Frameproof and IPP Codes

Palash Sarkar*and Douglas Robert Stinson
Centre for Applied Cryptographic Research
Department of Combinatorics and Optimization
University of Waterloo
200 University Avenue West
Waterloo, Ontario
Canada N2L 3G1
e-mail: {psarkar,dstinson}@cacr.math.uwaterloo.ca

Abstract

Frameproof and identifying parent property codes and their relationship to hash families and error correcting codes are studied. A new upper bound on the size of such codes is presented. A generalization of identifying parent property is introduced and studied in terms of a new class of hash families - the strong separating hash families. Asymptotic consequences of some recursive construction techniques are described.

1 Introduction

Codes providing certain forms of traceability have been studied for their applications to protection of intellectual property rights. The weakest form of such codes provides protection against the possibility of a user being framed by a coalition of users. Such codes are called frameproof codes and were introduced by Boneh and Shaw [4]. A stronger notion introduced in [10] and called secure frameproof code is protection against two disjoint coalition being able to produce a common descendant. The strongest version studied in this paper is called the identifying parent property (IPP) and requires the condition that if a set of coalitions is able to produce a common descendant then the coalitions must have a codeword in common (the identifying parent). This concept was introduced in [7].

The class of codes with traceability property have been studied by several authors. A systematic combinatorial study have been carried out in [8]. In a recent work, Blackburn [2] have provided an upper bound on the cardinality of frameproof codes. Also [2] provides the counterexample of

*On leave from the Indian Statistical Institute, Calcutta, INDIA. e-mail: palash@isical.ac.in

length 5, 3-frameproof code to show that the coefficient of the leading term in the upper bound is not tight. We build on this counterexample to show that in most cases the coefficient of the upper bound is not tight.

A sufficient condition for the existence of IPP codes has been obtained in terms of perfect hash families in [8]. This raised the question of existence of IPP codes for certain values of the parameters. This question was resolved in [3] by the probabilistic method and using a new family of hash functions called the partially hashing family. We introduce a natural generalization of perfect and separating hash family called the strong separating hash family which turns out to be equivalent to the class of partially hashing property. We introduce a generalization of IPP codes and obtain a hierarchy of codes between the frameproof and IPP codes. The existence of these codes are studied in terms of the strong separating hash family and error correcting codes.

Frameproof and IPP codes can be constructed from suitable hash families which in turn can be construted from error correcting codes. We explore this theme in conjunction with a recursive composition type construction used in [7, 11]. It turns out that for each size of the alphabet and strength of the code it is possible to construct these classes of codes having a fixed relation between the length and cardinality.

2 Hash Functions

Let $\mathcal{H}$ be a family of N functions from $\{1, \ldots, n\}$ to the alphabet Q of cardinality q. The family $\mathcal{H}$ is called an (N, n, q) hash family. Given an (N, n, q) hash family we can obtain an $N \times n$ matrix by enumerating the values of all the functions. Sometimes it is easier to consider hash families as such matrices.

Definition 2.1 *We consider the following kinds of hash families.*

1. **Perfect hash family :** *An (N, n, q) hash family $\mathcal{H}$ is called an (N, n, q, w) perfect hash family ((N, n, q, w)-PHF) if for any subset $S \subseteq \{1, \ldots, N\}$ of cardinality w, there is a function $f \in \mathcal{H}$ such that f is injective on S.*

2. **Separating Hash Family :** *An (N, n, q) hash family $\mathcal{H}$ is called an (N, n, q, w_1, w_2) separating hash family ((N, n, q, w_1, w_2)-SHF) if for any two disjoint subsets A, B of $\{1, \ldots, n\}$ with $|A| = w_1$ and $|B| = w_2$, there is a function f in $\mathcal{H}$ such that $f(A)$ and $f(B)$ are also disjoint.*

3. **Strong Separating Hash Family :** *An (N, n, q) hash family $\mathcal{H}$ is called an (N, n, q, w_1, w_2) strong separating hash family ((N, n, q, w_1, w_2)-SSHF) if for any two disjoint subsets $A, B \subset \{1, \ldots, n\}$ with $|A| = w_1$ and $|B| = w_2$, there is a function $f \in \mathcal{H}$ such that f is injective on A and $f(A) \cap f(B) = \emptyset$.*

4. **Partially Hashing Family :** *[3] An (N, n, q) hash family $\mathcal{H}$ is called an (N, n, q, t, u) partially hashing family ((N, n, q, t, u)-PAHF) if for any two subsets $T, U \subset \{1, \ldots, n\}$ such that $T \subset U$, $|T| = t$, $|U| = u$, there is a function f in $\mathcal{H}$ such that for any $x \in T$ and any $y \in U$, if $y \neq x$, we have $f(x) \neq f(y)$.*

See [5] for a survey on perfect hash families. The concept of separating hash families was introduced in [9] and that of partially hashing families was introduced in [3].

We now relate strong separating hash and partially hashing properties to the other hashing properties. The following is immediate from the definition of strong separating hash property.

