Traitor tracing against powerful attacks using combinatorial designs

McNicol, Simon; Boztas, Serdar; Rao, Asha

https://researchrepository.rmit.edu.au/esploro/outputs/9921859652501341/filesAndLinks?institution=61RMIT_INST&index=null

McNicol, S., Boztas, S., & Rao, A. (2006). Traitor tracing against powerful attacks using combinatorial designs. *Lecture Notes in Computer Science*, 3857, 215–224. <https://doi.org/10.1007/11617983>

Published Version: <https://doi.org/10.1007/11617983>

Repository homepage: <https://researchrepository.rmit.edu.au>

© Springer-Verlag Berlin Heidelberg 2006

Downloaded On 2023/04/27 00:01:03 +1000

Citation:

McNicol, S, Boztas, S and Rao, A 2006, 'Traitor tracing against powerful attacks using combinatorial designs', *Lecture Notes in Computer Science*, vol. 3857, pp. 215-224.

Traitor Tracing Against Powerful Attacks Using Combinatorial Designs

Simon McNicol, Serdar Boztas and Asha Rao

School of Mathematical and Geospatial Sciences,
RMIT University, GPO Box 2476V, Melbourne 3001, Australia
serdar.boztas@ems.rmit.edu.au

Abstract

This paper addresses the problem of threshold traitor tracing for digital content where, by embedding appropriate digital patterns into the distributed content, it is possible to trace and identify the source of unauthorised redistribution.

We use a set of marking assumptions where the adversaries have varying powers change or erase coordinates of the fingerprint where their individual fingerprints differ—and consider the implications.

We consider codes derived from combinatorial designs—and develop a method for concatenating the block design code to filter out the false positives and defend against some of the attacks considered.

Keywords: Traitor tracing, fingerprinting, digital rights management, coding theory, block designs, Reed-Solomon codes.

1 Introduction

Fingerprinting (see [4, 2, 1]) is a technique that aims to prevent the unauthorized redistribution of digital content. Identical or very close copies of a digital document are made available to a large number of authorized users. The locations where the copies differ, are where the fingerprint has been embedded into the digital object. Malicious users attempt to discover the fingerprint and alter it to construct rogue copies which will still “function”. The document is assumed to be a string over a finite alphabet, with the fingerprint being a randomly spread-out substring of the former, but much shorter than the document itself. *Traitor Tracing* (see [5, 3]) schemes enable the tracing of the user(s) whose fingerprints were used to construct the rogue copies. Here we are interested in tracing more powerful attacks—given in the example below.

Example 1. For two coalitions drawn from codewords 01234 and 00224 10323 we exhibit some descendants below:

<i>codeword</i>		<i>codeword</i>		(1)
<i>traitor1</i>	0 1 2 3 4	<i>traitor1</i>	0 1 2 3 4	
<i>traitor2</i>	0 0 2 2 4	<i>traitor3</i>	1 0 3 2 3	
<i>narrow</i>	0 0 2 3 4	<i>narrow</i>	1 1 3 3 3	
<i>wide</i>	0 4 2 # 4	<i>wide</i>	# 0 3 # 4	
<i>erasure</i>	0 # 2 # 4	<i>erasure</i>	# # # # #	
<i>hybrid</i>	0 # 2 3 4	<i>hybrid</i>	1 # 2 3 #	
<i>type</i>	<i>descendant</i>	<i>type</i>	<i>descendant</i>	

In a narrow attack, which has been widely studied, the attackers can choose any symbol which already appears in their copies at that location. Recently Tardos [11] has introduced an efficient solution. In a hybrid attack, the attackers are forced to erase symbols at locations where their copies mismatch. In a hybrid attack, the attackers can switch between the narrow and erasure attacks, position by position. In the wide attack they can choose any symbol from the alphabet or erase the location wherever their copies differ.

We have recently considered this hierarchy of attacks [7, 8] and demonstrated that the wide attack can be reduced to a hybrid attack—by the method of alphabet boosting—though at the cost of a considerable overhead in terms of expanding the codelength. Here, we focus on the details of tracing the hybrid and erasure attacks, after a brief overview of the attack hierarchy.

