Traitor Tracing Against Powerful Attacks

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—We describe narrow, erasure, hybrid and wide attacks on fingerprinting schemes used for securing digital content, and introduce new classes of codes, namely vector space block design codes and δ–nonlinear codes, which can be used to defend against such attacks by means of traitor tracing.

I. INTRODUCTION

In this paper we describe different attacks which can be mounted against a digital fingerprinting system. We then discuss the general problem of tracing descendants under the various attacks and describe techniques that can reduce the wide attack problem to solving a hybrid attack.

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 document–software, images, other digital media–are made available to a large number of users of the system by paid subscription. The parts where the fingerprints differ, are where the fingerprint has been embedded into the digital object, and it is the aim of malicious users to discover these parts and construct rogue copies which will still “function”. Traitor Tracing (see [5], [3]) schemes enable the tracing of the identity of the user whose fingerprints were used to construct rogue copies of the digital content, so that they can be prosecuted.

Unless otherwise stated, we always consider a finite alphabet $Q$, usually $F_q$, the finite field of $q$ elements.

II. THE NARROW ATTACK

A. Coalitions

A coalition is a collection of users who wish to perform a dishonest task within the system. We identify each user with the codeword that he/she has been assigned–thus the code $M$ represents the collection of all users. There are two restrictions that we will place on coalitions: (i) What can a coalition do? (ii) How large can a coalition be?

Definition 1 (Coalition): Let $C(M, \omega)$ denote the collection of all possible coalitions of $M$ with size at most $\omega$:

$$C(M, \omega) = \{A \subseteq M : |A| \leq \omega\}. \tag{1}$$

B. Descendants

A descendant is a word that has been constructed by the ordered cutting and pasting of a collection of codewords from $M$.

Example 1: Consider a coalition containing two codewords $(0,1,2,3)$ and $(3,1,0,3)$. The set of all possible descendants from this coalition is $\{(0,1,0,3), (0,1,2,3), (3,1,0,3), (3,1,2,3)\}$ or by slight abuse of notation $\{0,3\} \times \{1\} \times \{0,2\} \times \{3\}$, and include the original codewords themselves.

A fundamental measure concerned with sharing properties between sets of codewords, is the spectrum.

Definition 2 (Spectrum): Let $A \subseteq F_q^N$ be a collection of vectors $v = (v^{(1)}, \ldots, v^{(N)})$. The spectrum of $A$ at the $z$-th coordinate is,

$$\text{spec}(A, z) = \{v^{(z)} : v \in A\}. \tag{2}$$

Definition 3 (Narrow Descendant): Let $M$ be a code with a coalition $T \in C(M, \omega)$. A narrow descendant $d$ is any word that is constructed pointwise from a coalition by selecting symbols from the spectrum:

$$d^{(z)} \in \text{spec}(T, z), \quad 1 \leq i \leq N. \tag{3}$$

The word $d$ is a narrow descendant of the coalition $T$, denoted $d \in \text{Ndesc}_\omega(T)$ provided, for $1 \leq z \leq N$, $d = (d^{(1)}, \ldots, d^{(N)})$ where $d^{(z)} \in \text{spec}(T, z). \tag{4}$

Definition 4: For any given code $M$, the set of all narrow descendants from coalitions of size at most $\omega$, is denoted:

$$\text{Ndesc}_\omega(M) = \bigcup_{T \in C(M, \omega)} \text{Ndesc}_\omega(T) \tag{5}$$

Given only a descendant, how can we find a coalition that could have constructed it?

Example 2: Let $M$ be the code $\{0000,1111,0123,2301\}$. We are supplied with a descendant 0123, which is different from the codewords in $M$. However, we can trace the descendant back to a coalition. The set of coalitions that could have constructed the descendant is determined by finding the smallest coalition that contains the digit 0123. We can determine this by checking the spectrum:

$$\text{spec}(T, 1) = \{0,3\}, \quad \text{spec}(T, 2) = \{1\}, \quad \text{spec}(T, 3) = \{0,2\}, \quad \text{spec}(T, 4) = \{3\}.$$
C. Strength of Codes Resisting the Narrow Attack

It was shown by Staddol et al. [8] that if the minimum distance of \( M \) is \( \geq N(1 - 1/q) \), any descendant is traceable to at least one traitor, using \( M \). In general, the strength of a code determines the ease of detecting or tracing a coalition (see e.g., [8], [9], [10]).

