Códigos de fingerprinting binarios. Nuevos paradigmas de identificación

Marcel Fernandez
Departament d'Enginyeria Telemàtica
Universitat Politècnica de Catalunya

Elena Egorova
Higher School of Economics
Moscow, Russia

Grigory Kabatiansky
Institute for
Information Transmission Problems RAS
Moscow, Russia

***Resumen*—Los códigos con la Identifiable Parent Property garantizan, con probabilidad 1, la identificación de como mínimo de uno de los traidores en sistemas de Traitor Tracing. Desafortunadamente, para el caso de códigos binarios la propiedad IPP no se garantiza ni tan solo en el caso de dos traidores. En recientes trabajos se ha mostrado que códigos almost t-IPP de tasa positiva existen para el caso $t = 2$. De manera sorprendente, los esquemas de fingerprinting no poseen la propiedad IPP de manera automática. En la práctica, ello significa que para una huella digital z falsificada, un usuario identificado como culpable por el algoritmo de identificación, puede negar dicha acusación ya que podrá presentar una coalición de usuarios, en la que él no pertenece y que pueden crear el mismo z. En este artículo estudiamos los códigos t-almost IPP para $t > 2$.**

I. Introducción

El término traitor tracing fue acuñado en [1], y se refiere a códigos para prevenir redistribución ilegal de copias. Más adelante una parte de dicho trabajo fue refinada en [2] bajo el nombre de códigos con la *Identifiable Parent Property*.

Necesitaremos de las siguientes definiciones. Un *distribuidor* que comercia con contenido digital para M usuarios. Para prevenir la redistribución ilegal, el distribuidor D, hace que cada copia de su producto sea única. D consigue esto, insertando en cada copia que distribuye un vector de símbolos único, llamado *fingerprint*. Más concretamente, en la copia para el usuario i, D inserta un vector $\mathbf{c}^{(i)} = (c_1^{(i)}, \ldots, c_n^{(i)})$, de longitud n sobre un alfabeto q-ario. El alfabeto es generalmente el cuerpo finito F_q. Desafortunadamente, estos esquemas son débiles frente a ataques de confabulación. Llamamos una *coalición* de traidores a un grupo $U = \{\mathbf{u}_1, \ldots, \mathbf{u}_t\}$ de t usuarios deshonestos (*traidores*) que intentan redistribuir ilegalmente. Debido a que tener copias únicas disuade la redistribución ilegal, la coalición necesita crear una nueva copia ilegal que contiene un fingerprint falso $\mathbf{z} \in F_q^n$ que oculta sus identidades. Además, a partir de dicho fingerprint falso, D debe poder identificar a uno de los traidores en U.

Debido a que los traidores saben que sus copias estan marcadas, comparándolas pueden detectar diferencias. La estrategia de las coaliciones para crear un fingerprint falso, está restringido por la *Marking Assumption*, es decir, una coalición no puede cambiar las posiciones en las que todos sus miembros tienen el mismo valor, ya que dichas posiciones no son detectadas por comparación. Bajo la *Strong Marking Assumption* (SMA) una coalición puede crear solamente fingerprints en los que cada una de sus n coordenadas tiene solamente valores que los miembros de la coalición tienen en esta coordenada. Esta restricción apareció por primera vez en [1], donde los elementos del alfabeto son claves secretas, y por lo tanto, para esta coordenada los usuarios sólo tienen conocimiento de las claves que se les han dado.

En [3] se introdujo la *Weak Marking Assumption* (MA) y se estudió en un escenario de aplicación diferente, que permite a una coalición colocar cualquier símbolo del alfabeto en coordenadas donde no todos los miembros de la coalición tienen el mismo valor (es decir, el alfabeto se conoce). En el lenguaje de [4] corresponde a modelos narrow and wide envelope, y estos dos modelos coinciden en el caso binario ($q = 2$).

Una vez que se encuentra una copia ilegal, D extrae el fingerprint falso $\mathbf{z}$ y aplica a un algoritmo de identificación. Si los fingerprints forman un código con la *Identifiable Parent Property* (IPP), entonces a partir de la falsa huella digital $\mathbf{z}$, al menos un usuario de U será identificado con probabilidad 1.

