Explicit Construction of Secure Frameproof Codes

Dongvu Tonien
dong@uow.edu.au

Reihaneh Safavi-Naini
rei@uow.edu.au

School of IT & CS, University of Wollongong, NSW, 2522, Australia

Abstract

Γ is a q-ary code of length L. A word w is called a descendant of a coalition of codewords $w^{(1)}$, $w^{(2)}$, ..., $w^{(t)}$ of Γ if at each position i, $1 \leq i \leq L$, w inherits a symbol from one of its parents, that is $w_i \in \{w_i^{(1)}, w_i^{(2)}, \ldots, w_i^{(t)}\}$. A k-secure frameproof code (k-SFPC) ensures that any two disjoint coalitions of size at most k have no common descendant. Several probabilistic methods prove the existance of codes but there are not many explicit constructions. Indeed, it is an open problem in [J. Staddon et al., *IEEE Trans. on Information Theory,* 47 (2001), pp. 1042–1049] to construct explicitly q-ary 2-secure frameproof code for arbitrary q.

In this paper, we present several explicit constructions of q-ary 2-SFPCs. These constructions are generalisation of the binary inner code of the secure code in [V.D. Tô et al., *Proceeding of IndoCrypt'02,* LNCS 2551, pp. 149–162, 2002]. The length of our new code is logarithmically small compared to its size.

1 Introduction

Codes with secure property are used for copyright protection and piracy tracing. In many situations, pirate objects can be modelled as descendant words. For example, in pay-per-view movies described in Fiat et al [6], a movie is divided into L segments and each segment has q different variations with different fingerprintings embeded. Subscribers receive different versions of the movie, and if some of them collude to form a pirate movie, they may select at each segment one of the versions that they have. Therefore, a version of a movie can be thought of as a codeword $(w_1, w_2, \ldots, w_L)$, where w_i is the version number at the segment i. The pirate movie is then a decendant word constructed from the codewords corresponding to the colluders' movie versions.

Other example is in broadcast encryption. Chor et al [4] describle a scheme in which the session key is divided into L shares. Each share is encrypted independently with q different keys. An authorized user is given a decoder box which contains L keys, each key enables it to decrypt a share. Again, we can view the decoder as a codeword $(w_1, w_2, \ldots, w_L)$, where w_i is a decryption key for the i^{th} share. A coalition of users can collude to create a pirate decoder where the i^{th} key is selected from the collection of the i^{th} keys in the colluders' decoder boxes. Therefore, a pirate decoder can be viewed as a descendant word produced from the coalition. Secure frameproof code ensures that a coalition of users cannot frame other disjoint coalition by creating a pirate word which is possibly created by the second coalition.

Several probabilistic methods prove the existance of codes but there are not many explicit constructions. Indeed, it is an open problem in Staddon et al [11] to construct explicitly q-ary 2-secure frameproof code for arbitrary q.

Since the code size corresponds to the number of users and the transmission bandwidth is proportional to the code length and the alphabet size, it is desirable to construct code with large size but relatively small length and small alphabet. In [5], Encheva et al. have several explicit constructions of 2-secure frameproof codes. Using Hadamard matrix, a binary code is constructed, the size N and length L of this code are the same and equal to 2^n. For other q-ary codes, the size is linear to (length $\times$ q). In Tô et al [14], 2-secure code with efficient tracing algorithm is introduced. This is a two level construction which combines a binary 2-secure frameproof code with some outer structures such as error-correcting code or perfect hash family. With a parameter n, the inner code has size $N = n$ and length $L = \binom{n}{2}$.

In this paper, we present a number of new constructions of 2-secure frameproof codes. These codes are generalisation of the inner code used in Tô et al [14]. The size of our new codes are exponentially large compared to the length. The rest of the paper is organised as follows. In Section 2, we list all the basic definitions and known results that will be used throughout the paper. In Section 3, we define our new code families $\Gamma_{t,r}(n)$, $\Gamma_{t,\leq r}(n)$, $\Gamma^*_{t,r}(n)$ and $\Gamma^*_{t,\leq r}(n)$. We prove the equivalence of the Tô et al [14] inner code with our code $\Gamma_{1,2}(n)$ in Section 4. Secure frameproof property of code $\Gamma_{2,2}(n)$ is proved in Section 5. Sections 6 and 7 are dealing with the general code $\Gamma_{t,r}(n)$, $\Gamma_{t,\leq r}(n)$ and binary code $\Gamma^*_{t,r}(n)$, $\Gamma^*_{t,\leq r}(n)$ respectively. We conclude our paper by summarizing our code parameters, constraints and comparing them with the Encheva et al [5] codes in Section 8.

2 Preliminaries

Let Γ be a q-ary code of length L and size N. We have $\Gamma \subset Q^L$, where Q denotes a set of alphabets, and $|\Gamma| = N$. Each element of Γ is called a codeword and can be written as $w = (w_1, w_2, \ldots, w_L)$, where $w_i \in Q$. Elements of Q^L in general are called words.

For a subset $C \subset \Gamma$, we define the *projection* of C on the position i as

$$\pi_i(C) = \{w_i : w \in C\}$$

and the *descendants set* of C as

$$Desc(C) = \{w \in Q^L : w_i \in \pi_i(C),\ \forall i,\ 1 \leq i \leq L\}$$

$Desc(C)$ is the set of all words which can be constructed from the coalition C. An element w of $Desc(C)$ is called a *descendant* of C and elements of C are called *parents* of w. From the definition, at any position i, w inherits a symbol from one of its parents. Hence, $Desc(C)$ is the Cartesian product of $\pi_i(C)$

$$Desc(C) = \prod_{i=1}^{L} \pi_i(C)$$

Secure frameproof code ensures that two disjoint coalitions cannot create the same pirate word.