**Definition 6 (Tracing Hierarchy):** Let \( M \) be a code and let \( d \in \text{NDesc}_\omega(M) \). If \( M \) has a tracing or detecting property, it can be classified into one of four categories (from weakest to strongest):

- **\( \omega\)-FP (\( \omega\)-Frameproof):** No coalition can recreate a valid codeword belonging to a user not in the coalition;
- **\( \omega\)-SFP (\( \omega\)-Secure Frameproof):** No two disjoint coalitions can create the same descendant;
- **\( \omega\)-IPP (\( \omega\)-Identifiable Parent Property):** Any descendant can be traced to at least one codeword;
- **\( \omega\)-TA (\( \omega\)-Traceable Ancestor):** The codeword that is nearest to the given descendant is a traitor.

**Theorem 1 ([10]):** The following hierarchy of holds for any code:

\[
\omega\text{-TA} \Rightarrow \omega\text{-IPP} \Rightarrow \omega\text{-SFP} \Rightarrow \omega\text{-FP}. \tag{7}
\]

For a descendant to be traceable to at least one traitor, the code must be designed to be \( \omega\)-IPP or better.

As the number of codewords gets large, the number of descendants gets even larger, roughly \( O(\lvert M\rvert^N) \) (for small \( \omega \)), where \( N \) is the codewidth and \( \lvert M \rvert \) is the number of users for the system. The brute-force attack approach to tracing traitors thus quickly becomes computationally infeasible. McNicol and Boztas\textsuperscript{7} [7] have designed efficient randomized traitor tracing algorithms for the narrow attack.

D. Elementary Descendants

Note that a linear code \( M \) has the following property:

\[
f \in M, \ d \in \text{NDesc}_\omega(M) \Rightarrow d + f \in \text{NDesc}_\omega(M). \tag{8}
\]

We shall show that, more generally, descendants can be related by their *trace* which is useful for constructing an efficient tracing algorithm whereby manipulating a descendant \( d \) with a set of operations we obtain an *elementary descendant* \( d' \). Then, a lookup table can be constructed in advance to determine the trace for \( d' \).

**Definition 7 (Descendant Permutations):** Let \( M \) be a linear code, i.e., a vector subspace of \( F_q^N \). We define a class of mappings \( \Pi : F_q^N \rightarrow F_q^N \) as follows:

1) Every \( \pi \in \Pi \) is a bijection on \( F_q^N \). That is, \( \pi^{-1} \in \Pi \) exists.
2) For all codewords \( f \in M, \ \pi \in \Pi \Rightarrow \pi(f) \in M \).
3) For every descendant \( d \in \text{NDesc}_\omega(T) \) and every coalition \( T \),

\[
\pi \in \Pi \Rightarrow \pi(d) \in \text{NDesc}_\omega(\pi T) \tag{9}
\]

Thus \( \Pi \) is the set of bijections which leave the code \( M \) and the set of descendants \( \text{NDesc}_\omega(M) \) invariant.

**Theorem 2:** Let \( M \subset F_q^N \) be a linear code and \( \pi \in \Pi \) be a descendant permutation. Then, the trace of the permuted descendant is the permuted trace of the descendant, i.e.,

\[
\forall \pi \in \Pi, \forall d \in \text{Desc}_\omega(M), \ \text{trace}_\omega(\pi d) = \pi \text{trace}_\omega(d).
\]

We now state our main result on descendant permutations.

**Theorem 3:** Let \( M \) be a linear code with \( N = \lvert F_q \rvert \). The mapping \( \pi : v(x) \mapsto \lambda v(ax + b) + g(x) \) where \( v \in F_q^N \) and \( g \in M, \) with \( a, b \in F_q^*, \) and \( b \in F_q^* \) is a descendant permutation. Here, \( v(x) \) (resp. \( g(x) \)) is obtained by Lagrange interpolation from \( v \) (resp. \( g \)).