Desafortunadamente, para el caso más interesante de códigos binarios la propiedad IPP no se mantiene incluso en el caso de sólo dos traidores. El reciente trabajo [5] ha considerado una generalización natural de códigos IPP, donde la propiedad debe cumplirse con una probabilidad de casi 1. Se ha demostrado en [5] que almost t-IPP códigos de cualquier tasa $R < 0,2075...$ existen para el caso $t = 2$. Al mismo tiempo, nadie antes que en [5] observó que buenos códigos de fingerprinting (CSDF) no proporcionan automáticamente la propiedad IPP. Esto significa que el distribuidor, a pesar de estar seguro de que el acusado es culpable, puede no ser capaz de demostrarlo. De hecho, el usuario acusado, al ser llevado a un tribunal de justicia pedirá revelar las marcas dadas a otros usuarios y entonces mostrará una coalición, a la que él no pertenece y sin embargo, esta coalición también es capaz de crear la misma huella digital. En otras palabras, esto dice que códigos CSDF sin la propiedad almost t-IPP propiedad muestran evidencia pero no prueban que un usuario es culpable.

I-A. Nuestras contribuciones

En [5] se ha comprobado que existen códigos almost 2-IPP para tasas de menos de $0,2075$ ... En este trabajo se estudia el caso de códigos almost t-IPP para $t > 2$. Se demuestra que para el caso binario, no existen tales códigos. Mirando a este

resultado parece una idea relajar el requisito de fingerprinting y no insistir en señalar a usuarios de culpables, sino en su lugar de llegar a una *red de usuarios sospechosos*, es decir, una lista que contiene por lo menos un traidor real con una probabilidad cercana a 1 y el distribuidor puede demostrarlo. Establecemos cotas inferiores del tamaño de dicha lista (red). Por otra parte, para códigos de fingerprintring conocidos se muestra que el tamaño de la red está acotado superiormente por el tamaño de la coalición de traidores.

II. Notación y resultados previos

Sea F_q un cuerpo finito de tamaño q que llmaremos *alfabeto*. Sea F_q^n un espacio vectorial n-dimensional sobre F_q, entonces $C \subseteq F_q^n$ es un *código* y los elementos de C se llaman *palabras código*. Un código C se llama *código lineal* (n, k) si es un subespacio vectorial de F_q^n k dimensional.

Sean $\mathbf{u}, \mathbf{v} \in F_q^n$ dos vectores (palabras), entonces la *distancia de Hamming* $d(\mathbf{u}, \mathbf{v})$ entre $\mathbf{u}$ y $\mathbf{v}$ se define como el número de posiciones en las que $\mathbf{u}$ y $\mathbf{v}$ difieren. La *distancia mínima* d de C se define como la distancia más pequeña entre dos palabras código diferentes.

Con el fin de evitar la redistribución ilegal, el distribuidor incrusta en cada una de las copias de su producto digital una huella digital diferente (palabra código) a partir de un cierto código

$$\varphi : \{1, \dots, M\} \to C \subset F_q^n.$$

Para cualquier $V \subset F_q^n$ y cualquier coordenada i se define la i-ésima proyección $P_i(V)$ del conjunto V como

$$P_i(V) = \bigcup_{\mathbf{v} \in V} v_i.$$

Llamaremos $< \varphi(U) >$ al conjunto de fingerprints falsos (también llamados descendientes) que la coalición U puede crear dentro de la suposición SMA, ver [1], [2]. Entonces

$$< \varphi(U) >= \{\mathbf{x} \in F_q^n : \forall i \; x_i \in P_i(\varphi(U))\} \qquad (1)$$

Observar que para $U = \{a, b\}$ el correspondiente conjunto $< \mathbf{a}, \mathbf{b} >$, donde $\mathbf{a} = \varphi(a)$ y $\mathbf{b} = \varphi(b)$, es el "segmento" $[\mathbf{a}, \mathbf{b}]$ en el espacio de Hamming F_q^n, ya que

$$< \mathbf{a}, \mathbf{b} >= \{\mathbf{x} \in F_q^n : d(\mathbf{a}, \mathbf{x}) + d(\mathbf{x}, \mathbf{b}) = d(\mathbf{a}, \mathbf{b})\}.$$

Por lo tanto para cualquier $V \subset F_q^n$ de tamaño arbitrario, llamaremos $< V >$ el convex hull de V (en [4] se le llama "narrow envelope").