Definition 1 *Let Γ be a q-ary code of length L and size N. If for any two subsets $C_1, C_2 \subset \Gamma$ of size up to k,*

$$C_1 \cap C_2 = \emptyset \quad \rightarrow \quad Desc(C_1) \cap Desc(C_2) = \emptyset \tag{1}$$

then Γ is called a k-secure frameproof code (k-SFPC).

Secure frameproof code is also called *partially identifying code* (Encheva et al [5]).

Since

$$Desc(C_1) \cap Desc(C_2) = \prod_{i=1}^{L} (\pi_i(C_1) \cap \pi_i(C_2)),$$

we have the following lemma immediately followed

Lemma 1 *If Γ is a k-SFPC then for any two subsets $C_1, C_2 \subset \Gamma$ of size up to k, there exists a position i such that their projections on this position are two disjoint sets*

$$\pi_i(C_1) \cap \pi_i(C_2) \quad = \quad \emptyset. \tag{2}$$

In (2), the position i is said to separate C_1 and C_2. Lemma 1 says that for any two disjoint coalitions C_1 and C_2 of size up to k, there must exist a position that separates them.

It is proved in Staddon et al [11] that when $|\Gamma| \geq 2k$, to prove Γ to be k-SFPC, one needs to check the condition (1) in the Definition 1 for only disjoint subsets $C_1, C_2 \subset \Gamma$ of size equal to k.

Theorem 1 *[11] Let Γ be a code of size $N \geq 2k$. Then Γ is a k-SFPC if and only if for any two subsets C_1, C_2 of size k ($|C_1| = |C_2| = k$),*

$$C_1 \cap C_2 = \emptyset \quad \rightarrow \quad Desc(C_1) \cap Desc(C_2) = \emptyset.$$

3 Our new code families $\Gamma_{t,r}(n)$, $\Gamma_{t,\leq r}(n)$, $\Gamma^*_{t,r}(n)$ and $\Gamma^*_{t,\leq r}(n)$

In this section, we describle our new code families $\Gamma_{t,r}(n)$, $\Gamma_{t,\leq r}(n)$, $\Gamma^*_{t,r}(n)$ and $\Gamma^*_{t,\leq r}(n)$.

Let (n) be the set $\{1, 2, \ldots, n\}$. By $(n)_t$ we denote the set of all subsets of (n) which contain exactly t elements. Similarly, $(n)_{\leq t}$ denotes the set of all nonempty subsets of (n) which contain less than or equal to t elements.

With parameters n, t, r, consider the matrix $\mathcal{M}_{t,r}(n)$ whose rows are labelled by elements of $(n)_t$ and columns are labelled by elements of $(n)_r$. For $U \in (n)_t$, $V \in (n)_r$, the entry at the row U and column V of the matrix $\mathcal{M}_{t,r}(n)$ is $|U \cap V|$. The code $\Gamma_{t,r}(n)$ is composed by rows of the matrix $\mathcal{M}_{t,r}(n)$. Without ambiguity, we identify a codeword of $\Gamma_{t,r}(n)$ with a set $U \in (n)_t$ and a position with a set $V \in (n)_r$. And so, by definition, the symbol of the codeword U at the position V is $U_V = |U \cap V|$.

We define the code $\Gamma_{t,\leq r}(n)$ in a similar way. Code $\Gamma_{t,\leq r}(n)$ is depicted by the matrix $\mathcal{M}_{t,\leq r}(n)$ whose rows and columns are labelled by elements of the sets $(n)_t$ and $(n)_{\leq r}$ respectively. For $U \in (n)_t$ and $V \in (n)_{\leq r}$, the symbol of the codeword U at the position V is $U_V = |U \cap V|$.

Codes $\Gamma^*_{t,r}(n)$ and $\Gamma^*_{t,\leq r}(n)$ are binary codes. They are constructed the same as code $\Gamma_{t,r}(n)$ and $\Gamma_{t,\leq r}(n)$ except that the symbol of the codeword U at the position V is $U_V = |U \cap V| \pmod 2$. We can think of codes $\Gamma^*_{t,r}(n)$ and $\Gamma^*_{t,\leq r}(n)$ as the modulo 2 of the previous codes $\Gamma_{t,r}(n)$ and $\Gamma_{t,\leq r}(n)$ respectively.

Example 1 *Codes* $\Gamma_{3,2}(5)$, $\Gamma^*_{3,2}(5)$, $\Gamma_{3,\leq 2}(4)$ *and* $\Gamma^*_{3,\leq 2}(4)$ *are shown below*

$\Gamma_{3,2}(5)$	$\{1,2\}$	$\{1,3\}$	$\{1,4\}$	$\{1,5\}$	$\{2,3\}$	$\{2,4\}$	$\{2,5\}$	$\{3,4\}$	$\{3,5\}$	$\{4,5\}$
$\{1,2,3\}$	2	2	1	1	2	1	1	1	1	0
$\{1,2,4\}$	2	1	2	1	1	2	1	1	0	1
$\{1,2,5\}$	2	1	1	2	1	1	2	0	1	1
$\{1,3,4\}$	1	2	2	1	1	1	0	2	1	1
$\{1,3,5\}$	1	2	1	2	1	0	1	1	2	1
$\{1,4,5\}$	1	1	2	2	0	1	1	1	1	2
$\{2,3,4\}$	1	1	1	0	2	2	1	2	1	1
$\{2,3,5\}$	1	1	0	1	2	1	2	1	2	1
$\{2,4,5\}$	1	0	1	1	1	2	2	1	1	2
$\{3,4,5\}$	0	1	1	1	1	1	1	2	2	2