Before an example, we introduce “square” Generalized Reed Solomon Codes (SGRSs):

**Definition 8 (Generalized Reed Solomon Code [6]):** Let \( X = (x_1, x_2, \ldots, x_N) \), where \( x_i \) are distinct elements of elements of \( F_q \), and let \( A = (a_1, a_2, \ldots, a_N) \), where the \( a_i \) are nonzero (but not necessarily distinct) elements of \( F_q \). Then the generalized Reed-Solomon code \( GRS_k(X, A) \), consist of all vectors,

\[
f = (a_1 f(x_1), a_2 f(x_2), \ldots, a_N f(x_N)), \tag{10}
\]

where \( f(x) \) ranges over all polynomials of degree \( < k \) with coefficients from \( F_q^\ast \).

A \( GRS_k(X, A) \) is an \([N, k] \) code over \( F_q \) with minimum distance \( N - k + 1 \) and hence is MDS. In the special case where \( X = F_q^N \) and \( A = (1, 1, \ldots, 1) \), we obtain a square GRS (SGRS) code, denoted \( SGRS_k(F_q) \) with codewidth \( N = q \). We now give an example of a 2-TA code which contains only 2 nontrivial descendants.

**Example 3:** The code \( M = GRS_1(F_q^5) = \{ f(x) : \deg f \leq 1 \} \), where the codewords are constructed by evaluation on the ordered set \( X_M = \{0, 1, 2, 3, 4\} \). Any of the \( 5^5 = 78125 \) words from \( F_q^5 \) corresponds to a polynomial of degree at most 4. How many are descendants? Consider the monomials \( x^i \), and their evaluations into words:

\[
\begin{array}{c|c|c}
   x & v & \text{trace}_2(v) \\
   \hline
   x^4 & 01111 & \{111111\} \\
   x^3 & 01324 & \{01321, 01234\} \\
   x^2 & 01441 & \{\} \\
   x^1 & 01234 & \{01234\} \\
   1 & 11111 & \{11111\}.
\end{array} \tag{11}
\]

From Theorem 3 we can show that all polynomials \( \lambda(ax + b)^i + cx + d \) are descendants, for \( i = 3, 4 \), with \( cx + d \) forming the codewords. However, the polynomials \( \lambda(ax + b)^2 + cx + d \) with \( \lambda \neq 0 \) are not descendants. Therefore only \( 4 \times 5 \times 5 = 100 \) words are not descendants and \( 3125 - 100 = 3025 \) are descendants—each can be related to one of the 4 allowed monomials except for \( x^2 \). Given any descendant \( d \), we can efficiently construct its trace by constructing a descendant permutation \( \pi \) that maps \( d \) to one of the monomials \( x^3 \) or \( x^4 \). If we apply the inverse \( \pi^{-1} \) to the corresponding trace, we get the trace of the descendant \( d \).
III. MORE POWERFUL ATTACKS

A. Attack Classification

We now define the wide attack and also define its’ special case sub-attacks.

Definition 9 (Extended Alphabet): For any finite alphabet $Q$, we define $Q^* = Q \cup \{\#\}$.

Definition 10 (Wide Attack): Let $M$ have codelength $N$ and alphabet $Q$. The coalition $T \in C(Q, \omega)$ can choose any descendant $d = (d(1), d(2), \ldots, d(N))$ where, for $1 \leq i \leq N$:

$$d(i) \in \{\text{spec}(T, i), \text{if } |\text{spec}(T, i)| = 1 \}
\text{else,}$$

Since traitors can identify part of the fingerprint by checking the differences in each other’s copy of the digital object, this model is reasonable. We define the special cases informally due to space constraints:

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

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

Example 4: For a coalition with codewords 01234 and 00224 we exhibit some descendants below:

<table>
<thead>
<tr>
<th>codeword</th>
<th>traitor1</th>
<th>traitor2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 1 2 3 4</td>
<td>narrow</td>
<td>narrow</td>
</tr>
<tr>
<td>0 0 2 3 4</td>
<td>wide</td>
<td>wide</td>
</tr>
<tr>
<td>0 1 2 0 4</td>
<td>erasure</td>
<td>erasure</td>
</tr>
<tr>
<td>0 4 2 # 4</td>
<td>hybrid</td>
<td>hybrid</td>
</tr>
<tr>
<td>0 0 2 3 4</td>
<td>descendant type</td>
<td></td>
</tr>
</tbody>
</table>

The narrow, erasure and hybrid attacks are special cases of the full wide attack and need to be defended against if we are to be able to defend against the wide attack. The descendant sets $E_{\text{desc}}(Q)|Q\rangle$, $H_{\text{desc}}(Q)|Q\rangle$, and $W_{\text{desc}}(Q)|Q\rangle$ can be defined in analogy with Definitions 3, and 4 and obey:

Lemma 4: For any code $M$, with $\omega \geq 2$ we have:

1) $N_{\text{desc}}(Q)|Q\rangle \subset H_{\text{desc}}(Q)|Q\rangle$.
2) $E_{\text{desc}}(Q)|Q\rangle \subset H_{\text{desc}}(Q)|Q\rangle$.
3) $H_{\text{desc}}(Q)|Q\rangle \subset W_{\text{desc}}(Q)|Q\rangle$.
4) $N_{\text{desc}}(Q)|Q\rangle \cap E_{\text{desc}}(Q)|Q\rangle = \emptyset$.

B. Attack Reduction by Alphabet Boosting

Assuming that a code $M$ is resistant to any combination of erasure and narrow attacks, we can generate a new code $M^*$ that is resistant to a wide attack:

Definition 11 (Alphabet Boosting): Let $f$ be a word with length $N$ over $Q$. Let $\xi = \{\xi_i\}_{i=1}^N$ be a family of $N$ injective mappings $\xi_i: Q \rightarrow Q^*$ where $|Q^*| > |Q|$. The alphabet boosted word $f^*$ is constructed by,

$$f^* = \xi(f) = (\xi_1(f(1)), \xi_2(f(2)), \ldots, \xi_N(f(N))).$$

This still a word of length $N$ but the symbols are now restricted to a subset of the alphabet $Q^*$.

Definition 12 (Alphabet Boosted Code): Let $M$ be a code of length $N$ and alphabet $Q$. Given the boosting map $\xi$, the corresponding boosted code $M^*$ is the set of all words $\xi(f)$ with $f \in M$.

$$M^* = \{\xi(f) : f \in M\}$$

We keep the mappings $\xi$ secret to limit the knowledge of any $\omega$-coalition to $\leq \omega$ legitimate characters in the boosted alphabet $Q^*$. Alphabet boosting doesn’t affect the distance distribution or the resistance of $M$ to a narrow attack. If the coalition employs a wide attack, with a certain probability the character they substitute won’t be a valid character, and therefore can be considered to be an erasure. By choosing $Q^*$ large enough we can reduce the wide attack to an erasure attack.

Lemma 5: Let $M$ be a code over $Q$ and $M^*$ be the corresponding boosted code over $Q^*$. The probability that a coalition of size $\omega$ will select a valid character that is not in its spectrum obeys:

$$p_t \leq \frac{|Q| - \omega}{|Q^*| - \omega}$$

For the moment, we will assume that a coalition will be successful in creating a descending that is untraceable if any of the coordinates are correctly spoofed, i.e., mapped to valid symbols.

Theorem 6 (Wide Attack Reduction): Let $d \in W_{\text{desc}}(T)|Q\rangle$ be a wide descending constructed from a coalition $T$. The probability $p_H$ that $d \in H_{\text{desc}}(T)|Q\rangle$ is at least,

$$p_H \geq (1 - p_t)^N.$$ 

Therefore, $d$ can be traced using hybrid attack tracing techniques with probability $p_H$.