2 Attacks on Fingerprinting Schemes

The symbol $\mathbf{v}$ denotes a vector; $v^{(i)}$ is its i -th coordinate. We assume the finite alphabet $\mathcal{Q}$ to be F_q , wherever convenient.

Definition 1. A code $\mathcal{M}$ is a subset of words from $\mathbb{F}_q^N$. A linear code is a vector subspace of $\mathbb{F}_q^N$.

2.1 Attack Taxonomy

We identify each user with the unique codeword that she has been assigned. A coalition is a set of attackers. We then ask: (i) *What can a coalition do?* (ii) *How large can a coalition be?*

Definition 2. Let $\mathcal{C}(\mathcal{M}, \omega)$ denote the collection of all possible coalitions of $\mathcal{M}$ with size at most ω : $\mathcal{C}(\mathcal{M}, \omega) = \{A \subseteq \mathcal{M} : |A| \leq \omega\}$.

Definition 3. Let $A \subseteq F_q^N$. The spectrum of A at the z -th coordinate is, $\text{spec}(A, z) = \{v^{(z)} : \mathbf{v} \in A\}$.

Definition 4. Let $\mathcal{T} \in \mathcal{C}(\mathcal{M}, \omega)$. The word $\mathbf{d}$ is a narrow descendant of the coalition $\mathcal{T}$, denoted $\mathbf{d} \in \text{Ndesc}_\omega(\mathcal{T})$ provided, for $1 \leq z \leq N$,

$$\mathbf{d} = (d^{(1)}, \dots, d^{(N)}) \text{ where } d^{(z)} \in \text{spec}(\mathcal{T}, z). \quad (2)$$

Definition 5. For any code $\mathcal{M}$, the set of all narrow descendants from coalitions of size $\leq \omega$ is denoted:

$$\text{NDesc}_\omega(\mathcal{M}) = \bigcup_{\mathcal{T} \in \mathcal{C}(\mathcal{M}, \omega)} \text{Ndesc}_\omega(\mathcal{T}) \quad (3)$$

Note that $\{\mathcal{T} : \mathbf{d} \in \text{Ndesc}_\omega(\mathcal{T})\}$ is the set of all coalitions that could construct the descendant $\mathbf{d}$. The set of codewords that *must* have been used to construct the descendant $\mathbf{d}$ is its Trace:

Definition 6 (Trace). The trace of the descendant $\mathbf{d}$ is:

$$\text{trace}_\omega(\mathbf{d}) = \bigcap_{\{\mathcal{T} : \mathbf{d} \in \text{Ndesc}_\omega(\mathcal{T})\}} \mathcal{T}. \quad (4)$$

2.2 Strength of Codes Resisting the Narrow Attack

It was shown by Staddon et al. [9] that if the minimum distance of $\mathcal{M}$ is $\geq N(1 - \frac{1}{\omega^2})$, any descendant is traceable to at least one traitor, using $\mathcal{M}$. In general, the strength of a code determines the ease of detecting or tracing a coalition (see e.g., [9, 10]). In this paper, we design codes for more general attacks.

Definition 7 (Extended Alphabet). Let $Q^\# = Q \cup \{\#\}$ be the extension of the finite alphabet Q .

Definition 8 (Wide Attack). Let $\mathcal{M}$ have codelength N and alphabet Q . The coalition $\mathcal{T} \in \mathcal{C}(\mathcal{M}, \omega)$ can choose any descendant $\mathbf{d} = (d^{(1)}, d^{(2)}, \dots, d^{(N)})$ where, for $1 \leq i \leq N$:

$$d^{(i)} \in \begin{cases} \text{spec}(\mathcal{T}, i), & \text{if } |\text{spec}(\mathcal{T}, i)| = 1 \\ Q^\# & \text{else.} \end{cases}$$

We define the special cases of the wide attack informally due to space constraints:

The **Erasure Attack**, where a coalition is restricted to converting the i^{th} symbol to $\#$ whenever $|\text{spec}(\mathcal{T}, i)| > 1$.