Theorem 2.1 *Let $\mathcal{H}$ be an (N, n, q) hash family.*

1. *If $\mathcal{H}$ is an (N, n, q, w_1, w_2)-SSHF, then it is simultaneously an (N, n, q, w_1, w_2)-SHF and an (N, n, q, w_1)-PHF.*
2. *If $\mathcal{H}$ is an $(N, n, q, w_1 + w_2)$-PHF, then it is an (N, n, q, w_1, w_2)-SSHF.*
3. *If $\mathcal{H}$ is an (N, n, q, w_1, w_2)-SSHF, then it is an (N, n, q, t_1, t_2)-SSHF for any $t_1 \leq w_1$ and $t_2 \leq w_2$.*

It is not difficult to see that the converse of Theorem 2.1(1) does not in general hold. We now show the equivalence of strong separating hash and partially hashing property.

Theorem 2.2 *A hash family $\mathcal{H}$ is an (N, n, q, w_1, w_2)-SSHF iff it is an $(N, n, q, w_1, w_1 + w_2)$-PAHF.*

Proof : Suppose $\mathcal{H}$ is an (N, n, q, w_1, w_2)-SSHF. We will show that $\mathcal{H}$ is an $(N, n, q, w_1, w_1 + w_2)$-PAHF. Let $A \subset B \subset \{1, \ldots, n\}$, with $|A| = w_1$, $|B| = w_1 + w_2$. Let $C = B - A$. Then $|C| = w_2$ and $C \cap A = \emptyset$. By the strong separating hash property there is an f such that f is injective on A and $f(A) \cap f(C) = \emptyset$. Let $x \in A$ and $y \in B$ with $x \neq y$. There are two cases to consider.

1. $y \in A$. Since f is injective on A, $f(x) \neq f(y)$.
2. $y \in B - A$. Since $C = B - A$ and $f(A) \cap f(C) = \emptyset$, $f(x) \neq f(y)$.

Thus in both cases $f(x) \neq f(y)$ and hence $\mathcal{H}$ is an $(N, n, q, w_1, w_1 + w_2)$-PAHF.

Conversely, suppose $\mathcal{H}$ is an $(N, n, q, w_1, w_1 + w_2)$-PAHF. We show $\mathcal{H}$ is an (N, n, q, w_1, w_2)-SSHF. Let $A, B \subset \{1, \ldots, n\}$, with $A \cap B = \emptyset$, $|A| = w_1$, $|B| = w_2$. Let $C = A \cup B$. Then $|C| = w_1 + w_2$ and $A \subset C$. Thus by the partially hashing property there is a function f such that if $x \in A$ and $y \in C$ and $x \neq y$, then $f(x) \neq f(y)$. It is not difficult to argue that this f must be injective on A and $f(A) \cap f(B) = \emptyset$. Hence $\mathcal{H}$ is an (N, n, q, w_1, w_2)-SSHF. ■

We present a sufficient condition for the existence of separating hash families based on error correcting codes. An (n, N, q, D) error correcting code is a set $C \subset Q^N$ of n codewords where $|Q| = q$ and the distance between any pair of codewords is at least D. We denote by $hd(x, y)$ the Hamming distance between two codewords x, y.

The ideas in the following theorem are present in [6, 1, 11].

Theorem 2.3 *Let C be an (n, N, q, D) error correcting code with*

$$D > N \left(1 - \frac{1}{\binom{w_1}{2} + w_1 w_2} \right). \tag{1}$$

Then $\mathcal{H}(C)$ is an (N, n, q, w_1, w_2)-SSHF.

Proof : Let A, B be disjoint subsets of $\{1, \ldots, n\}$, with $|A| = w_1$ and $|B| = w_2$. Let

$$S_1 = \sum_{\{x,y\} \subseteq A} hd(x, y),\ S_2 = \sum_{x \in A, y \in B} hd(x, y) \text{ and } S = S_1 + S_2.$$

Clearly $S_1 \geq \binom{w_1}{2} D$, $S_2 \geq w_1 w_2 D$ and

$$S \geq D(\binom{w_1}{2} + w_1 w_2). \tag{2}$$

Suppose that $\mathcal{H}(C)$ is not (w_1, w_2)-SSHF. Then for each $j \in \{1, \ldots, N\}$ the maximum contribution of the jth column to S is $(\binom{w_1}{2} + w_1 w_2 - 1)$. Since there are N such indices, we must have $S \leq N(\binom{w_1}{2} + w_1 w_2 - 1)$. This contradicts the combination of (1) and (2). ∎

It is now easy to see that all the construction techniques (direct and recursive) used in [11] for the construction of perfect and separating hash functions can be applied to construct strong separating hash functions. In particular, we get the following result using the recursive construction technique of [11].

