Strongly Separable Codes

Ying Miao

joint work with Minquan Cheng and Jing Jiang

University of Tsukuba, Japan

March 16, 2015, ALCOMA 15

1 Introduction

Exam. 1.1 A $(3,4,2)$ code $\quad \mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4\}$.

$$\begin{array}{cccc} \mathbf{c}_1 & \mathbf{c}_2 & \mathbf{c}_3 & \mathbf{c}_4 \end{array}$$
$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \Longrightarrow \overset{\mathbf{c}_1}{\left\{\begin{matrix} 1 \\ 0 \\ 0 \end{matrix}\right\}} \quad \overset{\mathbf{c}_2}{\left\{\begin{matrix} 0 \\ 1 \\ 0 \end{matrix}\right\}} \quad \overset{\mathbf{c}_3,}{\left\{\begin{matrix} 0 \\ 0 \\ 1 \end{matrix}\right\}} \quad \overset{\mathbf{c}_4}{\left\{\begin{matrix} 0 \\ 1 \\ 1 \end{matrix}\right\}}$$

$$\overset{\mathbf{c}_1 \cup \mathbf{c}_2}{\left\{\begin{matrix} 0,1 \\ 0,1 \\ 0 \end{matrix}\right\}} \quad \overset{\mathbf{c}_1 \cup \mathbf{c}_3}{\left\{\begin{matrix} 0,1 \\ 0 \\ 0,1 \end{matrix}\right\}} \quad \overset{\mathbf{c}_1 \cup \mathbf{c}_4}{\left\{\begin{matrix} 0,1 \\ 0,1 \\ 0,1 \end{matrix}\right\}} \quad \overset{\mathbf{c}_2 \cup \mathbf{c}_3}{\left\{\begin{matrix} 0 \\ 0,1 \\ 0,1 \end{matrix}\right\}} \quad \overset{\mathbf{c}_2 \cup \mathbf{c}_4}{\left\{\begin{matrix} 0 \\ 1 \\ 0,1 \end{matrix}\right\}} \quad \overset{\mathbf{c}_3 \cup \mathbf{c}_4}{\left\{\begin{matrix} 0 \\ 0,1 \\ 1 \end{matrix}\right\}}$$

Question: Given a subset of

$$\left\{ \begin{array}{c} 0,1 \\ 0,1 \\ 0,1 \end{array} \right\}$$

say,

$$\left\{ \begin{array}{c} 0 \\ 1 \\ 0,1 \end{array} \right\}$$

Can we trace back to the codewords $\mathbf{c}_2, \mathbf{c}_4$ who produced it?

Answer: Yes, we can. The subsets produced by up to two codewords are all distinct.

Remark: Such kind of codes are used in multimedia fingerprinting where the identification of malicious authorized users taking part in the linear collusion attack is required to prevent pirate copies of multimedia contents.

2 General Definitions and Tracing Properties

Let n, M, q be positive integers, and $Q = \{0, 1, \ldots, q-1\}$.

A set $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_M\} \subseteq Q^n$ is called an (n, M, q) code and each $\mathbf{c}_i$ a codeword (or a fingerprint).

$\forall\ \mathcal{C}' \subseteq \mathcal{C}$, define the descendant code of $\mathcal{C}'$ as

$$\mathsf{desc}(\mathcal{C}') = \mathcal{C}'(1) \times \cdots \times \mathcal{C}'(n),$$

where

$$\mathcal{C}'(i) = \{\mathbf{c}(i) \in Q \mid \mathbf{c} = (\mathbf{c}(1), \ldots, \mathbf{c}(n))^T \in \mathcal{C}'\}.$$

Remark: $\mathsf{desc}(\mathcal{C}')$ consists of the n-tuples that could be produced by a coalition holding the codewords (fingerprints) in $\mathcal{C}'$.

Def. 2.1 Let $\mathcal{C}$ be an (n, M, q) code and $t \geq 2$ be an integer. $\mathcal{C}$ is a $\bar{t}$-separable code, $\bar{t}$-SC(n, M, q), if $\forall$ distinct $\mathcal{C}_1, \mathcal{C}_2 \subseteq \mathcal{C}$ with $|\mathcal{C}_1| \leq t, |\mathcal{C}_2| \leq t$, we have $\mathsf{desc}(\mathcal{C}_1) \neq \mathsf{desc}(\mathcal{C}_2)$.

Tracing: Given $\mathsf{desc}(\mathcal{C}_0)$, to trace $\mathcal{C}_0$, we need check $\mathsf{desc}(\mathcal{C}')$ for all $\mathcal{C}' \subseteq \mathcal{C}$ (separable code) with $|\mathcal{C}'| \leq t$, that is, the computational complexity of the tracing is $O(M^t)$.

Question: Is it possible to find an efficient tracing, say, with computational complexity $O(M)$?

Answer: In general, **NOT**. But in some cases, **OK**.