The proof is omitted due to space limitations. The main implication of this theorem is that by varying $p_t$, i.e., by varying the boosted alphabet size, a wide attack can be reduced to a hybrid attack with a certain probability that we can choose.

C. Tracing

For this subsection we will assume that a coalition will solely perform an erasure attack.

Definition 13 (Coalition Erasure Profile): Consider a coalition $T$, let $E_T$ be defined by,

$$E_T = \{(i, \text{spec}(T, i)) : |\text{spec}(T, i)| = 1\}.$$ 

The erasure profile $E_T$ can also be constructed from the intersection of the profiles from each codeword in $T$:

$$E_T = \bigcap_{f \in T} E_f.$$ 

Note that, the profile $E_f$ is the set of pairs \{($i, f(i) : 1 \leq i \leq N$), i.e., the erasure profile of a coalition containing only the codeword $f$. Given a descendant $d$ constructed under an erasure attack, the erasure profile is simply the sets of pairs $(i, d(i))$ of the descendant that haven’t been erased:

$$E_T = \{(i, d(i)) : d(i) \neq \#, 1 \leq i \leq N\}.$$ 

Authorized licensed use limited to: IEEE Xplore. Downloaded on November 17, 2008 at 19:26 from IEEE Xplore. Restrictions apply.
To construct a code that is resistant to an erasure attack we impose two conditions: (i) $E_T$ must be nonempty for all coalitions $T$; (ii) To ensure that every coalition can be uniquely defined by $E_T$, we require the code $\mathcal{M}$ to have the property:

$$T_1 \neq T_2 \iff E_{T_1} \neq E_{T_2}, \forall T_1, T_2 \in \mathcal{C}(\mathcal{M}, \omega). \quad (20)$$

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

**Definition 14 (Partial Trace):** Given the Erasure profile of a coalition $E_T$, the partial trace denoted $B_{i,x}$ for the coalition $T$ on the pairs $(i, x) \in E_T$ is:

$$B_{i,x} = \{ f \in M : f^{(i)}(x) = x \}. \quad (21)$$

The full trace can be obtained from the partial trace.

**Definition 15 (Trace of Erasure Profile):** Given an Erasure profile $E_T$, the intersection of all partial traces gives the list (trace) of all possible codewords that could make up the coalition $T$.

$$\text{trace}(E_T) = \bigcap_{(i,x) \in E_T} B_{i,x} \quad (22)$$

For a code to be traceable, we impose the requirement

$$\text{trace}(E_T) = T.$$ 

**Lemma 7:** The erasure profile $E_T$ is traceable to the coalition $T$ if $\text{trace}(E_T) = T$.

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

This concludes our discussion of the erasure attack. We omit the discussion of the details of the hybrid attack and the hybrid trace due to space constraints.

**IV. VECTOR SPACE BLOCK DESIGN CODES**

We now construct a VSBD, which is a type of Resolvable BIBD, by using the set of all $(\omega - 1)$-dimensional quotient vector spaces of a vector space $V$. Instead of a detailed definition, we give an example of a VSBD with $\omega = 3$.

**Example 5:** Let $V$ be the vector space $Z_2^2$. The VSBD $\mathcal{M} = \text{BDC}(V, B_{V,2})$ is constructed from the set all vector subspaces of $V$ with dimension 2, i.e. $S_{V,2}$. There are 7 of these; the ordered set $S_{V,2} = \{ H_1, H_2, H_3, H_4, H_5, H_6, H_7 \}$ is:

$$\begin{align*}
&\{ (001, 010), (001, 100), (001, 110), (010, 100), \\
&\{ (010, 101), (100, 011), (011, 101), \}
\end{align*} \quad (23)$$

From $S_{V,2}$, we can construct all the quotient vector spaces $B_{V,2}$:

$$\begin{align*}
V/H_1 &= \{ (000, 000, 010, 011), (010, 010, 111, 111) \}, \\
V/H_2 &= \{ (000, 001, 100, 101), (010, 011, 110, 111) \}, \\
V/H_3 &= \{ (000, 001, 110, 111), (010, 100, 011, 101) \}, \\
V/H_4 &= \{ (000, 010, 100, 110), (001, 011, 101, 111) \}, \\
V/H_5 &= \{ (000, 010, 101, 111), (001, 100, 110, 011) \}, \\
V/H_6 &= \{ (000, 100, 111, 111), (001, 010, 110, 111) \}, \\
V/H_7 &= \{ (001, 011, 110, 111), (001, 010, 100, 111) \}.
\end{align*} \quad (24)$$

The set of quotient vector spaces $B_{V,2}$ form a block design, say $\mathcal{B}$.

