GENERALIZED TRACEABILITY CODES

Majid MAZROOEI¹, Ali ZAGHIAN²

In this paper, we introduce generalizations of frameproof, secure frameproof, traceability and identifiable parent property codes and will study some of their basic combinatorial properties.

Keywords: Frameproof, Identifiable parent property, Hash family, Traceability.

1. Introduction

Throughout this paper, $Q$ is an alphabet of size $q$ and $\mathcal{C} \subseteq Q^n$ is a $q$–array code of length $n$. If $|\mathcal{C}| = M$, then we call $\mathcal{C}$ an $(n,M,q)$–code. The elements of $\mathcal{C}$ are called code words and each code word will have the form $x = (x_1, \ldots, x_n)$, where $x_i \in Q$, $1 \leq i \leq n$. The matrix representation of the code $\mathcal{C}$, denoted $H(\mathcal{C})$, is an $M \times n$ matrix on $q$ symbols where each row of the matrix corresponds to one of the code words.

Codes providing some forms of traceability (TA, for short) to protect copyrighted digital data against piracy have been extensively studied in the recent years. The weak forms of such codes are frameproof codes introduced by Boneh and Shaw [2], and secure frameproof codes. The more strong form of these codes are identifiable parent property (IPP, for short) codes which have been introduced by Hollmann, Van Lint, Linnartz and Tolhuizen [4] and defined as follows.

Let $\mathcal{C} \subseteq Q^n$ be a $q$–array code. For any subset of code words $\mathcal{C}_0 \subseteq \mathcal{C}$, the set of descendants of $\mathcal{C}_0$, denoted $\text{desc}(\mathcal{C}_0)$, is defined by

$$\text{desc}(\mathcal{C}_0) = \{x \in Q^n | x_i \in \{a_i : a \in \mathcal{C}_0\}, 1 \leq i \leq n\}.$$ 

Thus $\text{desc}(\mathcal{C}_0)$ consists of all $n$–tuples that could be produced by a coalition holding the code words in $\mathcal{C}_0$. If $x \in \text{desc}(\mathcal{C}_0)$, then we say that $\mathcal{C}_0$ produces $x$.

Let $w$ be an integer. Define the $w$–descendant code, denoted $\text{desc}_w(C)$, as follows:

$$\text{desc}_w(C) = \bigcup_{\mathcal{C}_0 \subseteq \mathcal{C} | \mathcal{C}_0 \leq w} \text{desc}(\mathcal{C}_0)$$

Thus $\text{desc}_w(C)$ consists of all $n$–tuples that could be produced by some coalition of size at most $w$. Now the code $\mathcal{C}$ is called an $(n,M,q,w)$–identifiable parent property code ($w$–IPP) provided that, for all $x \in \text{desc}_w(C)$, it holds that

¹ Assistant Prof., Dept. of Mathematical Sciences, University of Kashan, Isfahan, Iran, P. O. Box: 87317-51167, e-mail: m.mazrooei@kashanu.ac.ir
² Assistant Prof., Dept. of Mathematics, Malek Ashtar University of Technology, Isfahan, Iran, P. O. Box: 83145-115, e-mail: a_zaghian@mut-es.ac.ir
Another strong versions of such codes are TA schemes and TA codes introduced by Chor, Fiat and Naor in [3]. In fact, TA codes turn out to be a subclass of IPP codes and are defined as follows.

Let define \( I(x, y) = \{ i : x_i = yi \} \) for any \( x, y \in Q^n \). Suppose \( C \subseteq Q^n \) is an \((n, M, q)\) - code and \( w \geq 2 \) is an integer. \( C \) is called a \( w \)-traceability code (\( w \)-TA) provided that, for all \( i \) and all \( x \in \text{desc}(C_i) \), \( C_i \subseteq C \) and \(|C_i| \leq w\), there is at least one code word \( y \in C_i \) such that \(|I(x, y)| > |I(x, z)|\) for any \( z \in CC_i \).

Combinatorial properties of IPP codes and TA codes have been studied by Staddon, Stinson and Wei [9], Sarkar and Stinson [7], Barg, et al. [1], and also in [11]. The question of traitor tracing algorithms for IPP and TA codes is treated in [8], e.g. certain classes of TA codes are shown to have a faster tracing algorithm than their initially known linear runtime by using the list decoding techniques. New results on bounds of frame-proof codes and TA schemes can be found in [6].

