_i| = \alpha_i \leq w$. We prove that $\bigcap_{1 \leq i \leq r}(C_i) \neq \emptyset$, i.e. C is a w-IPP code.

Let $C_i = \{\mathbf{c}_1^{(i)}, ..., \mathbf{c}_{\alpha_i}^{(i)}\} \subseteq C$, where $\mathbf{c}_j^{(i)} = (\mathbf{b}_{j1}^{(i)}, ..., \mathbf{b}_{jn_2}^{(i)})$. For any $1 \leq i \leq r$ and any $1 \leq \ell \leq n_2$ define $D_\ell^{(i)} = \{\mathbf{b}_{1\ell}^{(i)}, ..., \mathbf{b}_{\alpha_i\ell}^{(i)}\}$, i.e. $D_\ell^{(i)}$ is the collection of all ℓth blocks of the codewords of C_i. In other words, $D_\ell^{(i)} \subseteq C_2$ is a subset of α_i codewords. As $\mathbf{x} \in \operatorname{desc}(C_i)$ by the assumption, we have $\mathbf{x}_\ell \in \operatorname{desc}(D_\ell^{(i)})$ for $1 \leq i \leq r$ and $1 \leq \ell \leq n_2$. Since C_2 is a w-IPP code, we have

$$\bigcap_{1 \leq i \leq r} D_\ell^{(i)} \neq \emptyset.$$

Let $\mathbf{b}_\ell \in \bigcap_{1 \leq i \leq r} D_\ell^{(i)}$ be an arbitrary but fixed codeword, i.e. $\mathbf{b}_\ell$ is a guilty user for $\mathbf{x}_\ell$ in code C_2. Set $\mathbf{y} = (\mathbf{b}_1, ..., \mathbf{b}_{n_2})$. Let $\bar{\mathbf{y}} = (a_1, ..., a_{n_2}) \in Q^{n_2}$ be the corresponding sequence obtained from $\mathbf{y}$ using θ, i.e. $a_i = \theta^{-1}(\mathbf{b}_i)$. In the same way let $\bar{C}_i = \{\bar{\mathbf{c}}_1^{(i)}, ..., \bar{\mathbf{c}}_{\alpha_i}^{(i)}\} \subseteq C_1$ denote the corresponding subset of C_i.

Since $\mathbf{y} \in \mathrm{desc}(C_i)$ by the construction, we have $\bar{\mathbf{y}} \in \mathrm{desc}(\bar{C}_i)$. for $1 \leq i \leq r$. Hence

$$\bar{\mathbf{y}} \in \bigcap_{1\leq i\leq r} \mathrm{desc}(\bar{C}_i).$$

Since C_1 is a w-IPP code, we have

$$\bigcap_{1\leq i\leq r} \bar{C}_i \neq \emptyset.$$

Let $\bar{\mathbf{z}}' = (a'_1, ..., a'_{n_2}) \in \bigcap_{1\leq i\leq r}(\bar{C}_i)$ be a guilty user for $\bar{\mathbf{y}}$ in C_1. Then $\mathbf{z}' = (\mathbf{b}'_1, ..., \mathbf{b}'_{n_2}) \in C_i$ for $1 \leq i \leq r$, where $\mathbf{z}'$ the codeword of C corresponding to $\bar{\mathbf{z}}'$. Therefore

$$\bigcap_{1\leq i\leq r} C_i \neq \emptyset.$$

Thus C is an w-IPP code. □

References

[1] Biehl, I., Meyer, B., Protocols for collusion-secure asymmetric fingerprinting. 14th Symposium on Theoretical Aspects of Computing, *Lecture Notes in Computer Science,* Vol. 1200 (1997). Springer, Berlin, pp. 399-412.

[2] Boneh, D., Shaw, J., Collusion-secure fingerprinting for digital data, *Advances in Cryptology-Crypto '95. Lecture Notes in Computer Science,* Vol. 963. Springer, Berlin, (1995) pp. 452-465.

[3] Bush, K.A., Federer, W.T., Pesotan, H., Raghavarao, D., New combinatorial designs and their application to group testing, *J. Statist. Plann. Inference* 10, (1984) 335-343.

[4] Chee, Y.M., Turán-type problems in group testing, coding theory and cryptography, *Ph.D. Thesis, University of Waterloo. (1996)*

[5] Chor, B., Fiat, A., Naor, M., Tracing traitors. Advances in Cryptology-Crypto '94, *Lecture Notes in Computer Science,* Vol. 839. Springer, Berlin, (1994) pp. 257-270.

[6] Du, D.-Z., Hwang, F.K., Pooling designs and nonadaptive group testing, *World Scientific, Singapore. (2006)*

[7] A. J. Macula. A simple construction of d-disjunct matrices with certain constant weights, *Discrete Math. (1996)* 162:311–312.

[8] A. J. Macula,. Error-correcting nonadaptive group testing with d^e-disjunct matrices, *Discrete Appl. Math. (1997)* 80:217-222.

[9] Marcel Fernandez and Miguel Soriano, Intellectual property protection of e-learing contents, *International conference on network universities and e-Learing.* 8-9 May (2003), Valencia. Spain.

[10] Pfitzmann, B., Trials of traced traitors, *Workshop on Information Hiding, Lecture Notes in Computer Science,* Vol. 1174. Springer, Berlin, (1996) pp. 49-64.

[11] Pfitzmann, B., Waidner, M., Asymmetric fingerprinting for larger collusions, *Fourth ACM Conference on Computer and Communications Security. (1997a)*

[12] Pfitzmann, B., Waidner, M., Anonymous fingerprinting. Advances in Cryptology-Eurocrypt '97, *Lecture Notes in Computer Science,* Vol. 1233. Springer, Berlin, (1997b) pp. 88-102.

[13] J. N. Staddon, D. R. Stinson and R. Wei, Combinatorial properties of frameproof and traceability codes, *IEEE Trans. Inform. Theory,* Vol. 47 (2001) pp. 1042-1049.

[14] D. R. Stinson, Tran van Trung and R. Wei, Secure frameproof codes, key distribution patterns, group testing algorithms and related structures, *Journal of Statistical Planning and Inference* 86 (2000) 595-617.

[15] Stinson, D.R., Wei, R., Combinatorial properties and constructions of traceability schemes and frameproof codes, *SIAM J. Discrete Math.* 11, (1998)41-53.

[16] Tran van Trung and Sosina Martirosyan, On a class of traceability codes, *Design, Codes and Cryptography,* 31, (2004) 125-132.

[17] Tran van Trung and Sosina Martirosyan, New constructions for IPP codes, *Design, Codes and Cryptography,* 35, (2005) 227-239.