Theorem 2.4 *For positive integers q, w_1 and w_2, there exists an infinite class of (N, n, q, w_1, w_2)-SSHF for which N is $O((w_1(w_1 + w_2))^{\log^* n} \log n)$.*

3 Traceability Codes

Let Q be an alphabet of size q and C be a subset of Q^N of cardinality n. We will call C an (n, N, q) code and the elements of C to be codewords or simply words. It is sometimes convenient to consider $C = (c_{ij})$ to be an $n \times N$ matrix with elements from the set Q. Let $\mathsf{elem}_C(j)$ be the set of all elements that occur in the jth column of C, i.e., $\mathsf{elem}_C(j) = \{a \in Q : a = c_{ij} \text{ for some } 1 \leq i \leq n\}$. When the code is understood we will write $\mathsf{elem}(i)$ instead of $\mathsf{elem}_C(i)$. Define

$$\mathsf{desc}(C) = \mathsf{elem}_C(1) \times \ldots \times \mathsf{elem}_C(N).$$

The set $\mathsf{desc}(C)$ is called the descendant code of C. Clearly $C \subseteq \mathsf{desc}(C)$. We make a few remarks on the connection between C and $\mathsf{desc}(C)$.

The connection of codes and hash families follows from the fact that if C is an (n, N, q) code then C^T can be considered to be the matrix obtained from an (N, n, q) hash family and conversely. Here C^T denotes the transpose of the matrix C. Thus given an (n, N, q) code C, we think of the columns of the matrix C as the functions of the associated hash family. This makes it easier to state many of the results below. Following [8], we will denote by $\mathcal{H}(C)$ the hash family obtained from the code C.

Lemma 3.1 *Let $x \in Q^N$ and C be a minimal (n, N, q) code over Q such that $x \in \mathsf{desc}(C) - C$. Then $\mathcal{H}(C)$ is an $(N, n, q, n-1, 1)$-SHF. Conversely, if C is an (n, N, q) code such that $\mathcal{H}(C)$ is a minimal $(N, n, q, n-1, 1)$-SHF, then there is an $x \in Q^N$ such that $x \in \mathsf{desc}(C) - C$.*

Proof : Suppose $x \in \mathsf{desc}(C) - C$ and C is a minimal code such that $x \in \mathsf{desc}(C)$. Since C is minimal if we remove any row from C we have that $x \notin \mathsf{desc}(C)$. This means that given any row c_i of C there is a column j such that $c_{ij} = x_j$ and $x_j \notin \mathsf{elem}(j) - \{c_{ij}\}$. Thus the function corresponding to the jth column separates the set $\{i\}$ from the set $\{1, \ldots, n\} - \{i\}$. Thus $\mathcal{H}(C)$ is an $(N, n, q, n-1, 1)$-SHF.

Conversely, if $\mathcal{H}(C)$ is a minimal $(N, n, q, n-1, 1)$-SHF, we construct an $x \in \mathsf{desc}(C) - C$ as follows. Since $\mathcal{H}(C)$ is minimal if we drop any column j of C, then $\mathcal{H}(C)$ is no longer an $(N, n, q, n-1, 1)$-SHF. This means that there is a row c_i such that $c_{ij} \neq c_{kj}$ for any $k \in \{1, \ldots, i-1, i+1, \ldots, n\}$. We put $x_j = c_{ij}$. Clearly $x \in \mathsf{desc}(C)$ and by construction x cannot be equal to any row c_i. ■

Suppose C is an (n, N, q) code such that $|\mathsf{elem}(j)| = q$ for all $1 \leq j \leq N$. Then $C = \mathsf{desc}(C)$ iff $C = Q^N$. Here we note that if we consider the codewords to be one way infinite strings over Q (i.e., elements of $Q^{\mathbb{N}}$) and C is an infinite subset of $Q^{\mathbb{N}}$, then C is a proper subset of $\mathsf{desc}(C)$, since by diagonalization it is possible to construct a string in $\mathsf{desc}(C)$ which is not in C.

Let C be an (n, N, q) code. We will call a subset D of C to be a coalition and if a codeword x is in $\mathsf{desc}(D)$, then we say that D produces x. We define $\mathsf{desc}_w(C)$, the w-descendant code, in the following manner.

$$\mathsf{desc}_w(C) = \bigcup_{C_0 \subset C, |C_0| \leq w} \mathsf{desc}(C_0).$$

The set $\mathsf{desc}_w(C)$ consists of all codewords which could have been produced by subsets of C of size atmost w. We next define the important classes of codes.

Definition 3.1 *Let C be an (n, N, q) code. We now define certain traceability properties such a code might possess. These are the proerties we will consider in this paper. A stronger notion of traceability exists (see for example [8]).*

1. **Frameproof :** *We say that C is an (n, N, q, w)-FP code if the following condition holds. For any subset C_0 of C of cardinality atmost w, if $x \in \mathsf{desc}(C_0) \cap C$ then $x \in C_0$.*

2. **Secure Frameproof :** *We say that C is (n, N, q, w)-SFP code if the following condition holds. For any two subsets C_0, C_1 of C of cardinality atmost w, we have that $\mathsf{desc}(C_0) \cap \mathsf{desc} C_1 \neq \emptyset$ implies $C_0 \cap C_1 \neq \emptyset$.*

3. **Identifying Parent Property :** *We say that C is an (n, N, q, w)-IPP code if the following condition holds. Let $\{C_1, \ldots, C_\alpha\}$ be a family of subsets of C where each C_i is of cardinality atmost w. Then we must have that*

$$\bigcap_{1 \leq i \leq \alpha} \mathsf{desc}(C_i) \neq \emptyset \textit{ implies } \bigcap_{1 \leq i \leq \alpha} C_i \neq \emptyset.$$

The relationships among these classes of codes and their relationships to other combinatorial structures such as PHF, SHF and cover free families have been studied in [8]. Here we briefly mention some of these relations. We use the notation $P_1 \leq P_2$ to denote the fact that property P_2 implies property P_1.