The **Hybrid Attack**, where for each coordinate, the coalition selects whether they employ an erasure or a narrow attack.

The narrow, erasure and hybrid attacks need to be defended against if we are to defend against the wide attack. The descendant sets $\text{EDesc}_\omega(\mathcal{M})$, $\text{HDesc}_\omega(\mathcal{M})$, and $\text{WDesc}_\omega(\mathcal{M})$ can be defined in analogy with those for the narrow attack and obey:

Lemma 9. $\forall \mathcal{M}$, with $\omega \geq 2$ we have: (i) $\text{NDesc}_\omega(\mathcal{M}) \subset \text{HDesc}_\omega(\mathcal{M})$. (ii) $\text{EDesc}_\omega(\mathcal{M}) \subset \text{HDesc}_\omega(\mathcal{M})$. (iii) $\text{HDesc}_\omega(\mathcal{M}) \subset \text{WDesc}_\omega(\mathcal{M})$. (iv) $\text{NDesc}_\omega(\mathcal{M}) \cap \text{EDesc}_\omega(\mathcal{M}) = \emptyset$.

3 Defending Against the Various Attacks

3.1 Tracing the Erasure Attack

Note that since most codes have codewords that are very far apart, a classical erasure decoding approach won't work here.

Definition 10 (Coalition Erasure Profile). Consider a coalition $\mathcal{T}$, let $E_{\mathcal{T}}$ be defined by, $E_{\mathcal{T}} = \{(i, \text{spec}(\mathcal{T}, i)) : |\text{spec}(\mathcal{T}, i)| = 1\}$.

The erasure profile $E_{\mathcal{T}}$ also obeys $E_{\mathcal{T}} = \bigcap_{f \in \mathcal{T}} E_f$. Note that, $E_f = E_{\{f\}} = \{(i, f^{(i)}) : 1 \leq i \leq N\}$, i.e., E_f is the erasure profile of a coalition containing only the codeword f . Given $\mathbf{d} \in \text{Edesc}(\mathcal{T})$, the erasure profile is simply the sets of pairs $(i, d^{(i)})$ of the descendant that haven't been erased:

$$E_{\mathcal{T}} = \{(i, d^{(i)}) : d^{(i)} \neq \#, 1 \leq i \leq N\}. \quad (5)$$

To construct a code that is resistant to an erasure attack we impose two conditions: (i) $E_{\mathcal{T}}$ must be nonempty for all coalitions $\mathcal{T}$; (ii) To ensure that every coalition can be uniquely defined by $E_{\mathcal{T}}$, we require:

$$\mathcal{T}_1 \neq \mathcal{T}_2 \Leftrightarrow E_{\mathcal{T}_1} \neq E_{\mathcal{T}_2}, \forall \mathcal{T}_1, \mathcal{T}_2 \in \mathcal{C}(\mathcal{M}, \omega). \quad (6)$$

We have an erasure profile analog to the minimum distance $d_{\mathcal{M}}$ of a code $\mathcal{M}$.

Definition 11 (Minimum Erasure Profile). Let $\mathcal{M}$ be a code and let $\mathcal{C}(\mathcal{M}, \omega)$ be given by Definition 2. The minimum erasure profile $\text{epro}_{\mathcal{M}}$ is given by,

$$\text{epro}_{\mathcal{M}} = \min_{\mathcal{T} \in \mathcal{C}(\mathcal{M}, \omega)} |E_{\mathcal{T}}|. \quad (7)$$

Note that, $\text{epro}_{\mathcal{M}}$ may equal zero. For tracing descendants $\mathbf{d} \in \text{EDesc}(\mathcal{M})$, we are only interested in codes with $\text{epro}_{\mathcal{M}} > 0$.

Each component of the erasure profile contains tracing information about the coalition that constructed it.