$\Gamma^*_{3,2}(5)$	$\{1,2\}$	$\{1,3\}$	$\{1,4\}$	$\{1,5\}$	$\{2,3\}$	$\{2,4\}$	$\{2,5\}$	$\{3,4\}$	$\{3,5\}$	$\{4,5\}$
$\{1,2,3\}$	0	0	1	1	0	1	1	1	1	0
$\{1,2,4\}$	0	1	0	1	1	0	1	1	0	1
$\{1,2,5\}$	0	1	1	0	1	1	0	0	1	1
$\{1,3,4\}$	1	0	0	1	1	1	0	0	1	1
$\{1,3,5\}$	1	0	1	0	1	0	1	1	0	1
$\{1,4,5\}$	1	1	0	0	0	1	1	1	1	0
$\{2,3,4\}$	1	1	1	0	0	0	1	0	1	1
$\{2,3,5\}$	1	1	0	1	0	1	0	1	0	1
$\{2,4,5\}$	1	0	1	1	1	0	0	1	1	0
$\{3,4,5\}$	0	1	1	1	1	1	1	0	0	0

$\Gamma_{3,\leq 2}(4)$	$\{1\}$	$\{2\}$	$\{3\}$	$\{4\}$	$\{1,2\}$	$\{1,3\}$	$\{1,4\}$	$\{2,3\}$	$\{2,4\}$	$\{3,4\}$
$\{1,2,3\}$	1	1	1	0	2	2	1	2	1	1
$\{1,2,4\}$	1	1	0	1	2	1	2	1	2	1
$\{1,3,4\}$	1	0	1	1	1	2	2	1	1	2
$\{2,3,4\}$	0	1	1	1	1	1	1	2	2	2

$\Gamma^*_{3,<2}(4)$	$\{1\}$	$\{2\}$	$\{3\}$	$\{4\}$	$\{1,2\}$	$\{1,3\}$	$\{1,4\}$	$\{2,3\}$	$\{2,4\}$	$\{3,4\}$
$\{1,2,3\}$	1	1	1	0	0	0	1	0	1	1
$\{1,2,4\}$	1	1	0	1	0	1	0	1	0	1
$\{1,3,4\}$	1	0	1	1	1	0	0	1	1	0
$\{2,3,4\}$	0	1	1	1	1	1	1	0	0	0

Code $\Gamma_{1,2}(n)$ has $L = o(N^2)$ and code $\Gamma_{2,2}(n)$ has $L = N$. However, the general code $\Gamma_{t,r}(n)$, $\Gamma_{t,\leq r}$, $\Gamma^*_{t,r}(n)$ and $\Gamma^*_{t,\leq r}(n)$ have much shorter length. Especially when r is a small number and t is a large number near $n/2$ then these codes have the length logarithmically small compared to the size.

4 Code $\Gamma_{1,2}(n)$

In this section, we will show that the Tô et al [14] inner code γ is equivalent to our code $\Gamma_{1,2}(n)$.

The Tô et al [14] inner code γ is defined as follows. First, choose two arbitrary permutations $(r_1, r_2, \ldots, r_n)$, $(c_1, c_2, \ldots, c_n)$ of $(1, 2, \ldots, n)$. Code γ contains n codewords $\{w^{(1)}, \ldots, w^{(n)}\}$ of length n^2. Each codeword $w^{(i)}$ is a two dimentional binary array indexed by $[r, c]$ where $1 \leq r \leq n$,

$1 \leq c \leq n$, and

$$w^{(i)}[r,c] = \begin{cases} 1, & \text{if } r = r_i \text{ and } c \neq c_i \\ 1, & \text{if } r \neq r_i \text{ and } c = c_i \\ 0, & \text{otherwise} \end{cases}$$

If we choose $r_i = c_i = i$ then

$$w^{(i)}[r,c] = \begin{cases} 1, & \text{if } r = i \text{ and } c \neq i \\ 1, & \text{if } r \neq i \text{ and } c = i \\ 0, & \text{otherwise} \end{cases}$$

Since all the codewords of γ has the value 0 at the positions $[1,1]$, $[2,2]$, $\ldots$, $[n,n]$, these n positions are redundant and can be removed. Moreover, each codeword $w^{(i)}$ is symmetric in the sense that $w^{(i)}[r,c] = w^{(i)}[c,r]$. Therefore, we only need $\binom{n}{2}$ position $[r,c]$ where $1 \leq r < c \leq n$; and we have

$$w^{(i)}[r,c] = \begin{cases} 1, & \text{if } i \in \{r,c\} \\ 0, & \text{if } i \notin \{r,c\} \end{cases}$$

If we make the correspondance from $w^{(i)}$ to $\{i\} \in (n)_1$ and from $[r,c]$ to $\{r,c\} \in (n)_2$ then it follows that the code γ is equivalent to our code $\Gamma_{1,2}(n)$. And so we have the following theorem. For the sake of completeness, we include a proof here.