Proposition 3.1 *[8] The following relationships hold for any (n, N, q) code.*

$$(n, N, q, w)\text{-}FP \leq (n, N, q, w)\text{-}SFP \leq (n, N, q, w)\text{-}IPP.$$

Theorem 3.1 *[9]*

1. *A code C is an (n, N, q, w)-FP code iff $\mathcal{H}(C)$ is an $(N, n, q, w, 1)$-SHF.*

2. *A code C is an (n, N, q, w)-SFP code iff $\mathcal{H}(C)$ is an (N, n, q, w, w)-SFP.*

Theorem 3.2 *[7, Lemma 1] A code C is an $(n, N, q, 2)$-IPP iff $\mathcal{H}(C)$ is both an $(N, n, q, 3)$-PHF and an $(N, n, q, 2, 2)$-SHF.*

A characterization of $(n, N, q, 3)$-IPP codes appear in [3].

4 Frameproof Codes

In [8], it was shown that for an (n, N, q, w)-FP code $n \leq w(q^{\lceil \frac{N}{w} \rceil} - 1)$. This was improved in [2] where it has been shown that

$$n \leq \max\{q^{\lceil \frac{N}{w} \rceil}, r(q^{\lceil \frac{N}{w} \rceil} - 1) + (w - r)(q^{\lfloor \frac{N}{w} \rfloor} - 1)\}, \tag{3}$$

where r is the unique integer in $\{1, \ldots, w\}$ such that $r \equiv N \bmod w$. Also in [2], an example of an $(n, 5, q, 3)$-FP code was presented where $n \leq \frac{5}{3}q^2 + 3q$. This example shows that the coefficient of the leading term of (3) is not always tight. Here we build on the counterexample of [2] to provide an upper bound which shows that the coefficient of the leading term of (3) is in most cases not tight.

Following [2], we define

$$U_S = \{x \in C : \text{ there exists no } y \in C - \{x\} \text{ such that } x_i = y_i \text{ for all } i \in S\}.$$

The following crucial result has been proved in [2, 8].

Lemma 4.1 *Let C be an (n, N, q, w)-FP code and $S_1, \ldots, S_w$ is a partition of $\{1, \ldots, N\}$. Then*

$$C = U_{S_1} \cup \ldots \cup U_{S_w}.$$

We require some preliminary results to prove the main theorem. These results are generalizations of the ideas present in [2].

In the following discussion we will fix an (n, N, q, w)-FP code C, where we write $N = wt + r$ with r to be the unique integer in the range $\{1, \ldots, w\}$. Further we will only consider codes with r in the range $1 < r < w$. Let $L = \{1, \ldots, N\}$. Let T_1, T_2 be subsets of L satisfying

1. $|T_1| = |T_2| = t + 1$,

2. $T_1 \cap T_2 = \emptyset$,

3. $|U_{T_1} \cap U_{T_2}|$ is the maximum among all pairs of subsets T_1, T_2 satisfying 1 and 2 above.

We put

$$|U_{T_1} \cap U_{T_2}| = kq^{t+1}. \tag{4}$$

Clearly $0 \leq k \leq 1$.

Lemma 4.2 *Let S_1, S_2 be subsets of $\{1, \ldots, N\}$ satisfying*

1. $|S_1|$=$|S_2|$=*t+1,*
2. $|S_1 \cap S_2|$=*1.*

Then $|U_{S_2} - U_{S_1}| \leq (r-1)kq^{t+1} + (w-r)q^t$.

Proof : Let $L' = L - (S_1 \cup S_2)$. Then $|L'| = (w-2)t + r - 1$. We form a partition of L' into $S_3, \ldots, S_w$, where $(r-1)$ of these sets have cardinality $(t+1)$ and the other $(w-r-1)$ sets have cardinality t. Clearly $S_1, S_2 - S_1, S_3, \ldots, S_w$ is a partition of L and hence by Lemma 4.1, we have

$$C = U_{S_1} \cup U_{S_2 - S_1} \cup U_{S_3} \cup \ldots \cup U_{S_w}.$$

Let x be in $U_{S_2} - U_{S_1}$. Since x is in U_{S_2} it must be in C and since it is not in U_{S_1} it must be either in $U_{S_2-S_1}$ or in one of $U_{S_i} \cap U_{S_2}$, $3 \leq i \leq w$. Hence we have

$$U_{S_2} - U_{S_1} \subseteq U_{S_2-S_1} \cup (U_{S_2} \cap U_{S_3}) \cup \ldots (U_{S_2} \cap U_{S_w}).$$