Recently, a major theoretical challenge is to derive more codes which have efficient tracing algorithms. To do this, some authors have given some generalizations of IPP and TA codes. Sarkar and Stinson [7] defined an \((n, M, q, w, t)\)-IPP code as a code \( C \subseteq Q^n \) with the property that for every \( k \) coalitions \((1 \leq k \leq t)\) \( C_1, \ldots, C_k \) of \( C \), each of size at most \( w \), \( x \in \bigcap_{i=1}^k \text{desc}(C_i) \) implies \( \bigcap_{i=1}^k C_i \neq \emptyset \). They have studied this generalization in part. For example, they proved that a code \( C \) is an \((n, M, q, 2)\)-IPP code if and only if it is an \((n, M, q, 2, 3)\)-IPP code. Another generalization is given by Wu and Sattar [12] for TA codes. They have defined TA codes relative to any well-defined metric and an efficient tracing algorithm is presented for Lee-metric TA codes.

In this paper, we introduce new generalizations of frame-proof, secure frame-proof, identifiable parent property and traceability codes and will study them.

2. Generalized Traceability Codes

From now, \( \alpha_1, \ldots, \alpha_n : C \rightarrow Q \) are maps and \( \alpha = \{\alpha_1, \ldots, \alpha_n\} \). For any \( x \in C \) and for any \( X \subseteq C \), let \( a(x) = (\alpha_1(x), \ldots, \alpha_n(x)) \) and \( a(X) = \{a(x) \mid x \in X\} \). For any subset of code words \( C_0 \subseteq C \), we define the set of \( a \)-descendants of \( C_0 \), denoted \( \text{desc}_a(C_0) \), by

\[
\text{desc}_a(C_0) = \{x \in Q^n \mid x_i \in \{a_i(a) : a \in C_0\}, 1 \leq i \leq n\}.
\]

Now, for any integer \( w \geq 2 \) define the \((a, w)\)-descendant code, denoted \( \text{desc}_{a,w}(C) \), as follows:

\[
\text{desc}_{a,w}(C) = \bigcup_{C_0 \subseteq C \mid |C_0| = w} \text{desc}_a(C_0).
\]
Definition 2.1 Let $C$ be an $(n,M,q)$–code and $w \geq 2$ be an integer. We say that:

1. $(a,w)$–Frameproof code: $C$ is an $(n,M,q,a,w)$–frameproof code ($(a,w)$–FP), if for any $C_0 \subseteq C$ of size at most $w$, $x \in \text{desc}_a(C_0) \cap a(C)$ implies $x \in a(C_0)$.

2. $(a,w)$–Secure frameproof code: $C$ is an $(n,M,q,a,w)$–secure frameproof code ($(a,w)$–SFP), if for any subsets $C_0,C_i$ of $C$ of cardinality at most $w$, $\text{desc}_a(C_0) \cap \text{desc}_a(C_i) \neq \emptyset$ implies $C_0 \cap C_i \neq \emptyset$.

3. $(a,w)$–Traceability code: $C$ is an $(n,M,q,a,w)$–traceability code ($(a,w)$–TA) if for any $C_0 \subseteq C$ of cardinality at most $w$ and for any $x \in \text{desc}_a(C)$, there exists $y \in C_0$ such that $|l(x,a(y))| > |l(x,a(z))|$ for all $z \in C C_0$.

4. $(a,w)$–Identifiable parent property code: $C$ is an $(n,M,q,a,w)$–identifiable parent property code ($(a,w)$–IPP) provided that, for all $x \in \text{desc}_{a,w}(C)$, it holds that

$$\bigcap_{\{i:j \leq w, x \in \text{desc}_a(C_j)\}} C_i \neq \emptyset$$

Note that the original definitions of FP, SFP, IPP and TA codes will be obtained if we take $a = \{\pi_1, \ldots, \pi_n\}$, where $\pi_i : C \rightarrow Q, 1 \leq i \leq n$, is the natural projection on the $i$–th component. Our first result determines the relations between our generalized traceability codes.