Definition 12 (Partial Trace). *Given the Erasure profile of a coalition $E_{\mathcal{T}}$, the partial trace for the coalition $\mathcal{T}$ on the pairs $(i, x) \in E_{\mathcal{T}}$ is $B_{i,x} = \{\mathbf{f} \in M : f^{(i)} = x\}$.*

The full trace can be obtained from the partial trace.

Definition 13 (Trace of Erasure Profile). *Given an Erasure profile $E_{\mathcal{T}}$, the intersection of all partial traces gives the list (trace) of all possible codewords that could make up the coalition $\mathcal{T}$.*

$$\text{trace}(E_{\mathcal{T}}) = \bigcap_{(i,x) \in E_{\mathcal{T}}} B_{i,x} \quad (8)$$

For a code to be traceable, we impose the requirement $\text{trace}(E_{\mathcal{T}}) = \mathcal{T}$.

Lemma 14. *The erasure profile $E_{\mathcal{T}}$ is traceable to the coalition $\mathcal{T}$ if $\text{trace}(E_{\mathcal{T}}) = \mathcal{T}$.*

Proof. Let $\mathcal{T}$ be a coalition and let $E_{\mathcal{T}}$ be its corresponding erasure profile. Each codeword $\mathbf{f} \in \mathcal{T}$ must be in $\text{trace}(E_{\mathcal{T}})$ since $\mathbf{f} \cap B_{i,x} \neq \emptyset$ for each $(i, x) \in E_{\mathcal{T}}$. If there is no other codewords in $\text{trace}(E_{\mathcal{T}})$ then, $\text{trace}(E_{\mathcal{T}}) = \mathcal{T}$ and the coalition is traced. $\square$

3.2 Tracing the Hybrid Attack

For a hybrid descendant $\mathbf{d}$, we define the corresponding profile as the “tracing profile”.

Definition 15 (Tracing Profile). *For any descendant $\mathbf{d} \in \text{HDesc}_{\omega}(\mathcal{M})$, the tracing profile of $\mathbf{d}$ is given by,*

$$I_{\mathbf{d}} = \left\{ (i, d^{(i)}) : d^{(i)} \neq \#, 1 \leq i \leq N \right\}. \quad (9)$$

Note that this appears to be the exact same definition as for the erasure profile $E_{\mathcal{T}}$ (see Definition 10). The key difference here is that $E_{\mathcal{T}}$ represents the set of (coordinate, value) that a coalition **can't** erase, where as $I_{\mathbf{d}}$ is the set of (coordinate, value) that a coalition **has not** deleted.

Definition 16 (Trace of hybrid attack). *The trace of a descendant $\mathbf{d} \in \text{HDesc}_\omega(\mathcal{M})$ is the intersection of all coalitions $\mathcal{T} \in \mathcal{C}(\mathcal{M}, \omega)$ which can construct $\mathbf{d}$. A coalition $\mathcal{T}$ can construct $\mathbf{d}$ provided $\mathcal{T} \cap B_{i,z} \neq \emptyset$ for all $(i, z) \in I_{\mathbf{d}}$. The intersection of all such coalitions gives all codewords required to construct the descendant $\mathbf{d}$, i.e.,*

$$\text{trace}(\mathbf{d}) = \bigcap_{\{\mathcal{T}: \mathbf{d} \in \text{Hdesc}(\mathcal{T})\}} \mathcal{T}. \quad (10)$$

It has been noted in [7] that the hybrid attack can be defended against by using codes described in the next section, which have been designed against the erasure attack, for slightly smaller coalition sizes. We don't consider this further in this paper.

3.3 Tracing the Wide Attack

We have shown in [7, 8] that tracing wide descendants is equivalent to tracing hybrid descendants, subject to a failure probability which can be made small enough, at the cost of increasing the codelength. Therefore, we don't consider the wide attack in the sequel.

4 Vector Space Block Design Codes (VSBDCs)

A VSBDC is a type of Resolvable BIBD [12], constructed by using the set of all $(\omega - 1)$ -dimensional quotient vector spaces of a vector space $\mathcal{V}$. We give an example of a VSBDC with $\omega = 3$; See [8] for more.