Def. 2.2 Let $\mathcal{C}$ be an (n, M, q) code and $t \geq 2$ be an integer. $\mathcal{C}$ is a t-frameproof code, t-FPC(n, M, q), if $\forall\, \mathcal{C}' \subseteq \mathcal{C}$ with $|\mathcal{C}'| = t$, and $\forall\, \mathbf{c} \in \mathcal{C} \setminus \mathcal{C}'$, $\exists\, 1 \leq i \leq n$ s.t. $\mathbf{c}(i) \notin \mathsf{desc}(\mathcal{C}')(i)$.

Exam. 2.3 A 2-FPC$(3, 3, 2)$ $\mathcal{C}$. Any t-FPC(n, M, q) is a $\bar{t}$-SC(n, M, q).

$$\mathcal{C} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \mathsf{desc}(\mathcal{C}_0) = \left\{ \begin{matrix} 0 \\ 0,1 \\ 0,1 \end{matrix} \right\}$$

Tracing: Given $\mathsf{desc}(\mathcal{C}_0)$, to trace $\mathcal{C}_0$, we eliminate all codewords $\mathbf{c}$ with $\mathbf{c}(i) \notin \mathsf{desc}(\mathcal{C}_0)(i)$. From the definition of FPC, the set of remaining codewords is necessarily $\mathcal{C}_0$. The computational complexity of the tracing is $O(M)$.

3 Strongly Separable Codes

Question: The constraints posed on frameproof codes are quite strong so that the number of codewords is not large enough. Can we find a new code weaker than a frameproof code but stronger than a separable code, so that the its computational complexity is the same with a frameproof code, i.e., $O(M)$, but the number of codewords in such a code is larger than that of a frameproof code?

Answer: Yes, we can.

Def. 3.1 Let $\mathcal{C}$ be an (n, M, q) code and $t \geq 2$ be an integer. $\mathcal{C}$ is a strongly $\bar{t}$-separable code, $\bar{t}$-SSC(n, M, q), if $\forall\ \mathcal{C}_0 \subseteq \mathcal{C}$, $|\mathcal{C}_0| \leq t$, we have

$$\bigcap_{\mathcal{C}' \in S(\mathcal{C}_0)} \mathcal{C}' = \mathcal{C}_0,$$

where $S(\mathcal{C}_0) = \{\mathcal{C}' \subseteq \mathcal{C} \mid \mathsf{desc}(\mathcal{C}') = \mathsf{desc}(\mathcal{C}_0)\}$.

Exam. 3.2 A $\bar{2}$-SSC$(3, 4, 2)$ $\mathcal{C}$. Any $\bar{t}$-SSC(n, M, q) is a $\bar{t}$-SC(n, M, q).

$$\mathcal{C} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \qquad \mathsf{desc}(\mathcal{C}_0) = \begin{Bmatrix} 0 \\ 0,1 \\ 0,1 \end{Bmatrix}$$

Tracing: Given $\mathsf{desc}(\mathcal{C}_0)$, to trace $\mathcal{C}_0$, we eliminate all codewords $\mathbf{c}$ with $\mathbf{c}(i) \notin \mathsf{desc}(\mathcal{C}_0)(i)$. The computational complexity of the tracing is $O(M)$.

It is obvious that the set

$$\mathcal{C}_L = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

of remaining codewords necessarily contains $\mathcal{C}_0$. We have to find the exact $\mathcal{C}_0$.

- $\mathcal{C}_L \in S(\mathcal{C}_0)$, that is, $\mathsf{desc}(\mathcal{C}_L) = \mathsf{desc}(\mathcal{C}_0)$.
- $\forall\ \mathbf{x} \in \mathcal{C}_0 = \bigcap_{\mathcal{C}' \in S(\mathcal{C}_0)} \mathcal{C}'$, $\exists\ 1 \leq j \leq n$ s.t. $\mathbf{x}(j) = 1, \mathbf{c}(j) = 0$, or $\mathbf{x}(j) = 0, \mathbf{c}(j) = 1$ for any $\mathbf{c} \in \mathcal{C}_L \setminus \{\mathbf{x}\}$. Otherwise $\mathsf{desc}(\mathcal{C}_L \setminus \{\mathbf{x}\}) = \mathsf{desc}(\mathcal{C}_L)$, i.e., $\mathcal{C}_L \setminus \{\mathbf{x}\} \in \mathbf{S}(\mathcal{C}_\mathbf{0})$, so $\mathbf{x} \notin \bigcap_{\mathcal{C}' \in S(\mathcal{C}_0)} \mathcal{C}'$, a contradiction.
- Any $\mathbf{x} \in \mathcal{C}_0 = \bigcap_{\mathcal{C}' \in S(\mathcal{C}_0)} \mathcal{C}'$ is a colluder. Otherwise, $\forall\ \mathcal{C}' \in S(\mathcal{C}_0)$, $\mathcal{C}' \setminus \{\mathbf{x}\} \in S(\mathcal{C}_0)$, so $\mathbf{x} \notin \bigcap_{\mathcal{C}' \in S(\mathcal{C}_0)} \mathcal{C}'$, a contradiction.