Sea $E_t(C)$ el conjunto de todos los fingerprints falsos que pueden ser creados por todas las coaliciones $U \subset C$ de tamaño como màximo t.

Definición 1 ([2]): Diremos que un código C posee la identifiable parent property de orden t, o que C es un t-IPP código, si para todo $\mathbf{z} \in E_t(C)$

$$C_t(\mathbf{z}) := \bigcap_{U:\, \mathbf{z} \in <\varphi(U)>,\ |U| \le t} U \neq \emptyset \qquad (2)$$

También

Definición 2: Sea C un código t-IPP y sea $\mathbf{z}$ un descendiente de alguna coalición de tamaño máximo t. Llamaremos a las palabras en la intersección de (2) el conjunto de *padres positivos* de $\mathbf{z}$.

En este artículo los trminos padre y traidor se usan de manera indistinta.

II-A. Familias de códigos de fingerprinting

Sea $\mathbf{z} \in F_q^n$ un falso fingerprint extraído de una copia ilegal. Un código t-IPP permite al distribuidor siempre identificar positivamente al menos un miembro de una coalición de tamaño como máximo t que ha creado $\mathbf{z}$. De hecho, al observar $\mathbf{z}$ el distribuidor puede afirmar con seguridad que todos los usuarios que pertenecen al conjunto $C_t(\mathbf{z})$ definido en (2) son traidores positivos. Esto es porque cada coalición que puede crear $\mathbf{z}$ contiene $C_t(\mathbf{z})$.

Es fácil ver que no existen códigos t-IPP q-arios para $t \ge q$. Por otro lado, para $t < q$ exiten familias de códigos t-IPP con tasa postiva, i.e. con un número de palabras código exponencial en n, ver [2], [6]. Desfortunadamente, para casos de interés prácticos, es decir el caso *binario*, no hay códigos t-IPP incluido $t = 2$. Para superar este inconveniente la idea de una familia de códigos con una distribución de probabilidad conocida, pero una elección particular de código desconocida fue elaborada bajo el nombre de *collusion-secure digital fingerprinting codes* (CSDF codes) ver [3]. Es este caso, el distribuidor usa una familia conocida de codificaciones $\{\varphi_k : k \in K\}$

$$\varphi_k : \{1, \dots, M\} \to F_q^n$$

y asigna a cada i-ésimo usuario el fingerprint $\varphi_k(i)$, donde la codificación φ_k se elige mediante la distribución de probabilidad (conocida) $P(k)$.

Mediante su conjunto de fingerprints asignados $V = \{\mathbf{v}_1 = \varphi_k(u_1), \dots, \mathbf{v}_t = \varphi_k(u_t)\}$, los miembros de una coalición $U = \{u_1, \dots, u_t\} \subset \{1, \dots, M\}$ crean un fingerprint $\mathbf{x}$ de acuerdo con la distribución de probabilidad $g_{U,V}(\mathbf{x})$. La elección de la función $g_{U,V}(\mathbf{x})$ es la estrategia más general posible de la coalición.

También existe por parte del distribuidor un algoritmo de *identificación*

$$\Psi : F_q^n \times K \to \{1, \dots, M\},$$

para el que dado un fingerprint $\mathbf{x}$ y un $k \in K$, devuelve $j = \Psi(\mathbf{x}, k)$, y entonces el distribuidor identifica al usuario j-ésimo como traidor. El algoritmo de identificación, así como todos los demás elementos de la estrategia del distribuidor, son conocidos por los miembros de una coalición. Lo único que se mantiene en secreto es la elección particular de la "clave" k.

El objetivo del distribuidor debe ser, por un lado, maximizar la probabilidad de identificar un usuario culpable y por el otro minimizar la probabilidad de error

$$P_{err} = Pr\{\Psi(\mathbf{x}, k) \notin U\}$$

del algoritmo de identificación, es decir, reducir al mínimo la probabilidad de acusar a un usuario inocente. Ver [4] para más detalles.