Clearly $|S_2 - S_1| = t$ and hence $|U_{S_2-S_1}| \leq q^t$. For $3 \leq i \leq w$, if S_i is one of the $(r-1)$ sets of cardinality $t+1$, using (4), we must have $|U_{S_2} \cap U_{S_i}| \leq kq^{t+1}$. If S_i is one of the $(w-r-1)$ sets of cardinality t, then $|U_{S_2} \cap U_{S_i}| \leq q^t$. Hence we get

$$\begin{aligned} |U_{S_2} - U_{S_1}| &\leq |U_{S_2-S_1}| + |U_{S_2} \cap U_{S_3}| + \ldots + |U_{S_2} \cap U_{S_w}| \\ &\leq (r-1)kq^{t+1} + (w-r)q^t. \end{aligned}$$

This proves the lemma. ■

Corollary 4.1 *Let T_1, T_2 be subsets of L such that $|T_1| = |T_2| = t+1$ and $T_1 \cap T_2 = \emptyset$. Then* $|U_{T_2} - U_{T_1}| \leq 2(r-1)kq^{t+1} + 2(w-r)q^t$.

Proof : Let V be a subset of L such that $|V| = t+1$ and $|V \cap T_1| = |V \cap T_2| = 1$. From the conditions on the parameters w, r, N it is not difficult to check that such a V exists. We apply Lemma 4.2 with T_1 and V to get

$$|U_V - U_{T_1}| \leq (r-1)kq^{t+1} + (w-r)q^t. \tag{5}$$

A second application of Lemma 4.2 with V, T_2 gives

$$|U_{T_2} - U_V| \leq (r-1)kq^{t+1} + (w-r)q^t. \tag{6}$$

Combining (5) and (6), we get

$$|U_{T_2} - U_{T_1}| \leq 2(r-1)kq^{t+1} + 2(w-r)q^t. \qquad \blacksquare$$

Now we are in a position to prove the main result.

Theorem 4.1 *Let C be an (n, N, q, w)-FP code where $N = wt + r$ and r is the unique integer in the set $\{1, \ldots, w\}$ such that $N \equiv r \bmod w$. If $1 < r < w$, then*

$$n \leq \left(r - \frac{\lfloor \frac{r}{2} \rfloor}{1 + 2\lfloor \frac{r}{2} \rfloor (r-1)}\right) q^{t+1} + (w-r)\left(2 + \left\lfloor \frac{r}{2} \right\rfloor\right) q^t.$$

Proof : Let $T_1, \ldots, T_w$ be a partition of $L = \{1, \ldots, N\}$, such that $T_1, \ldots, T_r$ are of cardinalities $(t+1)$ each and $T_{r+1}, \ldots, T_w$ are of cardinalities t each. Further we choose T_1, T_2 such that $|U_{T_1} \cap U_{T_2}| = kq^{t+1}$. Let $r_1 = \lfloor \frac{r}{2} \rfloor$ and assume $U_{T_0} = \emptyset$. We compute $|C|$ as follows.

$$\begin{aligned} |C| &= |U_{T_1} \cup U_{T_2} \cup \ldots \cup U_{T_w}| \\ &\leq (|U_{T_1}| + |U_{T_2}| - |U_{T_1} \cap U_{T_2}|) + |U_{T_3}| + \ldots + |U_{T_w}| \end{aligned}$$

Using (4), we get the following bound on $|C|$.

$$|C| \leq (r-k)q^{t+1} + (w-r)q^t. \tag{7}$$

Computing $|C|$ in a different way we get the following.

$$\begin{aligned} |C| &= |U_{T_1} \cup U_{T_2} \cup \ldots \cup U_{T_w}| \\ &\leq (|U_{T_1} \cup U_{T_2}|) + \ldots + (|U_{T_{2r_1-1}} \cup U_{T_{2r_1}}|) + |U_{T_{r-2r_1}}| + |U_{T_{r+1}}| + \ldots + |U_{T_w}| \\ &= (|U_{T_1}| + |U_{T_2} - U_{T_1}|) + \ldots + (|U_{T_{2r_1-1}}| + |U_{T_{2r_1}} - U_{T_{2r_1-1}}|) \\ &\quad + |U_{T_{r-2r_1}}| + |U_{T_{r+1}}| + \ldots + |U_{T_w}| \end{aligned}$$

For $1 \leq i \leq r_1$, we have using Corollary 4.1,

$$|U_{T_{2i}} - U_{T_{2i-1}}| \leq 2(r-1)kq^{t+1} + (w-r)q^t.$$

We use this to bound $|C|$ in another way as follows.

$$|C| \leq (r - r_1 + 2kr_1(r-1))q^{t+1} + (w-r)(2+r_1)q^t. \tag{8}$$

Combining (7) and (8) and eliminating k we get,

$$|C| \leq \left(r - \frac{r_1}{1 + 2r_1(r-1)}\right) q^{t+1} + (w-r)(2+r_1)q^t. \qquad \blacksquare$$

We briefly consider the construction problem for frameproof codes. The first construction is simple and doubles both the alphabet size and cardinality of the code while maintaining the parameters w and N constant.

Union Construction : Let C be an (n, N, q, w)-FP code on alphabet $Q = \{a_1, \ldots, a_q\}$. Define $Q' = \{a'_1, \ldots, a'_q\}$. This ensures that $Q \cap Q' = \emptyset$. Let C' be an (n, N, q, w)-FP code on alphabet Q'. Clearly C' can be obtained from C by changing each element $c_{i,j}$ of C to $c'_{i,j}$. Let $D = C \cup C'$. Then D is an $(2n, N, 2q, w)$-FP code on alphabet $Q \cup Q'$.