$\Gamma_{1,2}(4)$	$\{1,2\}$	$\{1,3\}$	$\{1,4\}$	$\{2,3\}$	$\{2,4\}$	$\{3,4\}$
$\{1\}$	1	1	1	0	0	0
$\{2\}$	1	0	0	1	1	0
$\{3\}$	0	1	0	1	0	1
$\{4\}$	0	0	1	0	1	1

Theorem 2 *[14] For any $n \geq 4$, $\Gamma_{1,2}(n)$ is a binary 2-SFPC with size $N = n$ and length $L = \binom{n}{2}$.*

Proof. Let $\{i_1\}$, $\{i_2\}$, $\{i_3\}$ and $\{i_4\}$ be four distinct elements of $(n)_1$. Then $\{i_1\}_{\{i_1,i_2\}} = \{i_2\}_{\{i_1,i_2\}} = 1$ and $\{i_3\}_{\{i_1,i_2\}} = \{i_4\}_{\{i_1,i_2\}} = 0$. Thus $\{i_1,i_2\} \in (n)_2$ separates $\{\{i_1\},\{i_2\}\}$ and $\{\{i_3\},\{i_4\}\}$. Since $|\Gamma_{1,2}(n)| = n \geq 4$, it follows from Theorem 1 that $\Gamma_{1,2}(n)$ is a 2-SFPC. □

5 The ternary secure frameproof code $\Gamma_{2,2}(n)$

The code $\Gamma_{2,2}(n)$ is depicted by an $\binom{n}{2} \times \binom{n}{2}$ matrix where rows and columns are labelled by elements of the set $(n)_2$, which are subsets of (n) of size 2. The entry at row U and column V is $|U \cap V|$.

When $n = 4$, we have the following code $\Gamma_{2,2}(4)$

$\Gamma_{2,2}(4)$	$\{1,2\}$	$\{1,3\}$	$\{1,4\}$	$\{2,3\}$	$\{2,4\}$	$\{3,4\}$
$\{1,2\}$	2	1	1	1	1	0
$\{1,3\}$	1	2	1	1	0	1
$\{1,4\}$	1	1	2	0	1	1
$\{2,3\}$	1	1	0	2	1	1
$\{2,4\}$	1	0	1	1	2	1
$\{3,4\}$	0	1	1	1	1	2

Code $\Gamma_{2,2}(4)$ is not a 2-SFPC because the word $(1,1,1,1,1,1)$ can be constructed from both disjoint coalitions $\{\{1,2\},\{1,3\}\}$ and $\{\{1,4\},\{2,4\}\}$.

However, we will show that for any $n > 4$, $\Gamma_{2,2}(n)$ is a 2-SFPC.

Theorem 3 *For any $n > 4$, $\Gamma_{2,2}(n)$ is a ternary 2-SFPC with size $N = \binom{n}{2}$ and length $L = \binom{n}{2}$.*

Proof. We prove by contradiction. Suppose $\Gamma_{2,2}(n)$ is not a 2-SFPC for some $n > 4$. Since $N = \binom{n}{2} \geq 4$, it follows from Theorem 1 and Lemma 1 that there exist four distinct codewords S_1, S_2, S_3, S_4 such that for any position V

$$\pi_V(S_1, S_2) \cap \pi_V(S_3, S_4) \neq \emptyset$$

(note that $S_1, S_2, S_3, S_4, V \in (n)_2$.)

This is equivalent to

$$\{|V \cap S_1|, |V \cap S_2|\} \cap \{|V \cap S_3|, |V \cap S_4|\} \neq \emptyset, \quad \forall V \in (n)_2$$

Take $V = S_1$ we obtain

$$\{2, |S_1 \cap S_2|\} \cap \{|S_1 \cap S_3|, |S_1 \cap S_4|\} \neq \emptyset$$

Since for any $i \neq j$, $|S_i \cap S_j|$ is either equal to 0 or 1, we have

$$|S_1 \cap S_2| \in \{|S_1 \cap S_3|, |S_1 \cap S_4|\} \tag{3}$$

Similar arguments with $V = S_2, S_3, S_4$ give

$$|S_1 \cap S_2| \in \{|S_2 \cap S_3|, |S_2 \cap S_4|\} \tag{4}$$
$$|S_3 \cap S_4| \in \{|S_1 \cap S_3|, |S_2 \cap S_3|\} \tag{5}$$
$$|S_3 \cap S_4| \in \{|S_1 \cap S_4|, |S_2 \cap S_4|\} \tag{6}$$

To proceed with the proof, we will use graph. Each element of (n) is represented as a point. A set $S = \{x, y\} \in (n)_2$ is represented as an edge joining x and y. To make it clear, edges S_1 and S_2 will be drawn as unbroken lines and edges S_3, S_4 are broken lines.

We consider three cases

Case 1: $S_1 \cap S_2 = S_3 \cap S_4 = \emptyset$
Equations (3)-(6) become

$$0 \in \{|S_1 \cap S_3|, |S_1 \cap S_4|\}$$
$$0 \in \{|S_2 \cap S_3|, |S_2 \cap S_4|\}$$
$$0 \in \{|S_1 \cap S_3|, |S_2 \cap S_3|\}$$
$$0 \in \{|S_1 \cap S_4|, |S_2 \cap S_4|\}$$

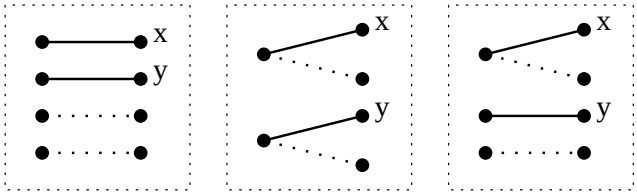

Figure 1: Case 1

Thus, in this case, each of S_1 and S_2 must be disjoint with either S_3 or S_4, and each of S_3, S_4 must be disjoint with either S_1 or S_2. Figure 1 shows all the possibilities.