It is obvious that the set

$$\mathcal{C}_L = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \qquad \Longrightarrow \quad \mathcal{C}_0 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

of remaining codewords necessarily contains $\mathcal{C}_0$. We have to find the exact $\mathcal{C}_0$.

- $\mathcal{C}_L \in S(\mathcal{C}_0)$, that is, $\mathsf{desc}(\mathcal{C}_L) = \mathsf{desc}(\mathcal{C}_0)$.
- $\forall\ \mathbf{x} \in \mathcal{C}_0 = \bigcap_{\mathcal{C}' \in S(\mathcal{C}_0)} \mathcal{C}'$, $\exists\ 1 \leq j \leq n$ s.t. $\mathbf{x}(j) = 1, \mathbf{c}(j) = 0$, or $\mathbf{x}(j) = 0, \mathbf{c}(j) = 1$ for any $\mathbf{c} \in \mathcal{C}_L \setminus \{\mathbf{x}\}$. Otherwise $\mathsf{desc}(\mathcal{C}_L \setminus \{\mathbf{x}\}) = \mathsf{desc}(\mathcal{C}_L)$, i.e., $\mathcal{C}_L \setminus \{\mathbf{x}\} \in \mathbf{S}(\mathcal{C}_\mathbf{0})$, so $\mathbf{x} \notin \bigcap_{\mathcal{C}' \in S(\mathcal{C}_0)} \mathcal{C}'$, a contradiction.
- Any $\mathbf{x} \in \mathcal{C}_0 = \bigcap_{\mathcal{C}' \in S(\mathcal{C}_0)} \mathcal{C}'$ is a colluder. Otherwise, $\forall\ \mathcal{C}' \in S(\mathcal{C}_0)$, $\mathcal{C}' \setminus \{\mathbf{x}\} \in S(\mathcal{C}_0)$, so $\mathbf{x} \notin \bigcap_{\mathcal{C}' \in S(\mathcal{C}_0)} \mathcal{C}'$, a contradiction.

4 Constructions

Thm. 4.1 (Concatenation) A $\bar{t}$-SSC(n, M, q) implies a $\bar{t}$-SSC $(nq, M, 2)$.

Thm. 4.2 A code C is a $\overline{2}$-SSC$(2, M, q)$ iff it is a $\overline{2}$-SC$(2, M, q)$.

Lemma 4.3 (Cheng et. al., 2012, 2015) Let $k \geq 2$ be a prime power. Then $\exists$ optimal $\overline{2}$-SC$(2, M \approx q^{3/2}, q)$ for any $q \in \{k^2 - 1, k^2 + k - 2, k^2 + k - 1, k^2 + k, k^2 + k + 1\}$.

Coro. 4.4 Let $k \geq 2$ be a prime power. Then $\exists$ optimal $\overline{2}$-SSC $(2, M \approx q^{3/2}, q)$ for any $q \in \{k^2 - 1, k^2 + k - 2, k^2 + k - 1, k^2 + k, k^2 + k + 1\}$.

Remark: A 2-FPC$(2, M, q)$ can have at most $2q$ codewords (Blackburn, 2003), but the above $\overline{2}$-SSC$(2, M, q)$ can have about $q^{3/2}$ codewords.

A Direct Construction

Let q be a positive integer, s a non-negative integer, $0 \le s \le q$, $q - s$ odd. Let $Q = \{\infty_0, \ldots, \infty_{s-1}\} \cup Z_{q-s}$.

Let

$$M_i = \begin{pmatrix} \infty_i & i & 0 \\ 0 & \infty_i & i \\ i & 0 & \infty \end{pmatrix} \qquad M_s = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & q-s-1 \\ 0 & 2 & \cdots & 2(q-s-1) \end{pmatrix}$$

Define $\mathcal{D}_j = \{\mathbf{c} + g \mid \mathbf{c} \in M_j,\ g \in Z_{q-s}\}$, and $\mathcal{C} = \bigcup_{0 \le j \le s} \mathcal{D}_j$.

Thm. 4.5 $\mathcal{C}$ is a $\overline{2}$-SSC$(3, q^2 + sq - 2s^2, q)$.

Coro. 4.6 $\forall\ q \in N,\ \exists\ \overline{2}$-SSC$(3, \frac{1}{8}(9q^2 - w^2), q)$, with $m \equiv q \pmod 8$, and

$$w = \begin{cases} 4 - m, & \text{if } m \equiv 0 \pmod 4, \\ \min\{m, 8 - m\}, & \text{otherwise} \end{cases}$$

Remark: A 2-FPC$(3, M, q)$ can have at most q^2 codewords (Bazrafshan-Tran van Trung, 2008), but the above $\overline{2}$-SSC$(3, M, q)$ can have about $\frac{9}{8}q^2$ codewords. It is even possible to construct $\overline{2}$-SSC$(3, M, q)$ with more codewords.

Problem: What is the largest number of codewords in a $\overline{t}$-SSC (n, M, q)?

Any Questions?

Thanks for Your Attention!