Los códigos collusion-secure fingerprinting (BCSF) binarios, fueron los primeros en ser estudiados en [3] donde se construyen familias de códigos "buenos", es decir, con una tasa positiva

$$R = n^{-1} \log_q M \geq a > 0 \text{ y } P_{err} \to 0 \text{ cuando } n \to \infty.$$

Diferentes familias de códigos BCSF con mejores cotas se construyeron más tarde, véase [8].

II-B. Almost t-IPP codes

Como se ha mencionado antes un buen codigo (BCSF) no proporciona automáticamente la propiedad IPP a pesar de que la probabilidad de acusar a un usuario inocentes es muy pequeña. Ello significa que, en un juicio el usuario acusado pedirá revelar la correspondiente codificación φ_k y entonces él mostrará otra coalición, en la que el no es miembro que también es capaz de crear el mismo fingerprint falso.

A tal evento lo llamamos "coalition wins a court of law" (CWCL) y su probabilidad en buenos códigos de fingerprintint debería estar cerca de 1!

Este nuevo paradigma para códigos fingerprinting fue introducido en [5] bajo el nombre de *almost IPP codes*.

Para un conjunto de fingerprints dados $V = \{\mathbf{v}_1, \ldots, \mathbf{v}_t\}$, denotar por $K_{U,V} = \{k \in K : \varphi_k(u_1) = \mathbf{v}_1, \ldots, \varphi_k(u_t) = \mathbf{v}_t\}$ el correspondiente conjunto de claves admisibles k y denotar mediante

$$K_{U,V}^{(\mathbf{z})} = \{k \in K_{U,V} : \bigcap_{U:\, \mathbf{z} \in <\varphi_k(U)>,\ |U| \leq t} U = \emptyset\}$$

el subconjunto de claves admisibles, que permite a la coalición U "win a court of law" mediante el conjunto de fingerprints V generando un false fingerprint $\mathbf{z}$. La probabilidad de este evento (CWCL) es igual a

$$P_{win}(\mathbf{z}; U, V) = \Pr(K_{U,V})^{-1} \Pr(K_{U,V}^{(\mathbf{z})}) \tag{3}$$

Claramente, para cualquier algoritmo de indentificación Ψ la probabilidad que exista una coalición $\hat{U}$ tal que $\Psi(\mathbf{z}, k) \notin \hat{U}$ pero $\mathbf{z} \in <\varphi_k(\hat{U})>$ es al menos $P_{win}(\mathbf{z}; U, V)$.

Definición 3: Decimos que una familia de codificaciones

$$\{\varphi_k : \{1, \ldots, M\} \to F_q^n\}$$

junto con una distribución de probabilidad dada $P(k)$ tiene la ε-identifiable parent property de orden t, o es un código (ε, t)-IPP, si para todo $\mathbf{z} \in E_{t,K,\varphi}$

$$Pr\{K_{U,V}^{(\mathbf{z})}\} \leq \varepsilon, \tag{4}$$

donde $E_{t,K,\varphi} = \bigcup_{U,k:\ k \in K,\ |U| \leq t} \varphi_k(U)$ es el conjunto de todos los posibles fingerprints que pueden ser creados por coaliciones de tamaño como mucho t.

Definición 4: Una familia de códigos (ε, t)-IPP se denomina una familia de *almost t-IPP codes* si los correspondientes valores de ε tienden a cero cuando la longitud del código tiende a infinito.

A partir de [7] en [5] se demuestra el siguiente lemma.

Lema 1: Existen familias de almost 2-IPP codes de tasa

$$R \leq R_2 = 1 - \frac{\log_2 3}{2} = 0{,}2075... \tag{5}$$

En este artículo extendemos los resultados de [5] considerando el caso $t > 2$.

III. Almost t-IPP codes. El caso general - red de usuarios sospechosos

El Lemma 1 muestra que existen familias de almost 2-IPP codes. Es esta sección se muestra que no es el caso para $t > 2$.