Case 2: $S_1 \cap S_2 \neq \emptyset$, $S_3 \cap S_4 \neq \emptyset$
Equations (3)-(6) become

$$1 \in \{|S_1 \cap S_3|, |S_1 \cap S_4|\}$$
$$1 \in \{|S_2 \cap S_3|, |S_2 \cap S_4|\}$$
$$1 \in \{|S_1 \cap S_3|, |S_2 \cap S_3|\}$$
$$1 \in \{|S_1 \cap S_4|, |S_2 \cap S_4|\}$$

Thus, in this case, each of S_1 and S_2 must intersect with either S_3 or S_4 and each of S_3, S_4 must intersect with either S_1 or S_2. Figure 2 shows all the possibilities.

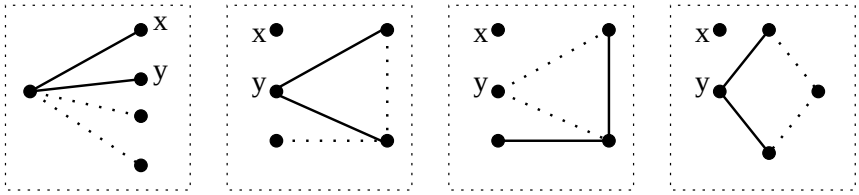

Figure 2: Case 2

Case 3: $S_1 \cap S_2 \neq \emptyset$, $S_3 \cap S_4 = \emptyset$
Equations (3)-(6) become

$$1 \in \{|S_1 \cap S_3|, |S_1 \cap S_4|\}$$
$$1 \in \{|S_2 \cap S_3|, |S_2 \cap S_4|\}$$
$$0 \in \{|S_1 \cap S_3|, |S_2 \cap S_3|\}$$
$$0 \in \{|S_1 \cap S_4|, |S_2 \cap S_4|\}$$

Thus, in this case, each of S_1 and S_2 must intersect with either S_3 or S_4 and each of S_3, S_4 must be disjoint with either S_1 or S_2. There is only one possibility as shown in Figure 3.

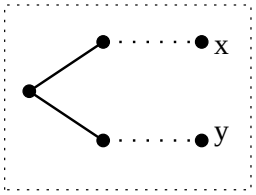

Figure 3: Case 3

It is easy to see in Figures 1, 2 and 3 that in all these cases if we choose $V = \{x, y\}$ then V separates the two coalitions $\{S_1, S_2\}$, $\{S_3, S_4\}$. This contradicts to the assumption that there does not exist any set $V \in (n)_2$ that can separate $\{S_1, S_2\}$ and $\{S_3, S_4\}$. And so, the theorem is proved. Note here that we have used the fact that $n > 4$ in the Figure 2. □

6 General codes $\Gamma_{t,r}(n)$ and $\Gamma_{t,\leq r}(n)$

We use the following theorem to prove the secure frameproof property of $\Gamma_{t,r}(n)$ and $\Gamma_{t,\leq r}(n)$

Theorem 4 *If S_1, S_2, S_3 and S_4 are arbitrary subsets of (n) such that*

$$S_i \not\subset S_j \text{ and } S_j \not\subset S_i \quad \text{for all } i \in \{1,2\},\ j \in \{3,4\}$$

then there must exist an elements $V \in (n)_{\leq 3}$ such that the following two sets

$$\{|V \cap S_1|, |V \cap S_2|\} \text{ and } \{|V \cap S_3|, |V \cap S_4|\}$$

are disjoint.

Proof. We prove by contradiction. Assume that there exists subsets S_1, S_2, S_3 and S_4 of (n) such that

$$S_i \not\subset S_j \text{ and } S_j \not\subset S_i \quad \text{for all } i \in \{1,2\},\ j \in \{3,4\} \tag{7}$$

and for any $V \in (n)_{\leq 3}$ we have

$$\begin{aligned} &(|V \cap S_1| - |V \cap S_3|)(|V \cap S_1| - |V \cap S_4|) \\ &(|V \cap S_2| - |V \cap S_3|)(|V \cap S_2| - |V \cap S_4|) \quad = \quad 0 \end{aligned} \tag{8}$$

We introduce a few notations.
If S is a subset of (n), we define

$$\begin{aligned} S\langle 1\rangle &= S \\ S\langle 0\rangle &= S^c = (n) \setminus S \end{aligned}$$

and for an element $z = (z_1, z_2, z_3, z_4) \in \{0,1\}^4$, we denote

$$S_z = S_1\langle z_1\rangle \cap S_2\langle z_2\rangle \cap S_3\langle z_3\rangle \cap S_4\langle z_4\rangle.$$

For instance, $S_{1010} = S_1 \cap S_2^c \cap S_3 \cap S_4^c$.

