Two-level Fingerprinting Codes

Alexander Barg

University of Maryland, College Park

Joint work with

Prasanth Anthapadmanabhan

UMD

ISIT 2009, Seoul, Korea
Fingerprinting

Introduce *hidden marks* in every legal copy

Original: 0110000101000110001110101000101011110010

⇓

User 1: 01110000101000110000011110100100010111101101111011010
User 2: 01100000101100011000011100111010010010001011110110111101010

...  

User M: 01110000101100011000001110011100101000100010111100010

\[ x(1): \begin{array}{c} 100111 \\ x(2): \begin{array}{c} 010011 \\ \vdots \\ x(M): \begin{array}{c} 110000 \end{array} \end{array} \end{array} \]

“Fingerprints”

Collection is called a *code*, denoted by \( C : \{1, \ldots, M\} \rightarrow Q^n \)
The fingerprinting problem

Decoding algo outputs \textit{one} of the guilty users.
The fingerprinting problem

Decoding algorithm outputs one of the guilty users.
The Boneh-Shaw fingerprinting problem

Which forgeries are possible?

User 2: 0111000010100001100011110101000101110111101010
User 3: 01100000101100011000011100101000101110111101010

\[ x_1 : \quad 100111 \]
\[ x_2 : \quad 010011 \]
\[ y : \quad 000111 \]

Envelope

Set of all possible forgeries a coalition can create.

\[ \mathcal{E}(x_1, \ldots, x_t) = \{ y | y_i \in \{ x_{1i}, \ldots, x_{ti} \}, \forall i \} \]
Randomized codes

**Randomized code**

- **Family of codes:** \( \{ C_k, k \in K \} \), \( C_k : \{1, \ldots, M\} \rightarrow \mathbb{Q}^n \)
- **Decoding algos:** \( \{ D_k, k \in K \} \), \( D_k : \mathbb{Q}^n \rightarrow \{1, \ldots, M\} \cup \{0\} \)

\[(C, D) \equiv \text{rv over } \{(C_k, D_k), k \in K\}\]

**Rate** \( R = \frac{1}{n} \log_q M \)

Code family is public. *Selection of \( k \) known only to distributor.*
Coalition strategies

- Coalition $U \subseteq \{1, \ldots, M\}$ of size $t$ observes $C_k(U) = \{x_1, \ldots, x_t\}$
- $V(\cdot, \ldots, \cdot)$: coalition strategy
  
  $V(x_1, \ldots, x_t) =$ outputs a random forgery in $\mathcal{E}(x_1, \ldots, x_t)$

  $\mathbf{Y}_{c,u,v} := V(C(U))$

- Probability of error for a given coalition $U$ and strategy $V$:
  
  $e(C, \mathcal{D}, U, V) = \mathbb{P}[\mathcal{D}(\mathbf{Y}_{c,u,v}) \notin U]$
(One-level) Fingerprinting codes

**Definition**

$(C, D)$ is $t$-fingerprinting with $\varepsilon$-error if

$$\max_{V} e(C, D, U, V) \leq \varepsilon, \quad \forall U \text{ with } |U| \leq t.$$ 

**Informally:** A randomized code is $t$-fingerprinting if for every coalition of size $\leq t$ and any coalition strategy, it is possible to identify one of the pirates with high probability.
(One-level) Fingerprinting codes

Definition

$(C, D)$ is \textbf{$t$-fingerprinting} with $\varepsilon$-error if

\[
\max_V e(C, D, U, V) \leq \varepsilon, \quad \forall U \text{ with } |U| \leq t.
\]

Informally: A randomized code is \textbf{$t$-fingerprinting} if for every coalition of size $\leq t$ and any coalition strategy, it is possible to identify one of the pirates with high probability.

- N. Merhav and A. Somekh-Baruch, \textit{Capacity of private fingerprinting}, IT, March 2005
- E. Amiri and G. Tardos, \textit{High-rate codes and FP capacity}, ACM SODA 2009
- P. Moulin, \textit{Capacity and error exponents for FP}, 2008-09
- Y.-W. Huang and P. Moulin, previous talk.
Two-level codes

Users are organized in $M_1$ groups each containing $M_2$ users.