Proof : Clearly $|D| = 2n$ and the alphabet size is $2q$, so it remains to prove the w-FP property only. This is easy to see, since if x is in D it must be in exactly one of C or C'. If x is in C (respectively C') then x cannot be framed by any coalition of C' (respectively C). Moreover, since C and C' are w-FP, if x is in C (respectively C'), no coalition of C (respectively C') of size at most w can frame x. Thus D is w-FP. ■

This construction immediately gives us the following result.

Theorem 4.2 *Let C be an (n, N, q, w)-FP code. Then it is possible to construct $(2^i n, N, 2^i q, w)$-FP codes for each integer $i \geq 1$.*

Recursive Construction : From Theorem 3.1 we know that C is an (n, N, q, w)-FP code iff $\mathcal{H}(C)$ is an $(N, n, q, 1, w)$-SHF. A recursive construction for (N, n, q, w_1, w_2)-SHF was presented in [11]. Specializing this construction to the case of $w_1 = 1$ and $w_2 = w$ we get the following result.

Theorem 4.3 *For any positive integers q and w, there exists an infinite class of (n, N, q, w)-FP codes for which N is $O(w^{\log^* n} \log n)$.*

5 IPP codes

In this section we study IPP codes. We present a necessary condition for IPP codes and introduce a generaliztion of IPP codes.

Theorem 5.1 *Let C be an (n, N, q, w)-IPP code and $x \in \mathsf{desc}_w(C)$. Then there is a nonempty set of indices $S \subseteq \{1, \ldots, N\}$ and a unique $y \in C$, such that $x_i = y_i$ for all $i \in S$.*

Proof : If $x \in C$, then the result is clearly true on choosing $S = \{1, \ldots, N\}$. Suppose $x \in \mathsf{desc}_w(C) - C$. Since $x \in \mathsf{desc}_w(C)$, there are coalitions $C_1, \ldots, C_\alpha$ of size atmost w, such that $x \in \mathsf{desc}(C_i)$, for each $1 \leq i \leq \alpha$. Without loss of generality we require the following conditions on the collection C_i.

1. Each C_i is minimal, in the sense that if any codeword is dropped from C_i, then x is no longer in $\mathsf{desc}(C_i)$.

2. The collection $\{C_i : 1 \leq i \leq \alpha\}$ is to be the maximum over all collections satisfying 1. In other words we require $C_1, \ldots, C_\alpha$ to be a maximum collection of minimal coalitions such that $x \in \mathsf{desc}(C_i)$ for each $1 \leq i \leq \alpha$.

Since $x \in \mathsf{desc}(C_i)$ for each i and C is w-IPP it follows that the intersection of all the C_i's are nonempty. Let y be in the intersection of all the C_i's. By the minimality of each C_i, it follows that for each i, there is a nonempty set of columns $S_{y,i}$ such that $x_j = y_j$ for each $j \in S_{y,i}$. Let $S_y = \cup_{i=1}^{i=\alpha} S_{y,i}$. Suppose there is another codeword v in C such that $x_j = v_j$ for each $j \in S_y$. By the minimality of each C_i, this codeword cannot be in any of the C_i's. Define $C_1' = (C_1 - \{y\}) \cup \{v\}$. Clearly $x \in \mathsf{desc}(C_1')$. Remove codewords (if required) from C_1' to ensure that C_1' is a minimal coalition such that $x \in \mathsf{desc}(C_1')$. The above procedure ensures $|C_1'| \leq |C_1|$, $x \in \mathsf{desc}(C_1')$. Moreover, since $v \in C_1'$ and $v \notin C_i$, C_1' is distinct from each C_i, $1 \leq i \leq \alpha$. Thus there are $\alpha + 1$ distinct minimal coalitions which produces x. This contradicts the maximality of α. ∎

The converse of the above result is weaker.

Theorem 5.2 *Let C be an (n, N, q) code. Suppose for every $x \in \mathsf{desc}_w(C)$, there is a nonempty set $S \subseteq \{1, \ldots, N\}$ of columns and a unique $y \in C$, such that for any $i \in S$ and any codeword $z \in C$, $x_i = z_i$ implies $z = y$. Then C is an (n, N, q, w)-IPP code.*

Proof : Let $C_1, \ldots, C_\alpha$ be coalitions of size atmost w which produce x. Let S be the subset of columns and y be the unique codeword in C corresponding to x. We claim this y must be in all the C_i's. Suppose $j \in S$. Then for each C_i there is a codeword $y^i \in C_i$ such that $y^i_j = x_j$. But then by the condition on the theorem we must have $y^i = y$. Thus the intersection of all the C_i's is nonempty. Hence C is w-IPP. ∎

Corollary 5.1 *Let C be an (n, N, q) code. Suppose for every $x \in \mathsf{desc}_w(C)$, there is an index i and a $y \in C$, such that for any $z \in C$, $z_i = x_i$ implies $z = y$. Then C is an (n, N, q, w)-IPP code.*

We now introduce a hierarchy of codes between the SFP and the IPP codes.