Example 2. *Let $\mathcal{V} = \mathbb{Z}_2^3$. The VSBDC $\mathcal{M} = \text{BDC}(\mathcal{V}, \mathcal{B}_{\mathcal{V},2})$ is constructed from the set of all vector subspaces of $\mathcal{V}$ with dimension 2, denoted $S_{\mathcal{V},2}$. The codewords in $\text{BDC}(\mathcal{V}, \mathcal{B}_{\mathcal{V},2})$ are*

$$\begin{aligned} \mathbf{v}_{\mathcal{B}}(000) &= (1, 1, 1, 1, 1, 1), & \mathbf{v}_{\mathcal{B}}(001) &= (1, 1, 1, 2, 2, 2), \\ \mathbf{v}_{\mathcal{B}}(010) &= (1, 2, 2, 1, 1, 2), & \mathbf{v}_{\mathcal{B}}(100) &= (2, 1, 2, 1, 2, 1), \\ \mathbf{v}_{\mathcal{B}}(011) &= (1, 2, 2, 2, 2, 1), & \mathbf{v}_{\mathcal{B}}(101) &= (2, 1, 2, 2, 1, 2), \\ \mathbf{v}_{\mathcal{B}}(110) &= (2, 2, 1, 1, 2, 2), & \mathbf{v}_{\mathcal{B}}(111) &= (2, 2, 1, 2, 1, 2). \end{aligned} \quad (11)$$

The set $A_{\mathcal{T}}$, defined below, is the $\mathcal{V}$ -representation of the set of partition vectors in the coalition $\mathcal{T}$:

Definition 17. *Consider a coalition $\mathcal{T}$ of ω codewords from the vector space block design code $\text{BDC}(\mathcal{V}, \mathcal{B}_{\mathcal{V},\omega-1})$. Then $A_{\mathcal{T}}$ consists of all vectors $\mathbf{a}$ in $\mathcal{V}$ that satisfy: $A_{\mathcal{T}} = \{\mathbf{a} : \mathbf{v}_{\mathcal{B}_{\mathcal{V},\omega-1}}(\mathbf{a}) \in \mathcal{T}, \mathbf{a} \in \mathcal{V}\}$.*

Similarly, we can also redefine the partial trace $B_{i,x}$.

Definition 18. Let $\mathcal{M}$ be the code $\text{BDC}(\mathcal{V}, \mathcal{B}_{\mathcal{V},k})$. We now represent the partial trace $B_{i,x}$ (see Definition 12) by the vectors in $\mathcal{V}$ that construct the codewords in $\mathcal{T}$ instead of the codewords themselves. i.e., $B_{i,x} = \left\{ \mathbf{a} : \mathbf{v}_{\mathcal{B}_{\mathcal{V},k}}^{(i)}(\mathbf{a}) = x, \mathbf{a} \in \mathcal{V} \right\}$, where $\mathbf{v}_{\mathcal{B}_{\mathcal{V},k}}^{(i)}(\mathbf{a})$ is the i -th coordinate of the partition vector $\mathbf{v}_{\mathcal{B}_{\mathcal{V},k}}(\mathbf{a})$.

We characterize the partial trace below, but omit the long proof given in [7].

Theorem 19. Let $\mathcal{M}$ be the vector space block design code $\text{BDC}(\mathcal{V}, \mathcal{B}_{\mathcal{V},\omega-1})$, where $\mathcal{V} = F_q^n$. Let $\mathcal{T} \in \mathcal{C}(\mathcal{M}, \omega)$. The partial trace of $E_{\mathcal{T}}$ is the coset generator $[A_{\mathcal{T}}]_{\mathcal{V}}$. i.e.,

$$\text{ptrace}(E_{\mathcal{T}}) = \bigcap_{(i,x) \in E_{\mathcal{T}}} B_{i,x} = [A_{\mathcal{T}}]_{\mathcal{V}}. \quad (12)$$

Note that the VSBDC has the desirable property that any coalition of size ω must have at least one coordinate where all codewords are equal (since every coalition $\mathcal{T}$ must have a non empty erasure profile $E_{\mathcal{T}}$). This property enables a partial trace of the descendant which will contain the coalition plus other codewords (spoof words). We use concatenated codes to get around this problem.