We can construct the codewords from the vectors in $V$. For the vector $110 \in V$, the codeword $v_\mathcal{B}(110)$ is obtained as follows. First coordinate: the vector $110$ is in the second coset of $V/H_1$; therefore the first coordinate will be 2. Similarly for coordinates 2, 5 and 6. The other coordinates all have a value of 1.

$$v_\mathcal{B}(110) = (2, 2, 1, 1, 2, 1, 2, 1). \quad (25)$$

Following the same procedure, the complete list of codewords in $B_{V,2}$ is:

$$\begin{align*}
v_\mathcal{B}(000) &= (1, 1, 1, 1, 1, 1, 1, 1), \\
v_\mathcal{B}(001) &= (1, 1, 1, 1, 1, 2, 2, 2), \\
v_\mathcal{B}(010) &= (1, 2, 2, 1, 1, 2, 2), \\
v_\mathcal{B}(100) &= (2, 1, 2, 1, 1, 2, 2), \\
v_\mathcal{B}(011) &= (1, 2, 2, 2, 1, 1, 1), \\
v_\mathcal{B}(101) &= (2, 1, 2, 1, 1, 1, 1), \\
v_\mathcal{B}(110) &= (2, 2, 1, 1, 1, 1, 1, 1), \\
v_\mathcal{B}(111) &= (2, 2, 1, 1, 1, 1, 1, 1).
\end{align*} \quad (26)$$

**Definition 16:** Consider a coalition $T$ of $\omega$ codewords from the vector space block design code $BDC(V, B_{V,2})$. The set $A_T$ consists of all vectors $a$ in $V$ that satisfy:

$$A_T = \{ a : v_{\mathcal{B},V}^{(i)}(a) \in T, \ a \in V \}. \quad (27)$$

The set $A_T$ is the $V$ representation of the set of partition vectors in the coalition $T$. Similarly, we can also redefine the partial trace $B_{i,x}$.

**Definition 17:** Let $\mathcal{M}$ be the code $BDC(V, B_{V,2})$. We now represent the partial trace $B_{i,x}$ (see Definition 14) by the vectors in $V$ that construct the codewords in $T$ instead of the codewords themselves. i.e.,

$$B_{i,x} = \{ a : v_{\mathcal{B},V}^{(i)}(a) = x, \ a \in V \}, \quad (28)$$

where $v_{\mathcal{B},V}^{(i)}(a)$ is the $i$-th coordinate of the partition vector $v_{\mathcal{B},V}(a)$.

We characterize the partial trace below.

**Theorem 8:** Let $\mathcal{M}$ be the vector space block design code $BDC(V, B_{V,2})$, where $V = F_q^n$ Let $T \in C(\mathcal{M}, \omega)$ be a coalition of $\mathcal{M}$. The partial trace of the erasure profile $E_T$ is the coset generator $[A_T]_V$, i.e.,

$$\text{ptrace}(E_T) = \bigcap_{(i,x) \in E_T} B_{i,x} = [A_T]_V. \quad (29)$$

Note that the VSBD has the desirable property that any coalition of size $\omega$ must have at least one coordinate where
all codewords are equal (since every coalition $T$ must have a non empty erasure profile $E_T$).

This property enables a partial trace of the descendant which will contain the coalition plus other codewords (spoofer words). We use concatenated codes to get around this problem.