Definition 5.1 *Let C be an (n, N, q) code. Suppose for every k coalitions $(1 \leq k \leq t)$ $C_1, \ldots, C_k$ of C, each of size atmost w, $x \in \cap_{i=1}^{i=k} \mathsf{desc}(C_i)$ implies $\cap_{i=1}^{i=k} C_i \neq \emptyset$. Then we call C to be an (n, N, q, w, t)-IPP code.*

Clearly the case $t = 2$ correspond to the SFP codes. The hierarchy relation is made clear in the following result which is immediate from the definition of (n, N, q, w, t)-IPP codes.

Proposition 5.1 *For an (n, N, q) code the following holds.*

$$(n, N, q, w)\text{-}SFP = (n, N, q, w, 2)\text{-}IPP \leq (n, N, q, w, 3)\text{-}IPP \leq \ldots \leq (n, N, q, w,)\text{-}IPP.$$

Thus an (n, N, q, w)-IPP code is always an (n, N, q, w, t)-IPP code for any fixed t. Thus the new codes are between the SFP and the IPP codes.

An $(n, N, q, w, \binom{n}{w})$-IPP code is certainly an (n, N, q, w)-IPP code. So a relevant question is whether the hierarchy of codes in Definition 5.1 is strict from $t = 1$ to $t = \binom{n}{w}$. The next result shows that for $w = 2$, this is not the case.

Proposition 5.2 *A code C is an $(n, N, q, 2)$-IPP code iff it is an $(n, N, q, 2, 3)$-IPP code.*

Proof : The only if part is clear so we prove the if part. Suppose C is an $(n, N, q, 2, 3)$-IPP code. We claim this forces $\mathcal{H}(C)$ to be both a $(N, n, q, 3)$-PHF and an $(N, n, q, 2, 2)$-SHF. The proof consists of two parts.

(A) If $\mathcal{H}(C)$ is not an $(N, n, q, 3)$-PHF then C is not an $(n, N, q, 2, 3)$-IPP code.

(B) If $\mathcal{H}(C)$ is not an $(N, n, q, 2, 2)$-SHF then C is not an $(n, N, q, 2, 3)$-IPP code.

(A) If $\mathcal{H}(C)$ is not an $(N, n, q, 3)$-PHF, then there are three rows r_1, r_2, r_3 of C such that no column separates them. Let $C_1 = \{r_1, r_2\}$, $C_2 = \{r_2, r_3\}$ and $C_3 = \{r_3, r_1\}$. Use the majority rule on each column of r_1, r_2, r_3 to construct a word x. Then clearly $x \in \cap_{i=1}^{i=3}\mathsf{desc}(C_i)$, and $\cap_{i=1}^{i=3} C_i = \emptyset$. Thus C is not $(n, N, q, 2, 3)$-IPP.
(B) If $\mathcal{H}(C)$ is not an $(N, n, q, 2, 2)$-SHF there are rows r_1, r_2, r_3, r_4 of C such that $\{r_1, r_2\} \cap \{r_3, r_4\} = \emptyset$ but no column separates the rows r_1, r_2 from r_3, r_4. But this means that $\mathsf{desc}(\{r_1, r_2\}) \cap \mathsf{desc}(\{r_3, r_4\}) \neq \emptyset$. Hence C is not $(n, N, q, 2)$-SFP.

The result now follows on using Theorem 3.2. ■

A sufficient condition for the existence of (n, N, q, w)-IPP codes can be obtained in terms of perfect hash families [8] and partially hashing families [3]. We provide a sufficient condition for the existence of (n, N, q, w, t)-IPP codes. The idea of the proof is from Theorem 2.8 of [8].

Theorem 5.3 *Let C be an (n, N, q) code. If $\mathcal{H}(C)$ is an $(N, n, q, w, k(w+2-k) - w)$-SSHF, where k is such that $k(w+2-k) = \max_{2 \le i \le t}\{i(w+2-i)\}$. Then C is an (n, N, q, w, t)-IPP code.*

Proof : Let $D = \{C_1, \ldots, C_r\}$ be a minimal set of coalitions such that $x \in \cap_{i=1}^{i=r}\mathsf{desc}(C_i)$ and $\cap_{i=1}^{i=r} C_i = \emptyset$. Thus if we drop any C_i from D, then there is at least one codeword $y_i \in \cap_{j \neq i} C_j$. Let $E = \cup_{i=1}^{i=r} C_i$ and $|E| = \beta$. We now obtain an upper bound on β. Each C_i contains the codewords $y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_r$ and it can contain atmost $w - (r-1)$ other codewords. Hence

$$\beta \le r + r(w - r + 1) = r(w + 2 - r). \tag{9}$$

We have $\mathcal{H}(C)$ to be an $(N, n, q, w, \beta - w)$-SSHF. Let A be a subset of E such that $C_1 \subseteq A$ and $|A| = w$. Clearly such an A exists. Let $B = E - A$. Then $|B| = \beta - w$. By the SSHF property, there is an $f \in \mathcal{H}(C)$ (an index $j \in \{1, \ldots, n\}$) such that $|\mathsf{elem}_A(j)| = w$ and $\mathsf{elem}_A(j) \cap \mathsf{elem}_B(j) = \emptyset$. But this implies $|\mathsf{elem}_{C_1}(j)| = |C_1|$ and $\mathsf{elem}_{C_1}(j) \cap \mathsf{elem}_{E-C_1}(j) = \emptyset$. Since $x \in \mathsf{desc}(C_1)$, there is a codeword $z \in C_1$, such that $z_j = x_j$. Since no codeword in E is present in all coalitions in D, let C' be a coalition of D not containing z. But then $x \notin \mathsf{desc}(C')$, since no codeword in C' can equal x on index j. This is a contradiction. ■