Let $s_z = |S_z|$, then the condition (7) is equivalent to

$$|S_1 \cap S_3^c| = \sum_{z_1=1, z_3=0} s_z = s_{1000} + s_{1001} + s_{1101} + s_{1100} \geq 1 \tag{9}$$

$$|S_1 \cap S_4^c| = \sum_{z_1=1, z_4=0} s_z = s_{1000} + s_{1010} + s_{1110} + s_{1100} \geq 1 \tag{10}$$

$$|S_2 \cap S_3^c| = \sum_{z_2=1, z_3=0} s_z = s_{0100} + s_{0101} + s_{1101} + s_{1100} \geq 1 \tag{11}$$

$$|S_2 \cap S_4^c| = \sum_{z_2=1, z_4=0} s_z = s_{0100} + s_{0110} + s_{1110} + s_{1100} \geq 1 \tag{12}$$

$$|S_3 \cap S_1^c| = \sum_{z_1=0, z_3=1} s_z = s_{0010} + s_{0110} + s_{0111} + s_{0011} \geq 1 \tag{13}$$

$$|S_4 \cap S_1^c| = \sum_{z_1=0, z_4=1} s_z = s_{0001} + s_{0101} + s_{0111} + s_{0011} \geq 1 \tag{14}$$

$$|S_3 \cap S_2^c| = \sum_{z_2=0, z_3=1} s_z = s_{0010} + s_{1010} + s_{1011} + s_{0011} \geq 1 \tag{15}$$

$$|S_4 \cap S_2^c| = \sum_{z_2=0, z_4=1} s_z = s_{0001} + s_{1001} + s_{1011} + s_{0011} \geq 1 \tag{16}$$

For any $0 \leq v_z \leq s_z$ such that $\sum v_z \leq 3$, let V be an arbitrary element of $(n)_{\leq 3}$ which contains v_z elements from the set S_z for each $z \in \{0,1\}^4$. We have

$$|V \cap S_i| - |V \cap S_j| = \sum_{z_i=1} v_z - \sum_{z_j=1} v_z = \sum_{z_i=1, z_j=0} v_z - \sum_{z_i=0, z_j=1} v_z$$

Hence, (8) becomes

$$\begin{aligned} &(v_{1000} + v_{1001} + v_{1101} + v_{1100} - v_{0010} - v_{0110} - v_{0111} - v_{0011}) \\ &(v_{1000} + v_{1010} + v_{1110} + v_{1100} - v_{0001} - v_{0101} - v_{0111} - v_{0011}) \\ &(v_{0100} + v_{0101} + v_{1101} + v_{1100} - v_{0010} - v_{1010} - v_{1011} - v_{0011}) \\ &(v_{0100} + v_{0110} + v_{1110} + v_{1100} - v_{0001} - v_{1001} - v_{1011} - v_{0011}) \\ &= 0 \end{aligned} \tag{17}$$

If $s_{1100} \geq 1$ then we can substitute into (17) with $v_{1100} = 1$ and all others $v_z = 0$, we obtain $1 = 0$, a contradiction. Therefore, $\boxed{s_{1100} = 0}$. Similarly, we have $\boxed{s_{0011} = 0}$.

If $s_{1010}, s_{1001} \geq 1$ then we can substitute into (17) with $v_{1010} = v_{1001} = 1$ and all others $v_z = 0$, we obtain $1 = 0$, a contradiction. Therefore, at least one of s_{1010}, s_{1001} must be equal to 0. And so $s_{1010}s_{1001} = 0$. Similar argument shows that $s_{1010}s_{0110} = s_{0101}s_{1001} = s_{0101}s_{0110} = s_{1000}s_{0100} = s_{1000}s_{1011} = s_{0111}s_{0100} = s_{0111}s_{1011} = s_{0010}s_{0001} = s_{0010}s_{1110} = s_{1101}s_{0001} = s_{1101}s_{1110} = 0$. Hence,

$$(s_{1010} + s_{0101})(s_{1001} + s_{0110}) = 0 \tag{18}$$
$$(s_{1000} + s_{0111})(s_{0100} + s_{1011}) = 0 \tag{19}$$
$$(s_{0010} + s_{1101})(s_{0001} + s_{1110}) = 0 \tag{20}$$

From (18) we have either $s_{1010} = s_{0101} = 0$ or $s_{1001} = s_{0110} = 0$. Without lost of generality we can assume that $\boxed{s_{1010} = s_{0101} = 0}$.

From (19) and (20) we consider four cases:

Case 1: $s_{1000} = s_{0111} = 0$ and $s_{0010} = s_{1101} = 0$
Equations (10), (11) and (13) become $|S_1 \cap S_4^c| = s_{1110} \geq 1$, $|S_2 \cap S_3^c| = s_{0100} \geq 1$ and $|S_3 \cap S_1^c| = s_{0110} \geq 1$. Substitute into (17) with $v_{1110} = v_{0100} = v_{0110} = 1$ and all others $v_z = 0$, we obtain $(-3) = 0$, a contradiction.

Case 2: $s_{1000} = s_{0111} = 0$ and $s_{0001} = s_{1110} = 0$
Equations (10) and (14) become $|S_1 \cap S_4^c| = 0 \geq 1$ and $|S_4 \cap S_1^c| = 0 \geq 1$. This is a contradiction.

Case 3: $s_{0100} = s_{1011} = 0$ and $s_{0010} = s_{1101} = 0$
Equations (11) and (15) become $|S_2 \cap S_3^c| = 0 \geq 1$ and $|S_3 \cap S_2^c| = 0 \geq 1$. This is a contradiction.

Case 4: $s_{0100} = s_{1011} = 0$ and $s_{0001} = s_{1110} = 0$
Equations (10), (11) and (16) become $|S_1 \cap S_4^c| = s_{1000} \geq 1$, $|S_2 \cap S_3^c| = s_{1101} \geq 1$ and $|S_4 \cap S_2^c| = s_{1001} \geq 1$. Substitute into (17) with $v_{1000} = v_{1101} = v_{1001} = 1$ and all others $v_z = 0$, we obtain $(-3) = 0$, a contradiction.