Proposition 2.2. The following relationships hold for any $(n,M,q)$–code.

$(a,w)$–TA $\Rightarrow$ $(a,w)$–IPP $\Rightarrow$ $(a,w)$–SFP $\Rightarrow$ $(a,w)$–FP.

Proof. Assume that $C$ is an $(n,M,q,a,w)$–TA code. We show that $C$ is an $(a,w)$–IPP code. Let $x \in \text{desc}_{a,w}(C)$. Then there exists $C_0 \subseteq C, |C_0| \leq w$, such that $x \in \text{desc}_a(C_0)$. Let $y \in C_0$ such that $|l(x,a(y))| \geq |l(x,a(z))|$ for all $z \in C_0$. Thus $|l(x,a(y))| > |l(x,a(z))|$ for any $z \in C$ by assumption. We show that for any $C_i \subseteq C, |C_i| \leq w, x \in \text{desc}_a(C_i)$ implies $y \in C_i$, which proves that $C$ is an $(a,w)$–IPP code. By assumption, there is $y' \in C_i$ such that $|l(x,a(y'))| > |l(x,a(z))|$ for any $z \in C C_i$. If $y \neq y'$, then we will have $|l(x,a(y'))| > |l(x,a(y))|$, which is a contradiction. Hence $y = y' \in C_i$.

It is easy to see that any $(a,w)$–IPP code is an $(a,w)$–SFP code. We just prove that if $C$ is an $(a,w)$–SFP code, then $C$ is an $(a,w)$–FP code. Let $C_0$ be a subset of $C$ of size at most $w$ and $x \in \text{desc}_a(C_0) \cap a(C)$. Thus there exists a code word $u \in C$ such that $x = a(u)$. Let $C_i = \{u\}$. Then $x \in \text{desc}_a(C_i)$. By assumption, we have $C_0 \cap C_i \neq \emptyset$ showing that $u \in C_0$. Hence $x \in a(C_0)$ which completes the proof.

Proposition 2.3. Let $C$ be an $(n,M,q)$–code and let

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

Assume that $w \geq 2$ is an integer such that $d_a > n(1 - \frac{1}{w^2})$. Then $C$ is an $(a,w)$–TA code.
Theorem 3.3. Let \( \lambda = n(1 - w^2) \) and \( \beta = w^2 - 1 \). Assume that \( C_0 \subseteq C \), \( |C_0| \leq w \), be a set of code words and \( x \in \text{desc}_a(C_0) \). For any \( u \in \text{desc}_a(C_0) \), let \( M(u) = \max\{|l(u, \alpha(c))| \mid c \in C_0\} \) and \( M = \min_{u \in \text{desc}_a(C_0)} M(u) \). For any \( u \in \text{desc}_a(C_0) \) and \( c \in C_0 \), we have \( |l(u, \alpha(c))| \leq M(u) \) which shows that \( \sum_{c \in C_0} |l(u, \alpha(c))| \leq wM(u) \). On the other hand, we have \( n \leq \sum_{c \in C_0} |l(u, \alpha(c))| \). Thus \( n \leq wM(u) \) for any \( u \in \text{desc}_a(C_0) \). Hence \( \frac{n}{w} \leq M \). Now let \( y \in C_0 \) be a code word such that \( |l(x, \alpha(y))| = M(x) \). Then for any \( z \in CC_0 \) we have

\[
|l(x, \alpha(z))| \leq \sum_{c \in C_0} |l(\alpha(C), \alpha(z))| \leq \sum_{c \in C_0} \beta \leq w\beta < \frac{n}{w} \leq M(x).
\]

This completes the proof. 

Example 2.4. Let \( Q = F_{11} \) and \( C = \{100, 411, 511\} \). Define the maps

\[
\alpha_i : C \rightarrow Q (i = 1, 2, 3)
\]

such that \( \alpha_i(100) = 2, \alpha_i(411) = 3 \) and \( \alpha_i(511) = 6 \) for \( i = 1, 2, 3 \). Then we have \( d_a = 3 > 4 \). This shows that \( C \) is an \((3, 3, 11, \alpha, 2)\)–TA code while \( C \) is not a 2–TA code.