Definition 20 (Concatenated Vector Space Block Design Code). A CVSBDC $\mathcal{M}$ consists of an inner code $\mathcal{M}_I = \text{BDC}(F_q^{N_I}, \mathcal{B}_{F_q^{N_I},k})$, an outer code $\mathcal{M}_O \subset F_q^{N_O}$ and a scalar preserving homomorphism $\varphi : F_q^{N_I} \mapsto (F_q^{N_I}, +)$. Its codewords are constructed as follows:

$$\mathcal{M} = \left\{ \mathbf{v}_{\mathcal{B}} \left(\varphi^{-1}(f^{(1)}) \right), \dots, \mathbf{v}_{\mathcal{B}} \left(\varphi^{-1}(f^{(N_O)}) \right) : \mathbf{f} \in \mathcal{M}_O \right\}$$

where $\mathcal{B}$ is the set of quotient vector spaces $\mathcal{B}_{\mathcal{V},k}$, and $\mathbf{f} = (f^{(1)}, f^{(2)}, \dots, f^{(N_O)})$.

We cannot use a linear code as $\mathcal{M}_O$ as this re-introduces the problem of spoof words at the outer code level. In [8, 7], we have introduced δ -nonlinear codes—in the form of modified GRS codes—to get around this problem. These are codes where the sum of δ or fewer codewords is *not* a codeword, and are obtained by applying a special type of permutation polynomial to Generalized Reed Solomon (GRS) codewords—see [6] for details on GRS codes—which are MDS codes attaining the Singleton bound, and have length N , dimension k and minimum distance $N - k + 1$.

The δ -nonlinearity property ensures that the partial trace of the concatenated code would only contain the coalition $\mathcal{T}$ plus other non-codewords which can be discarded leaving only the coalition. See [7] for the proof, as well as the example of a concatenated vector space block design code (CVSBDC) below.

Example 3. Let $\mathcal{V}$ be the vector space $\mathbb{F}_2^2$. There are 3 vector subspaces $\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3$ of $\mathcal{V}$ with dimension 1. The set of all quotient vector spaces with dimension 1, $\mathcal{B}_{\mathcal{V},1}$, is:

$$\mathcal{B}_{\mathcal{V},1} = \left\{ \underbrace{\left\{ \overbrace{\{00, 01\}}^{\mathcal{H}_1}, \{10, 11\} \right\}}_{\mathcal{V}/\mathcal{H}_1}, \underbrace{\left\{ \overbrace{\{00, 10\}}^{\mathcal{H}_2}, \{01, 11\} \right\}}_{\mathcal{V}/\mathcal{H}_2}, \underbrace{\left\{ \overbrace{\{00, 11\}}^{\mathcal{H}_3}, \{01, 10\} \right\}}_{\mathcal{V}/\mathcal{H}_3} \right\}.$$

From $\mathcal{B}_{\mathcal{V},1}$, we generate the inner code $\mathcal{M}_I = \text{BDC}(\mathcal{V}, \mathcal{B}_{\mathcal{V},1})$;

$$\mathcal{M}_I = \{\mathbf{v}_{\mathcal{B}_{\mathcal{V},1}}(00), \mathbf{v}_{\mathcal{B}_{\mathcal{V},1}}(01), \mathbf{v}_{\mathcal{B}_{\mathcal{V},1}}(10), \mathbf{v}_{\mathcal{B}_{\mathcal{V},1}}(11)\} = \{111, 122, 212, 221\}.$$

Here $\mathbf{v}_{\mathcal{B}_{\mathcal{V},1}}(01) = 122$ is the partition vector of $01 \in \mathcal{V}$. Now, let $\mathcal{M}_O \subseteq \mathbb{F}_4^4$ be the outer code with a codeword $\mathbf{f} = (0, \alpha, 1, \alpha^2)$ (α is a primitive element of $\mathbb{F}_4$).