In all four cases we derive contradiction. This proves the theorem. □

From Theorem 4, the following theorem immediately follows

Theorem 5 *For any $0 < t < n$, the code $\Gamma_{t,\leq 3}(n)$ is a 2-SFPC.*

Proof. Let $\{S_1, S_2\}$, $\{S_3, S_4\}$ be two disjoint coalitions where $S_i \in (n)_t$. Since

$$S_i \not\subset S_j \text{ and } S_j \not\subset S_i \quad \text{for all } i \in \{1,2\},\ j \in \{3,4\}$$

it follows from Theorem 4 that $\exists V \in (n)_{\le 3}$ such that the two sets $\{|V \cap S_1|, |V \cap S_2|\}$ and $\{|V \cap S_3|, |V \cap S_4|\}$ are disjoint. Thus, V separates $\{S_1, S_2\}$ and $\{S_3, S_4\}$, and so $\Gamma_{t,\le 3}(n)$ is a 2-SFPC. □

Corollary 1 *For any $t > 0$, $r \ge 3$, and $n \ge 4t + r$, the code $\Gamma_{t,r}(n)$ is a 2-SFPC.*

Proof. For any four distinct elements S_1, S_2, S_3, S_4 of $(n)_t$, by Theorem 4, there exists $V \in (n)_{\le 3}$ such that the two sets $\{|V \cap S_1|, |V \cap S_2|\}$ and $\{|V \cap S_3|, |V \cap S_4|\}$ are disjoint. Since $n \ge 4t + r = |S_1| + |S_2| + |S_3| + |S_4| + r$, we can add more elements from the set $(n) \setminus (S_1 \cup S_2 \cup S_3 \cup S_4)$ to V to obtain a set $V' \in (n)_r$. We have $V \cap S_i = V' \cap S_i$, and thus, the two sets $\{|V' \cap S_1|, |V' \cap S_2|\}$ and $\{|V' \cap S_3|, |V' \cap S_4|\}$ are disjoint. This proves that the code $\Gamma_{t,r}(n)$ is a 2-SFPC. □

The codes $\Gamma_{t,\le 3}(n)$ in Theorem 5 has size $N = \binom{n}{t}$ and length $L = \binom{n}{1} + \binom{n}{2} + \binom{n}{3} = \frac{1}{6}n(n^2+5)$. To maximize the code size, choose $t = [n/2]$.

For $0 < \mu, \lambda < 1$, $\mu + \lambda = 1$, we have (Roman [10, page 445])

$$\frac{1}{\sqrt{8n\lambda\mu}}\lambda^{-\lambda n}\mu^{-\mu n} < \binom{n}{\lambda n} < \frac{1}{\sqrt{2\pi n\lambda\mu}}\lambda^{-\lambda n}\mu^{-\mu n}$$

Take $\mu = \lambda = 1/2$, then

$$\frac{1}{\sqrt{2n}}2^n < \binom{n}{\frac{1}{2}n} < \frac{1}{\sqrt{\frac{\pi}{2}n}}2^n$$

With $t = [n/2]$, we have $N > \frac{1}{\sqrt{2n}}2^n$, and therefore, $L = o((\log N)^3)$.

7 Binary codes $\Gamma^*_{t,r}(n)$ and $\Gamma^*_{t,\le r}(n)$

Similar to codes $\Gamma_{t,r}(n)$ and $\Gamma_{t,\le r}(n)$, the binary codes $\Gamma^*_{t,r}(n)$ and $\Gamma^*_{t,\le r}(n)$ can be proved to be 2-secure frameproof.

Theorem 6 *If S_1, S_2, S_3 and S_4 are arbitrary subsets of (n) such that*

$$S_i \not\subset S_j \text{ and } S_j \not\subset S_i \quad \text{for all } i \in \{1,2\}, j \in \{3,4\}$$

then there must exist an elements $V \in (n)_{\le 3}$ such that the following two sets

$$\{|V \cap S_1| \bmod 2, |V \cap S_2| \bmod 2\} \text{ and } \{|V \cap S_3| \bmod 2, |V \cap S_4| \bmod 2\}$$

are disjoint.

Proof. Similar to the proof of Theorem 4. Note that the Equation (17) becomes

$$\begin{aligned}
&(v_{1000} + v_{1001} + v_{1100} + v_{1101} - v_{0010} - v_{0011} - v_{0110} - v_{0111})\\
&(v_{1000} + v_{1010} + v_{1100} + v_{1110} - v_{0001} - v_{0011} - v_{0101} - v_{0111})\\
&(v_{0100} + v_{0101} + v_{1100} + v_{1101} - v_{0010} - v_{0011} - v_{1010} - v_{1011})\\
&(v_{0100} + v_{0110} + v_{1100} + v_{1110} - v_{0001} - v_{0011} - v_{1001} - v_{1011})\\
&= 0 \pmod 2
\end{aligned} \tag{21}$$

□

Corollary 2 *For any $0 < t < n$, the binary code $\Gamma^*_{t,\leq 3}(n)$ is a 2-SFPC with size $N = \binom{n}{t}$ and length $L = \frac{1}{6}n(n^2+5)$. When $t = [n/2]$ then $L = o((\log N)^3)$.*

Corollary 3 *For any $t > 0$, $r \geq 3$, and $n \geq 4t + r$, the binary code $\Gamma^*_{t,r}(n)$ is a 2-SFPC.*