Para proporcionar una cierta intuición y como un ejemplo introductorio se tiene el siguiente código

$$\begin{array}{ccccccc|c} 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{u}_1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 & \mathbf{u}_2 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 & \mathbf{u}_3 \\ 0 & 1 & 1 & 0 & 1 & 1 & 0 & \mathbf{u}_4 \\ 1 & 0 & 0 & 0 & 1 & 1 & 1 & \mathbf{u}_5 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 & \mathbf{u}_6 \\ 1 & 1 & 0 & 1 & 1 & 0 & 0 & \mathbf{u}_7 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & \mathbf{u}_8 \end{array}$$

Suponer que los usuarios con palabras código $\mathbf{u}_3 = (0101011)$, $\mathbf{u}_6 = (1011010)$ y $\mathbf{u}_8 = (1110001)$ confabulan y mediante una estrategia de decisión por *mayoría* crean el descendiente $\mathbf{z}$

$$\begin{array}{ccccccc|c} 0 & 1 & 0 & 1 & 0 & 1 & 1 & \mathbf{u}_3 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 & \mathbf{u}_6 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & \mathbf{u}_8 \\ \hline 1 & 1 & 1 & 1 & 0 & 1 & 1 & \mathbf{z} \end{array}$$

En este caso el distribuidor no podrá culpar a ningún usuario ya que el descendiente $\mathbf{z} = (1111011)$ puede ser creado por 3 coaliciones disjuntas $\{\mathbf{u}_3, \mathbf{u}_6\}$, $\{\mathbf{u}_3, \mathbf{u}_8\}$ y $\{\mathbf{u}_6, \mathbf{u}_8\}$. Así que si un usuario dado es señalado como culpable, siempre puede negar la acusación demostrando que el descendiente $\mathbf{z}$ puede ser creado por los otros dos usuarios. La generalización de este razonamiento para t arbitraria es sencilla. Considerar la siguiente estrategia de la coalición - cualquier coalición U de tamaño t crea su descendiente $\mathbf{z} = \mathbf{z}(U)$ tomando el valor mayoritario en todas las coordenadas (si t es un número par y en una posición dada el número de palabras de código de la coalición que tiene un símbolo 0 es igual al número de palabras de código que tiene un símbolo 1, se elije 0 ó 1 indistintamente como el símbolo de $\mathbf{z}$ para esta posición). Entonces, evidentemente, cualesquiera $\left\lfloor \frac{t}{2} \right\rfloor + 1$ palabras de código de U pueden generar $\mathbf{z}(U)$. Observamos que el distribuidor no puede probar que un único usuario, al que se ha acusado como traidor, es realmente culpable. Por lo tanto introducimos una versión suavizada de la *identificación*, es decir, una *red de usuarios sospechosos*.

Definición 5: Para un código C dado y un fingerprint falso $\mathbf{z}$ un conjunto de L usuarios es la *red de* $\mathbf{z}$ que denotamos por $N(\mathbf{z}, t, C)$ si $L \bigcap U \neq \emptyset$ para todo U tal que $\mathbf{z} \in <\varphi(U)>$

y $|U| \leq t$.

Es claro que el distribuidor desea tener un tamaño mínimo posible de red. Se observa que $|N(\mathbf{z}, t, C)| = 1$ para todo $|\mathbf{z} \in E_t(C)$ sii C es un código t-IPP. De la misma forma, a la idea de almost IPP-codes podemos introducir el concepto de ε-*red* exigiendo que $L \bigcap U \neq \emptyset$ para todo U tal que $\mathbf{z} \in < \varphi_k(U) >$ y $|U| \leq t$ con probabilidad como mínimo $1 - \varepsilon$. Estos argumentos muestran que el tamaño mínimo de red no puede ser inferior a $t - \left\lfloor \frac{t}{2} \right\rfloor$.

Lema 2: Para cualquier $0 < \varepsilon < 1$ y cualquier familia de códigos fingerprinting C_k la ε-red de usuarios sopechosos tiene tamaño al menos $\lceil t/2 \rceil$.