3. Generalized Traceability Codes and Hash Families

A finite set \( H \) of \( n \) functions \( h : A \rightarrow B \), where \( |A| = M > |B| = m \), is called an \((n, M, m)\)–hash family, denoted by \((n, M, m)\)–HF. An \((n, M, m)\)–HF \( H \) can be presented as an \( M \times n \) matrix on \( m \) symbols, where each column of the matrix corresponds to one of the functions in \( H \). Sometimes it is easier to consider hash families as such matrices.

Definition 3.1. We consider the following kinds of hash families.

1. Perfect hash family: An \((n, M, m)\)–HF \( H \) is called an \((n, M, m, w)\)–perfect hash family \((n, M, m, w)\)–PHF if for any subset \( X \) of \( A \) of size \( w \), there is a function \( f \in H \) such that \( f \) is injective on \( X \).

2. Separating hash family: An \((n, M, m)\)–HF \( H \) is called an \((n, M, m, w_1, w_2)\)–separating hash family \((n, M, m, w_1, w_2)\)–SHF if for any disjoint subsets \( X, Y \) of \( A \) with \( |X| = w_1 \) and \( |Y| = w_2 \), there is a function \( f \in H \) such that \( f(X) \) and \( f(Y) \) are also disjoint.

The following theorem, due to Staddon, Stinson and Wei [9], makes a connection between perfect hash families and IPP codes.

Theorem 3.2. Let \( C \) be an \((n, M, q)\)–code whose matrix representation is \( C \). If \( C \) is an \((n, M, q, 1)\)–PHF code, then \( C \) is an \((n, M, q, w)\)–IPP code.

Stinson, Van Trung and Wei [10] obtained another connections between (secure) frame-proof codes and separating hash families, which is stated in the following theorem.

Theorem 3.3.

1. A code \( C \) is an \((n, M, q, w)\)–FP code if \( H(C) \) is an \((n, M, q, w, 1)\)–SHF.
2. A code \( C \) is an \((n, M, q, w)\)–SFP code if \( H(C) \) is an \((n, M, q, w, w)\)–SHF.
In this section, we will derive similar results of our generalizations of FP, SFP and IPP codes. Before it, we need the following definition.

**Definition 3.4.** Let $C$ be an $(n, M, q)$-code, $\alpha_1, \ldots, \alpha_n : C \rightarrow \mathbb{Q}$ be maps and $\alpha = \{\alpha_1, \ldots, \alpha_n\}$. We say that $\alpha$ is an injective family if for any $x, y \in C$, $\alpha(x) = \alpha(y)$ implies $x = y$.

**Proposition 3.5.** Let $C$ be an $(n, M, q)$-code and $w \geq 2$ be an integer. If $\alpha$ is an $(n, M, q, w, 1)$–SHF, then $C$ is an $(a, w)$–FP code. The converse holds if $\alpha$ is an injective family.

**Proof.** Assume that $C_0 \subseteq C$, $|C_0| = w$ and $x \in \text{desc}_\alpha(C_0) \cap \alpha(C)$. Thus there exists $u \in C$ such that $x = \alpha(u)$. We claim that $u \in C_0$ which show that $C$ is an $(a, w)$–FP code. Let $C_1 = \{u\}$, if $u \notin C_0$ then $C_0 \cap C_1 = \emptyset$. Hence, by assumption, there is $1 \leq i \leq n$ such that $\alpha_i(\alpha) \in \alpha_i(C_0)$, which is a contradiction because $\alpha(u) = x \in \text{desc}_\alpha(C_0)$.

Now assume that $C$ is an $(a, w)$–FP code and $\alpha$ is an injective family. We prove that $\alpha$ is an $(n, M, q, w, 1)$–SHF. If not, then there exist disjoint subsets $C_0, C_1$ of $C$ with $|C_0| = 1$ and $|C_1| = w$ such that for any $1 \leq i \leq n$, $\alpha_i(C_0) \cap \alpha_i(C_1) \neq \emptyset$. This shows that there is a code word $a \in C_0$ such that $\alpha(a) \in \text{desc}_\alpha(C_1)$. Since $C$ is an $(a, w)$–FP code, $\alpha(a) \in \alpha(C_1)$. Thus there exists $b \in C_1$ such that $\alpha(a) = \alpha(b)$, which, by our assumption, implies $a = b$. Hence $C_0 \cap C_1 \neq \emptyset$, a contradiction. Thus $\alpha$ is an $(n, M, q, w, 1)$–SHF.