Corollary 5.2 *Let C be an (n, N, q) code. If $\mathcal{H}(C)$ is an $(N, n, q, w, \lfloor (\frac{w+2}{2})^2 \rfloor - w)$-SSHF, then C is an (n, N, q, w)-IPP code.*

Proof : If t is not fixed, then the maximum value of β in (9) is $\lfloor (\frac{w+2}{2})^2 \rfloor$. The rest of the argument is similar. ■

Corollary 5.2 was obtained in [3] in terms of partially hashing property. In [3], the probablistic method based on Corollary 5.2 was used to show the existence of (n, N, q, w)-IPP codes for all $q \geq w+1$. This also implies the existence of (n, N, q, w, t)-IPP codes for each t and $q \geq w+1$.

Theorem 2.3 can be used to provide a sufficient condition for the existence of IPP codes in terms of error correcting codes. Let C be an (n, N, q, D) error correcting code with

$$D > N\left(1 - \frac{1}{\binom{w}{2} + w(\beta - w)}\right). \tag{10}$$

Using Theorem 2.3 we have $\mathcal{H}(C)$ is an $(N, n, q, w, \beta - w)$-SSHF. From Theorem 5.3 and Corollary 5.2 we have the following.

1. If $\beta = k(w+2-k)$, where k is such that $k(w+2-k) = \max_{2 \leq i \leq t}\{i(w+2-i)\}$, then C is an (n, N, q, w, t)-IPP code.
2. If $\beta = \lfloor \frac{(w+2)^2}{4} \rfloor$, then C is an (n, N, q, w)-IPP code.

Using this and Theorem 2.4, we get the following result.

Theorem 5.4 *For any poitive integers q and w, there exists an infinite class of (n, N, q, w)-IPP codes for which N is $O((w^3)^{\log^* n} \log n)$.*

Proof : By the above discussion, existence of $(N, n, q, w, \beta - w)$-SSHF with $\beta = \lfloor \frac{(w+2)^2}{4} \rfloor$ is sufficient for the existence of (n, N, q, w)-IPP codes. Using Theorem 2.4, we know that $(N, n, q, w, \beta - w)$-SSHF exists with N to be $O((w(w+\beta-w))^{\log^* n} \log n)$. Since β is $O(w^2)$ we get the required result.
■

References

[1] N. Alon. Explicit construction of exponential sized families of k-independent sets. *Discrete Mathematics*, Volume 58(1986), pp 191-193.

[2] S.R. Blackburn. Frameproof codes. *Preprint*, 2000.

[3] A. Barg, G. Cohen, S. Encheva, G. Kabatiansky and G. Zémor. A hypergraph approach to identifying parent property: the case of multiple parents. DIMACS Technical Report 2000-20.

[4] D. Boneh and J. Shaw. Collision-secure fingerprinting for digital data. *IEEE Transactions on Information Theory*, Volume 44(1998), pp 1897-1905.

[5] Z.J. Czech, G. Havas and B.S. Majewski. Perfect hashing. *Theoretical Computer Science*, Volume 182 (1997), pp 1-143.

[6] J. Friedman. Constructing)$(n \log n)$ size monotone formulae for the kth elementary symmetric polynomial of n Boolean variables. *Proceedings of the 25th Annual Symposium on Foundations of Computer Science*, (1984), 506-515.

[7] H.D.L. Hollman, J.H. van Lint, J-P. Linnartz and L.M.G.M. Tolhuizen. On codes with the identifiable parent property. *Journal of Combinatorial Theory A*, Volume 82 (1998), pp 121-133.

[8] J.N. Staddon, D.R. Stinson and R. Wei. Combinatorial properties of frameproof and traceability codes. *CACR Technical Report CORR 2000-16*, http://cacr.math.uwaterloo.ca, 2000.

[9] D.R. Stinson, T. van Trung and R. Wei. Secure frameproof codes, key distribution patterns, group testing algorithms and related structures. *Journal of Statistical Planning and Inference*, Volume 86 (2000), pp 596-617.

[10] D.R. Stinson and R. Wei. Combinatorial properties and constructions of traceability schemes and frameproof codes. *SIAM Journal of Discrete Mathematics*, Volume 11 (1998), pp 41-53.

[11] D.R. Stinson, R. Wei and L. Zhu. New constructions for perfect hash families and related structures using combinatorial designs and codes. *Journal of Combinatorial Designs*, Volume 8 (2000), pp 189-200.