IV. Cotas superiores del tamaño de la red. Constructiones concatenadas de códigos CSDF con códigos separadores como códigos internos

En el Lema 2, hemos obtenido una cota inferior en el tamaño de la red. En esta sección mostramos una cota superior en el tamaño de la red para la familia de códigos de [4].

La familia de códigos de fingerprinting en [4] se basa en concatenación de códigos. El código interno es un código binario (t, t)-separador, y el código externo es un código q-ario con distancia mínima grande (códigos Reed-Solomon, o códigos AG).

Un código V se le llama (t, t)-separador si para dos conjuntos disjuntos cualesquiera de palabras código, por ejemplo X y Y, ambos de tamaño como màximo t, existe una coordenada i que los "separa", i.e. $P_i(X) \cap P_i(Y) = \emptyset$.

Denotaremos el código interno (t, t)-separador por V. Sea n la longitud del código V. Denotaremos el código externo q-ario por W. Sea N su longitud, M su cardinalidad, D su distancia mínima y sea $q = |V|$. Consideraremos las $q!$ biyecciones $\varphi : F_q \to V$ y las codificaciones $(q!)^N$ $\Phi = (\varphi^{(1)}, \ldots, \varphi^{(N)})$ que asignan a una palabra código del código externo $\mathbf{w} = (w_1, \ldots, w_N) \in W$ una palabra binaria del código concatenado resultante $\Phi(\mathbf{w}) = (\varphi^{(1)}(w_1), \ldots, \varphi^{(N)}(w_N))$ de longitud nN. Estas $(q!)^N$ codificaciones Φ junto con una distribución de probabilidad uniforme forman el código CSDF de [4].

El algoritmo de identificación de [4] como es usual para códigos concatenados consiste de dos etapas. El Tracing Algorithm 1 de [4] garantiza que la probabilidad de inculpar a un usuario inocente tiende a cero exponencialmente, ver Theorem 4.1 [4]. La coalición U genera un falso fingerprint binario $\mathbf{z} = (z_1, \ldots, z_{nN})$, que dividimos en bloques de longitud n, que denotamos por $X_1, \ldots, X_N$. Cada uno de estos bloques es generado por la correspondiente "coalición" de vectores del código V.

El primer paso es decodificar (tracing-decoding) el código interno V. Puede haber varias coaliciones de tamaño máximo t que pueden generar X_i, pero la intersección dos a dos es no vacía debido a la propiedad (t, t)-separadora. Por lo tanto, elegimos cualquiera de dichas coaliciones, H_i. Sabemos que $H_i \cap P_i(U) \neq \emptyset$.

Sea la función $s(\mathbf{w}, H)$ definida como

$$s(\mathbf{x}, H) = |\{i : x_i \in H_i\}|, \tag{6}$$

En el segundo paso el algoritmo de identificación devuelve la palabra código del código externo $\hat{\mathbf{w}} = (\hat{w}_1, \ldots, \hat{w}_N)$ que

$$s(\mathbf{w}, H) \geq s(\mathbf{v}, H), \ \forall \ \mathbf{v} \in W.$$

Es decir, el algoritmo de identificación devuelve la palabra código $\hat{\mathbf{w}}$ más cercana al conjunto $H = H_1 \times H_2 \times \ldots H_N$.

Es fácil ver que $s(\hat{\mathbf{w}}, H) \geq N/t$ y se demostró en [4] que si la distancia mínima del código externo W es lo suficientemente grande, entonces el algoritmo de identificación acusa a un verdadero traidor $\mathbf{w}$ (con una probabilidad que tiende a 1), pero, como hemos explicado antes, el distribuidor no puede probarlo. En la línea de [4] definamos la red $N_{H,t}$ de usuarios sospechosos como el conjunto de palabras código $\mathbf{w}$ tales que $s(\mathbf{w}, H) \geq N/t$. Entonces, cualquier coalición que pueda crear la misma fingerprint falsa tiene una intersección no vacía con $N_{H,t}$ y el distribuidor puede acusar al conjunto $N_{H,t}$ como la red de usuarios sospechosos porque al menos uno de ellos es un traidor. Por otra parte, "casi todos" los usuarios en $N_{H,t}$ son culpables, ya que se demostró en [4] que la probabilidad de que exista $\mathbf{u}$ con $s(\mathbf{u}, H) \geq N/t$ pero no que este en la coalición de traidores tiende a cero. Este razonamiento esboza la demostración del siguiente teorema.