**Proposition 3.6.** Let $C$ be an $(n, M, q)$-code and $w \geq 2$ be an integer. Then $C$ is an $(a, w)$–SFP code if $\alpha$ is an $(n, M, q, w, w)$–SHF.

**Proof.** Assume that $C$ is an $(a, w)$–SFP code. We show that $\alpha$ is an $(n, M, q, w, w)$–SHF. If not, then there exist disjoint subsets $C_0, C_1$ of $C$, both of size $w$, such that $\alpha_i(C_0) \cap \alpha_i(C_1) \neq \emptyset$ for all $1 \leq i \leq n$. This shows that for any $1 \leq i \leq n$, there exist $u_i \in C_0$ and $v_i \in C_1$ such that $\alpha_i(u_i) = \alpha_i(v_i)$. Since $\text{desc}_\alpha(C_0) \cap \text{desc}_\alpha(C_1) \neq \emptyset$, which is a contradiction because $C$ is an $(a, w)$–SFP code. Hence $\alpha$ should be an $(n, M, q, w, w)$–SHF.

Conversely, assume that $\alpha$ is an $(n, M, q, w, w)$–SHF. We prove that $C$ is an $(a, w)$–SFP code. Let $C_0, C_1$ be disjoint subsets of $C$ both of size $w$. We shall show that $\text{desc}_\alpha(C_0) \cap \text{desc}_\alpha(C_1) = \emptyset$. By assumption, there exists $1 \leq i \leq n$ such that $\alpha_i(C_0) \cap \alpha_i(C_1) = \emptyset$. This shows that $\text{desc}_\alpha(C_0) \cap \text{desc}_\alpha(C_1) = \emptyset$, as desired.

**Proposition 3.7.** Let $C$ be an $(n, M, q, a, w)$–IPP code where $M \geq w + 1$. Then the set $\alpha = \{\alpha_1, \ldots, \alpha_n\}$ is an $(n, M, q, w + 1)$–PHF.

**Proof.** If not, then there is $X \subseteq C$ such that $|X| = w + 1$ and for any $1 \leq j \leq n$, there are $x \neq x' \in X$ such that $\alpha_j(x) = \alpha_j(x') = u_j$. Let $u = (u_1, \ldots, u_n)$. Then for any $x \in X$, we have $u \in \text{desc}_\alpha(C_x)$ where $\bigcap_{x \in X} C_x = X \setminus \{x\}$.
But for any $x \in X$, $|C_x| = w$ and we have which contradicts the fact that $C$ is an $(a,w)$–IPP code. Hence the set $a$ is an $(n,M,q,w + 1)$–PHF.

**Proposition 3.8.** A code $C$ is an $(n,M,q,a,2)$–IPP code iff the family $a$ is both an $(n,M,q,3)$–PHF and an $(n,M,q,2,2)$–SHF.

**Proof.** If $C$ is an $(n,M,q,a,2)$–IPP code then $a$ is both an $(n,M,q,3)$–PHF and an $(n,M,q,2,2)$–SHF by propositions 3.6 and 3.7. Assume that $a$ is both an $(n,M,q,3)$–PHF and an $(n,M,q,2,2)$–SHF. We show that $C$ is an $(n,M,q,a,2)$–IPP code. Let $x \in \text{desc}_a(C_i), i = 1,\ldots,r$, where $C_i = \{a_{i1}, a_{i2}\}$, $1 \leq i \leq r$, are distinct subsets of $C$. If $C_1 \cap C_2 = \emptyset$, then by assumption, there exists $1 \leq i \leq n$ such that $a_i(C_1) \cap a_i(C_2) = \emptyset$, which is a contradiction because $x \in \text{desc}_a(C_i) \cap \text{desc}_a(C_2)$. Hence $C_1 \cap C_2 \neq \emptyset$. Without loss of generality we may assume that $a_{11} = a_{21}$. Note that in this case we have $a_{12} = a_{22}$. We claim that $a_{11} \in C_j$ for all $j = 3,\ldots,r$. If not, then there exists $3 \leq j \leq r$ such that $a_{11} \in C_j$. Since $C_1 \cap C_j \neq \emptyset$, we will have $a_{12} \in C_j$. Similarly, $a_{22} \in C_j$. Hence $C_j = \{a_{12}, a_{22}\}$. By assumption, there exists $1 \leq l \leq n$ such that $a_i(a_{11}), a_i(a_{12})$ and $a_i(a_{22})$ are distinct. Now there are 2 cases:

**Case 1.** $x_l = a_i(a_{11})$: In this case we will have $x \in \text{desc}_a(C_i)$, which is a contradiction.

**Case 2.** $x_l = a_i(a_{12})$: In this case we will have $x \in \text{desc}_a(C_2)$, which is a contradiction.

Hence $a_{11} \in C_k$ for all $k = 1,\ldots,r$, as desired.

**Theorem 3.9.** Let $C$ be an $(n,M,q)$–code. If the family $a = \{a_1,\ldots,a_n\}$ is an $(n,M,q,1(w + 2)^2/4)$–PHF, then $C$ is an $(a,w)$–IPP code.

**Proof.** If not, then there exist $x \in \bigcap_{i:x \in \text{desc}_a(C_i), |C_i| \geq w} C_i$ such that $\bigcap_{i:x \in \text{desc}_a(C_i), |C_i| \geq w} C_i = \emptyset$.

Let $D = \bigcap_{i:x \in \text{desc}_a(C_i), |C_i| \geq w} C_i$ and define

$$r = \min \left\{|D' \subseteq D, \bigcap_{C_i \in D'} C_i = \emptyset\right\}.$$

Without loss of generality, we may assume that $D' = \{C_1,\ldots,C_r\}$ is a family of $r$ distinct elements of $D$ which have no common code words and for any $1 \leq i \leq r$, $x \in \text{desc}_a(C_i)$. Let $C_i = \{y^{(1)},\ldots,y^{(\beta)}\}$. Hence for any $1 \leq i \leq r$, there exists $y^{(k_i)} \in \bigcup_{j=1,j\neq i} C_i$ such that $y^{(k_i)} \notin C_i$. This shows that for any $1 \leq i \leq r$, we have

$$|C_i - \{y^{(k_1)},\ldots,y^{(k_r)}\}| \leq w - (r - 1),$$

which implies $\beta \leq r + r(w - r + 1) = (w + 2 - r)r$. Since $w$ is an integer, the last inequality shows that $\beta \leq \lfloor (w + 2)2/4 \rfloor$. Now, since $a$ is an $(n,M,q,\beta)$–PHF,
there exists $1 \leq j \leq r$ such that $\alpha_j(y^{(1)}), \ldots, \alpha_j(y^{(\beta)})$ are distinct. On the other hand, there exists $1 \leq t \leq \beta$ such that $x_j = \alpha_t(y^{(t)})$. Also there is some $C_t \in D'$ such that $y^{(t)} \in C_t$. Since there is no code word $u \in \cup_{c_i \in D'} C_i$ such that $\alpha_j(u) = y^{(t)}$, we have $x \notin \text{desc}_{\alpha}(C_t)$ which is a contradiction. Hence $C$ is an $(a, w)$-IPP code.

In [5], it is proved that if $n \geq w^2 = \log M$, then there exists an $(n, M, m, w)$--PHF. Hence we have the following corollary.

**Corollary 3.10.** Let $c$ be an $(n, M, q)$--code and $w \geq 2$ be an integer such that $n \geq \left\lceil \frac{(w+2)^2}{4} \right\rceil q log M$. Then there are maps $\alpha_1, \ldots, \alpha_n : C \rightarrow Q$ such that $\tilde{C}$ is an $(a, w)$--IPP code.

**4. Conclusion**

We generalized and studied codes providing some forms of traceability (FP, SFP, TA and IPP codes) to generate new families of such codes. Our generalized traceability codes enjoy the basic combinatorial properties of original ones. This motivates us to study decoding algorithms of our families in the future.

**Acknowledgement**

The first author is partially supported by the University of Kashan under grant number 235604/1.

**References**