Accordingly,

\[ C_k : [M_1] \times [M_2] \rightarrow Q^n, \]
\[ D_k : Q^n \rightarrow ([M_1] \times [M_2]) \cup \{0\}, \]
\[ (R_1, R_2) = \left( \frac{1}{n} \log_q M_1, \frac{1}{n} \log_q M_2 \right). \]
Error probabilities

For user \( \mathbf{u} \equiv (u_1, u_2) \), let \( G(\mathbf{u}) = u_1 \), be its group index.

Error probabilities for a given coalition \( U \) and strategy \( V \):

\[
e(C, D, U, V) = \mathbb{P} \left[ D(Y_{C,U,V}) \notin U \right]
\]

\[
\bar{e}(C, D, U, V) = \mathbb{P} \left[ \underbrace{G(D(Y_{C,U,V}))}_{\text{group index of decoder output}} \notin \underbrace{G(U)}_{\text{groups containing guilty users}} \right]
\]
Two-level fingerprinting codes

Definition

$$(C, D)$$ is $$(t_1, t_2)$$-fingerprinting with $$(\varepsilon_1, \varepsilon_2)$$-error (where $$t_1 > t_2$$) if

1. $$\max_v e(C, D, U, V) \leq \varepsilon_2, \quad \forall U \text{ with } |U| \leq t_2. \quad \Leftrightarrow \quad t_2$$-fingerprinting

2. $$\max_v \tilde{e}(C, D, U, V) \leq \varepsilon_1, \quad \forall U \text{ with } |U| \leq t_1. \quad \Leftarrow \quad t_1$$-fingerprinting

In fact, with $$t_1 > t_2$$,

1. one-level $$t_1$$-fingerprinting $$\Rightarrow$$ $$(t_1, t_2)$$-fingerprinting

Q: Can we achieve higher rates?

2. $$(t_1, t_2)$$-fingerprinting of size $$M_1 M_2$$ $$\Rightarrow$$ one-level $$t_1$$-fingerprinting of size $$M_1$$. 
Want: \((C_n, D_n)\) which are \((2,1)\)-fingerprinting with \((\epsilon_{1,n}, \epsilon_{2,n})\)-error, such that \(\epsilon_{1,n} \rightarrow_n 0\) and \(\epsilon_{2,n} \rightarrow_n 0\)
(2,1)-fingerprinting code

Want: \((C_n, D_n)\) which are (2,1)-fingerprinting with \((\varepsilon_{1,n}, \varepsilon_{2,n})\)-error, such that \(\varepsilon_{1,n} \to_n 0\) and \(\varepsilon_{2,n} \to_n 0\)
(2,1)-fingerprinting code

Want: \((C_n, D_n)\) which are (2,1)-fingerprinting with \((\varepsilon_{1,n}, \varepsilon_{2,n})\)-error, such that \(\varepsilon_{1,n} \to_n 0\) and \(\varepsilon_{2,n} \to_n 0\)
(2,1)-fingerprinting code

Want: \((C_n, D_n)\) which are (2,1)-fingerprinting with \((\varepsilon_{1,n}, \varepsilon_{2,n})\)-error, such that \(\varepsilon_{1,n} \to_n 0\) and \(\varepsilon_{2,n} \to_n 0\)
(2,1)-fingerprinting code

**Decoding:** minimum distance

\[ D_{k,n}(y) = \arg \min_u d_H(C_{k,n}(u), y) \]

**Theorem**

\((C_n, D_n)\) is \((2,1)\)-fingerprinting if there exists an \(\omega \in [0, \frac{q-1}{2q}]\) such that

\[ R_1 < 1 - h \left( \frac{1}{2} \left( 1 - \frac{1}{q} \right) + \omega \right), \]

\[ R_2 < h(\omega). \]