A suitable choice of scalar preserving homomorphism $\varphi : \mathbb{F}_2^2 \longrightarrow (\mathbb{F}_4, +)$ is:

$x \in \mathbb{F}_2^2$	$\varphi(x) \in \mathbb{F}_{2^2}$	$\varphi(x) \text{ as } \alpha^i$
00	0	0
01	1	α^0
10	α	α^1
11	$\alpha + 1$	α^2

(13)

We construct the concatenated codeword as follows: The first coordinate of $\mathbf{f}$ is $f^{(1)} = 0$, applying the inverse of φ we get the vector 00. The corresponding partition vector is $\mathbf{v}_{\mathcal{B}_{\mathcal{V},1}}(00) = 111$. The second, third and fourth coordinates give us the partition vectors 212, 122 and 221. The concatenated codeword is then (111, 212, 122, 221). We can apply this process to all the codewords in $\mathcal{M}_O$ to construct $\mathcal{M}$.

Given a concatenated code $\mathcal{M}$, and a descendant $\mathbf{d} \in \text{EDesc}_\omega(\mathcal{M})$, what the partial trace of the erasure profile? If we view the concatenated code at the inner level, then the partial trace of the concatenated code is simply the sequence of partial traces for each inner code segment. We call this the “concatenated partial trace”.

Definition 21 (Concatenated Partial Trace). Let $\mathcal{M}$ be a CVSBDC with outer code $\mathcal{M}_O$ and inner code $\mathcal{M}_I$. Let $\mathcal{T} \in \mathcal{C}(\mathcal{M}, \omega)$ be a coalition of the concatenated code. Let $\mathcal{T}_O$ be the codewords in $\mathcal{M}_O$ that generate the concatenated codewords in $\mathcal{T}$. Let $\mathcal{T}^{(i)} \in \mathcal{C}(\mathcal{M}_I, \omega)$ be the coalition of inner codewords that are inserted into each codeword of $\mathcal{T}_O$ at the i -th coordinate. The concatenated partial trace of the erasure profile $E_{\mathcal{T}}$ denoted $\text{Cptrace}(E_{\mathcal{T}})$, is the sequence of N_O partial traces for each inner coalition $\mathcal{T}^{(i)}$. i.e., $\text{Cptrace}(E_{\mathcal{T}}) = (\text{ptrace}(E_{\mathcal{T}^{(1)}}), \dots, \text{ptrace}(E_{\mathcal{T}^{(N_O)}}))$.

Note that, if the inner code $\mathcal{M}_I$ is a VSBDC, then by Theorem 19 the concatenated partial trace is $\text{Cptrace}(E_{\mathcal{T}}) = ([\mathcal{T}^{(1)}]_{\mathbb{F}_q^n}, \dots, [\mathcal{T}^{(N_O)}]_{\mathbb{F}_q^n})$. The concatenated partial trace is useful for viewing the code at the inner level; at the outer level, we need to use a “coset image”.

Definition 22 (Coset Image). Consider $\mathbb{F}_q^n$ and $\mathbb{F}_{q^n}$. For a given scalar preserving homomorphism $\varphi : \mathbb{F}_q^n \rightarrow (\mathbb{F}_{q^n}, +)$ and a given subset $A \subseteq \mathbb{F}_{q^n}$, a coset image $[A]_{\varphi}$ is $[A]_{\varphi} = \varphi([\varphi^{-1}(A)]_{\mathbb{F}_q^n})$ is:

$$[A]_{\varphi} = \left\{ a + \sum_{b \in A} \lambda_b(b - a) : \lambda_b \in \mathbb{F}_q, b \in A \right\}, \text{ for any } a \in A, \quad (14)$$

where $\varphi^{-1}(A) = \{\varphi^{-1}(a) : a \in A\} \subset \mathbb{F}_q^n$.

The components λ_b are over the subfield $\mathbb{F}_q$ and not the field $\mathbb{F}_{q^n}$. The coset image contains $\leq q^{|A|-1}$ elements of $\mathbb{F}_{q^n}$. We can show that we can construct the concatenated partial trace by using a coset image on the outer code $\mathcal{M}_O$, see [7] for the proof.