8 Conclusion

Below is the summary of codes in our paper:

- Code $\Gamma_{1,2}(n)$, $n \geq 4$:
$$N = n,\ L = \binom{n}{2},\ q = 2$$
- Code $\Gamma_{2,2}(n)$, $n > 4$:
$$N = \binom{n}{2},\ L = \binom{n}{2},\ q = 3$$
- Code $\Gamma_{t,\geq 3}(n)$, $t < n$:
$$N = \binom{n}{t},\ L = \frac{1}{6}n(n^2+5),\ q = 4$$
- Code $\Gamma_{t,r}(n)$, $r \geq 3$, $n \geq 4t + r$:
$$N = \binom{n}{t},\ L = \binom{n}{r},\ q = \min(t, r) + 1$$
- Code $\Gamma^*_{t,\geq 3}(n)$, $t < n$:
$$N = \binom{n}{t},\ L = \frac{1}{6}n(n^2+5),\ q = 2$$
- Code $\Gamma^*_{t,r}(n)$, $r \geq 3$, $n \geq 4t + r$:
$$N = \binom{n}{t},\ L = \binom{n}{r},\ q = 2$$
- Choosing $t = [n/2]$, codes $\Gamma_{[n/2],\geq 3}(n)$ and $\Gamma^*_{[n/2],\geq 3}(n)$ have size N exponentially large compared to the length L:
$$N = \binom{n}{[n/2]} > \frac{1}{\sqrt{2n}}2^n,\ L = \frac{1}{6}n(n^2+5) = o((\log N)^3)$$

Our codes have much shorter length compared to the following Encheva et al [5] explicit 2-SFPC codes:

- Using Hadamard matrix: $N = 2^n$, $L = 2^n$, $q = 2$
constraint: Hadamard matrix obtained from a Sylvester type matrix.

- Using equi-distance code: $N = \frac{q^2\mu - 1}{q-1}$, $L = q^2\mu \approx (N \times q)$
  constraint: $\exists$ an affine design with the following parameters

$$v = q^2\mu,\ k = q\mu,\ \lambda = \frac{q\mu - 1}{q-1},\ r = \frac{q^2\mu - 1}{q-1},\ b = \frac{q^3\mu - q}{q-1}$$

- Using Mersenne prime and m-sequence: $N = p$, $L = p$, $q = 2$
  constraint: p is a Mersenne prime.

We note here that, using Nanya et al [9] recursive techniques, Encheva et al [5] showed that it is possible to construct codes with shorter length from existing binary codes. We have developed a general recursive technique in [15] which can be applied to q-ary code for any value of q. We can use this recursive technique to make our code even shorter.

Acknowledgment

This work is in part supported by Motorola Australian Research Centre under Strategic Partnership with Industry Program of Australian Research Council, grant number C00002687.

References

[1] N. Alon, E. Fischer and M. Szegedy, Parent-Identifying Codes, *Journal of Combinatorial Theory,* Series A, **95** (2001), 349–359.

[2] D. Boneh and J. Shaw, Collusion-Secure Fingerprinting for Digital Data, *Proceeding of CRYPTO'95,* LNCS **963** (1995), 453–465.

[3] D. Boneh and J. Shaw, Collusion-Secure Fingerprinting for Digital Data, *IEEE Transactions on Information Theory,* **44**, No. 5 (1998), 1897–1905.

[4] B. Chor, A. Fiat and M. Naor, Tracing Traitors, *Proceeding of CRYPTO'94,* LNCS **839** (1994), 257–270.

[5] S. Encheva and G. Cohen, Partially Identifying Codes for Copyright Protection, *Proceeding of 14th AAECC,* LNCS **2227** (2001), 260–267.

[6] A. Fiat and T. Tassa, Dynamic Traitor Tracing, *Proceeding of CRYPTO'99,* LNCS **1666** (1999), 354–371.

[7] A.D. Friedman, R.L. Graham and J.D.Ullman, Universal Single Transition Time Asynchronous State Assignments, *IEEE Transactions on Computers,* **C-18**, No. 6 (1969), 541–547.

[8] H.D. Hollmann, J.H. van Lint, J.P. Linnartz and L.M. Tolhuizen, On Codes with the Idetifiable Parent Property, *Journal of Combinatorial Theory,* Series A, **82** (1998), 121–133.

[9] T. Nanya and Y. Tohma, On Universal Transition Time Asynchronous State Assignments, *IEEE Transactions on Computers,* **C-27**, No. 8 (1978), 781–782.

[10] S. Roman, *Coding and Information Theory,* Springer-Verlag (1992).

[11] J. Staddon, D.R. Stinson and R. Wei, Combinatorial Properties of Frameproof and Traceability Codes, *IEEE Transaction on Information Theory,* **47**, No. 3 (2001), 1042–1049.

[12] D.R. Stinson and R. Wei, Combinatorial Properties and Constructions of Traceability Schemes and Frameproof Codes, *SIAM Journal of Discrete Mathematics,* **11**, No. 1 (1998), 41–53.

[13] 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**, No. 2 (2000), 595–617.

[14] V.D. Tô, R. Safavi-Naini and Y. Wang, A 2-Secure Code With Efficient Tracing Algorithm, *Proceeding of IndoCrypt'02,* LNCS **2551** (2002), 149–162.

[15] D. Tonien and R. Safavi-Naini, *Recursive Constructions of Secure Codes and Hash Families Using Difference Function Families,* Cryptology ePrint Archive, Report 2005/184. URL: `http://eprint.iacr.org/`.