**Remark:** For such \(\omega\)'s to exist we need large alphabet, e.g., \(q \geq t + 1\) (recall IPP codes)
(2,1)-fingerprinting: Achievable rates

Inner bound

2-fingerprinting

[ABD, 2008]
(2,1)-fingerprinting: Achievable rates

Inner bound
2-fingerprinting
[ABD, 2008]

Inner bound
2-FP with MD decoder
[Lin, Shahmohammadi, ElGamal, 2007]
(2,1)-fingerprinting: Achievable rates

- **Inner bound**
  - 2-fingerprinting
    - [ABD, 2008]
  - 2-FP with MD decoder
    - [Lin, Shahmohammadi, ElGamal, 2007]

- **This work**
(2,1)-fingerprinting: Achievable rates

Inner bound

2-fingerprinting
[ABD, 2008]

Inner bound
2-FP with MD decoder
[Lin, Shahmohammadi, ElGamal, 2007]

This work

Outer bound
From one-level codes
[Amiri and Tardos ’09]
[ Huang and Moulin ’09]
(t,1)-fingerprinting code

Extending the idea...

**Theorem**

(Cₙ, Dₙ) is (t, 1)-fingerprinting if there exists an ω with

\[
\frac{t-1}{t} \left(1 - \frac{1}{q^{t-1}}\right) + \omega \leq \frac{q-1}{q}
\]

such that

\[
R_1 < 1 - h \left(\frac{t-1}{t} \left(1 - \frac{1}{q^{t-1}}\right) + \omega\right),
\]

\[
R_2 < h(\omega).
\]
Similarly...

**Theorem**

$(C_n, D_n)$ is *(t, 2)-fingerprinting* if there exists an $\omega$ with

\[
\frac{t - 1}{t} \left(1 - \frac{1}{q^{t-1}}\right) + \omega \leq \frac{q - 1}{q}
\]

such that

\[
R_1 < 1 - h \left(\frac{t - 1}{t} \left(1 - \frac{1}{q^{t-1}}\right) + \omega\right),
\]

\[
R_2 < h(\omega) - (\ldots).
\]
Conclusion

- Introduced *two-level codes* for fingerprinting
  \[ \approx \text{unequal error protection} \]
- A bound on achievable rates under minimum-distance decoding for
  \((t, 1)\) fingerprinting, \((t, 2)\) fingerprinting
Conclusion

- Introduced *two-level codes* for fingerprinting
  \( \approx \) unequal error protection
- A bound on achievable rates under minimum-distance decoding for (\( t, 1 \)) fingerprinting, (\( t, 2 \)) fingerprinting

Open
Conclusion

- Introduced *two-level codes* for fingerprinting
  ≈ unequal error protection
- A bound on achievable rates under minimum-distance decoding for $(t, 1)$ fingerprinting, $(t, 2)$ fingerprinting

Open

- Two-level codes with polynomial-time decoding
Conclusion

- Introduced *two-level codes* for fingerprinting
  \[ \approx \text{unequal error protection} \]
- A bound on achievable rates under minimum-distance decoding for 
  \((t, 1)\) fingerprinting, \((t, 2)\) fingerprinting

Open

- Construction for \((t_1, t_2)\) FP
- Two-level codes with polynomial-time decoding
Conclusion

- Introduced *two-level codes* for fingerprinting
  ≈ unequal error protection
- A bound on achievable rates under minimum-distance decoding for $(t, 1)$ fingerprinting, $(t, 2)$ fingerprinting

Open

- Construction for $(t_1, t_2, \ldots, t_m)$ FP
- Construction for $(t_1, t_2)$ FP
- Two-level codes with polynomial-time decoding
Conclusion

- Introduced *two-level codes* for fingerprinting
  \approx unequal error protection

- A bound on achievable rates under minimum-distance decoding for
  \((t, 1)\) fingerprinting, \((t, 2)\) fingerprinting

Open

- **Upper bounds**
  - Construction for \((t_1, t_2, \ldots, t_m)\) FP
  - Construction for \((t_1, t_2)\) FP
  - Two-level codes with polynomial-time decoding