Theorem 23. Let $\mathcal{M}$ be a CVSBDC as defined in Definition 20. Let $\mathcal{T} \in \mathcal{C}(\mathcal{M}, \omega)$ be a coalition and let $\mathcal{T}_O$ and $\mathcal{T}^{(i)}$ be as defined in Definition 20. The coset image $[\mathcal{T}_O]_{\varphi}$ can generate every partial trace of the erasure profiles $E_{\mathcal{T}^{(i)}}$ via $\text{ptrace}(E_{\mathcal{T}^{(i)}}) = [\mathcal{T}^{(i)}]_{\mathbb{F}_q^n} = \varphi^{-1}(\text{spec}([\mathcal{T}_O]_{\varphi}, i))$.

5 Conclusions

We have investigated *powerful attacks against fingerprinting schemes* used for securing digital content and described *new types of codes which can be used to defend against such attacks*. We have made use of resolvable BIBDs (Balanced Incomplete Block Designs) in order to construct new code families we call VSBDC (Vector Space Block Design) codes. These codes, used with concatenation techniques, are resistant to narrow, erasure, and as far as experimentally observed, to hybrid attacks.

There are efficient algorithms for decoding Reed Solomon codes, which could be applied to the δ -nonlinear Generalized Reed Solomon codes. The algorithms for decoding the BIBD based component of the concatenated codes could be based on linear algebra operations for computing cosets and subspaces, and thus have a priori complexity no worse than a small power of the codelength, which is given by an appropriate q -binomial coefficient (see [7]). It remains an open problem to obtain more efficient decoding algorithms for the design based codes.

The main contribution of this paper is an explicit-nonrandomized-coding construction for addressing erasure and other attacks in digital fingerprinting.

Acknowledgment

The first author was supported by an APA scholarship from the ARC.

References

1. A. Barg, G.R. Blakely, and G.A. Kabatiansky. Digital finger printing codes: problem statements, constructions, identification of traitors. *IEEE Trans. Inform. Theory*, 49:852–865, 2003.
2. A. Barg, G. Cohen, S. Encheva, G. Kabatiansky, and G. Zemor. A hypergraph approach to the identifying parent property: the case of multiple parents. *SIAM Journal of Discrete Math.*, 14(3):423–431, 2001.
3. D. Boneh and M. Franklin. An efficient public key traitor tracing scheme. *Lecture Notes in Computer Science*, 1666:338–353, 1999.
4. D. Boneh and J. Shaw. Collusion-secure fingerprinting for digital data. *Lecture Notes in Computer Science*, 963:452–465, 1995.
5. B. Chor, A. Fiat, and M. Naor. Tracing traitors. *Lecture Notes in Computer Science*, 839:257–270, 1999.
6. F.J. MacWilliams and N.J.A. Sloane. *The Theory of Error Correcting Codes*. North-Holland Publishing Company., 1977.
7. S. McNicol. *Traitor Tracing Using Generalized Reed-Solomon Codes and Combinatorial Designs*. PhD thesis, RMIT University, July 2005.
8. S. McNicol, S. Boztaş, and A. Rao. Traitor tracing against powerful attacks. *Proceedings of the IEEE International Symposium on Information Theory*, pages 1878–1882, 2005.
9. J. Staddon, D. R. Stinson, and R. Wei. Combinatorial properties of frameproof and traceability codes. *IEEE Transactions on Information Theory*, 47:1042–1049, 2001.
10. D. Stinson, T. van Trung, and R. Wei. Secure frameproof codes, key distribution patterns, group testing algorithms and related structures. *Planning and Inference*, 86(2):595–617, 2000.
11. Gábor Tardos. Optimal probabilistic fingerprint codes. In *STOC '03: Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, pages 116–125, New York, NY, USA, 2003. ACM Press.
12. W.D. Wallis. *Combinatorial Designs*. Marcel Decker Inc., 1988.