Teorema 1: Para la familia de códigos t-CSFC de [4] el tamaño de la red de usuarios sospechosos es como máximo t con probabilidad tendiendo a 1 y casi todos los usuarios de la red son culpables.

V. Conclusiones

En este trabajo hemos discutido un nuevo paradigma para los códigos binarios de fingerprinting - almost t-IPP codes. Se ha demostrado muy recientemente que existen códigos almost t-IPP para coaliciones de tamaño $t = 2$. Mostramos que, lamentablemente, no es el caso para $t > 2$.

Al no ser capaces de demostrar que un usuario es un traidor, se presenta una versión relajada de la identificación. Ahora el algoritmo de identificación puede generar una lista, que nosotros llamamos red de usuarios sospechosos, de manera que el distribuidor puede demostrar que al menos un usuario de la red es culpable. Se han demostrado cotas inferiores y superiores para el tamaño mínimo de red, las últimas para una familia particular de códigos de fingerprinting. El trabajo futuro incluye el establecimiento de cotas en el tamaño de la lista correspondiente a otras familias de códigos de la literatura. Nótese que para cualquier código (t, t)-separador cualquier coalición, que puede crear un descendiente $\mathbf{z}$, puede ser elegida como la red de $\mathbf{z}$, ya por definición cualesquiera dos coaliciones que puede crear $\mathbf{z}$, tienen una intersección no vacía. La ventaja del Teorema 1, es que la red correspondiente es la red de "usuarios fuertemente sospechosos", es decir, que el distribuidor puede demostrar en un tribunal de justicia que al menos un usuario de la red es culpable y por otra parte la comprobación correspondiente mostrará que casi todos los usuarios en la red son culpables.

AGRADECIMIENTOS

. El trabajo de Marcel Fernandez está financiado en por el Gobierno de España a través de TEC2011-26491 “COPPI” y ANALISIS FORENSE AVANZADO (ANFORA) - TEC2015-68734-R El trabajo de G.Kabatiansky está financiado en parte por RFBR grants $13-07-00978$, $13-01-12458$ and $14-01-93108$.

REFERENCIAS

[1] B. Chor, A. Fiat, and M.Naor. Tracing traitors. *Advances in Cryptology-Crypto'94, LNCS*, 839:480–491, 1994.

[2] Henk D. L. Hollmann, Jack H. van Lint, Jean-Paul Linnartz, and Ludo M. G. M. Tolhuizen. On codes with the Identifiable Parent Property. *J. Combinatorial Theory*, 82(2):121–133, May 1998.

[3] D. Boneh and J. Shaw. Collusion-secure fingerprinting for digital data. *IEEE Trans. Inform. Theory*, 44(5):1897–1905, 1998.

[4] A. Barg, G. R. Blakley, and G. Kabatiansky. Digital fingerprinting codes: Problem statements, constructions, identification of traitors. *IEEE Trans. Inform. Theory*, 49(4):852–865, 2003.

[5] M. Fernandez, G.Kabatiansky, and J. Moreira. Almost IPP-codes or provably secure digital fingerprinting codes. In *Proc. IEEE International Symp. Information Theory (ISIT 2015)*, pages 1595–1599. IEEE Computer Society, 2015.

[6] A. Barg, G. Cohen, S. Encheva, G. Kabatiansky, and G. Zémor. A hypergraph approach to the identifying parent property: the case of multiple parents. *SIAM J. Disc. Math*, 14:423–431, 2001.

[7] G.R.Blakley and G.Kabatiansky. Random coding technique for digital fingerprinting codes: fighting two pirates revisited. In *Proc. IEEE International Symp. Information Theory (ISIT 2004)*, page 203. IEEE Computer Society, 2004.

[8] Gábor Tardos. Optimal probabilistic fingerprint codes. In *Proceedings of the 35th Annual ACM Symposium on Theory of Computing, June 9-11, 2003, San Diego, CA, USA*, pages 116–125. ACM, 2003.