**Definition 18 (Concatenated Vector Space Block Design Code):** A CVSBDC $M$ consists of an inner code $M_f = BDC(F_q^{N_1}, B_{F_q^{N_1}, k})$, an outer code $M_O \subseteq F_q^{N_2}$ and a scalar preserving homomorphism $\varphi : F_q^{N_1} \rightarrow (F_q^{N_2}, +)$, i.e., a homomorphism $\varphi$ which satisfies:

$$\varphi(a + b) = \varphi(a) + \varphi(b), \quad \varphi(\lambda a) = \lambda \varphi(a).$$

Its codewords are constructed as follows:

$$M = \{ vB(\varphi^{-1}(f(1))), \ldots, vB(\varphi^{-1}(f(N_2))) : f \in M_O \}$$

where $B$ is the set of quotient vector spaces $B_{\psi(k)}$, and $f = (f_1, f_2, \ldots, f(N_2))$.

We cannot use a linear code as $M_O$ as this re-introduces the problem of spoofer words at the outer code level (the partial trace of a descendant will contain the coalition plus some linear combinations of that coalition). We introduce $\delta$-nonlinear codes to get around this problem.

**V. MODIFIED GENERALIZED REED SOLOMON CODES**

**Definition 19 ($\delta$-Nonlinear Codes):** A code $M$ is $\delta$-nonlinear if for all subsets $S \subseteq M$ with $1 < |S| \leq \delta$ the sum $\sum_{f \in S} f$ is not another codeword.

Equivalently, if $M$ is defined over an alphabet with $\delta$ characteristic, for all $S \subseteq M$ with $1 < |S| \leq \delta + 1$, the sum $\sum_{f \in S} f$ is nonzero if $M$ is $\delta$-nonlinear. We give an algebraic construction for such a code.

**Definition 20 (Modified GRS Code):** Let $X = (x_1, x_2, \ldots, x_N)$, where $x_i$ are distinct elements of $F_q$, and let $\Psi : F_q^{N_1} \rightarrow F_q^{N_2}$ be a modified Reed-Solomon code, denoted by $MGRS_k(F_q, X, \Psi)$, consists of all vectors,

$$f = (\Psi(f)(x_1), \Psi(f)(x_2), \ldots, \Psi(f)(x_N)),$$

where $f(x)$ ranges over all polynomials of degree $\leq k$ with coefficients from $F_q$.

We can preserve the minimum distance of the original code, if we are selective about the mapping $\Psi$. Choosing a permutation polynomial for $\Psi$ will maintain the distance distribution, but won't necessarily give us the desired nonlinearity property. However, we have shown that the mapping $\Psi : f(x) \rightarrow \sigma(f(x) \cdot x)/x$ where $\sigma$ is a permutation polynomial, not only preserves the distance between all modified codewords, but that it also has the desired $\delta$-nonlinearity property provided we use finite fields with even characteristic and restrict the permutation polynomial $\sigma$ that we use.

**Theorem 9:** Consider the code $M = MGRS_k(F_2^N, X, \Psi)$ with codelength $N$. $M$ is $\delta$-nonlinear if it satisfies (i)-(iv):

(i) $\sigma \in F_2^N[x]$, $\sigma(x) = \sum_{i=1}^r a_i x^i$, $a_i \neq 0$, $1 \leq i \leq \delta$.

(ii) $\Psi : f(x) \rightarrow \sigma(f(x) \cdot x)$.

(iii) $2^n > r(s + 1) - 1$.

(iv) $N > \left( \frac{(\delta + 1)}{2} \right)^s$.

The $\delta$-nonlinearity property ensures that, provided enough block designs are concatenated, the partial trace of the concatenated code would only contain the coalition $T$ plus other non-codewords. The non-codewords can then be discarded leaving only the coalition.

**VI. CONCLUSIONS**

We have investigated powerful attacks against fingerprinting schemes used for securing digital content and described 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 hybrid attacks. For the outer code in the concatenation, we have defined modified GRS codes which have a certain nonlinearity property.

**ACKNOWLEDGMENT**

The first author was supported by an Australian Postgraduate Award from the ARC.

**REFERENCES**


