Capacities and Capacity-Achieving Decoders for Various Fingerprinting Games

Thijs Laarhoven

mail@thijs.com
http://www.thijs.com/

IH&MMSec 2014, Salzburg, Austria
(June 12, 2014)

Outline

Introduction

Related work
- Lower bounds
- Efficient decoders

Previously on IH&MMSec 2013

Contributions
- Lower bounds
- Efficient decoders

Conclusion

TU/e

Introduction

Problem: Illegal redistribution

User	Copyrighted content																
Antonino	0	1	1	1	0	0	1	1	1	0	1	1	0	0	1	0	...
Boris	0	1	1	1	0	0	1	1	1	0	1	1	0	0	1	0	...
Caroline	0	1	1	1	0	0	1	1	1	0	1	1	0	0	1	0	...
David	0	1	1	1	0	0	1	1	1	0	1	1	0	0	1	0	...
Eve	0	1	1	1	0	0	1	1	1	0	1	1	0	0	1	0	...
Fred	0	1	1	1	0	0	1	1	1	0	1	1	0	0	1	0	...
Gábor	0	1	1	1	0	0	1	1	1	0	1	1	0	0	1	0	...
Henry	0	1	1	1	0	0	1	1	1	0	1	1	0	0	1	0	...

TU/e

Introduction

Problem: Illegal redistribution

User	Copyrighted content																
Antonino	0	1	1	1	0	0	1	1	1	0	1	1	0	0	1	0	...
Boris	0	1	1	1	0	0	1	1	1	0	1	1	0	0	1	0	...
Caroline	0	1	1	1	0	0	1	1	1	0	1	1	0	0	1	0	...
David	0	1	1	1	0	0	1	1	1	0	1	1	0	0	1	0	...
Eve	0	1	1	1	0	0	1	1	1	0	1	1	0	0	1	0	...
Fred	0	1	1	1	0	0	1	1	1	0	1	1	0	0	1	0	...
Gábor	0	1	1	1	0	0	1	1	1	0	1	1	0	0	1	0	...
Henry	0	1	1	1	0	0	1	1	1	0	1	1	0	0	1	0	...
Copy	0	1	1	1	0	0	1	1	1	0	1	1	0	0	1	0	...

Introduction

Solution: Embed fingerprints

User	Copyrighted content (fingerprinted)															
Antonino	0	1	1	1	0	0	1	1	1	0	1	1	0	1	0	0 ...
Boris	0	1	1	1	0	1	0	1	1	0	1	1	1	1	1	0 ...
Caroline	0	1	0	1	0	1	0	1	1	0	0	1	1	0	1	0 ...
David	0	1	1	1	0	0	0	1	1	0	1	1	0	0	0	0 ...
Eve	0	1	0	1	0	1	0	1	1	0	1	1	1	0	0	0 ...
Fred	0	1	0	1	0	0	1	1	1	0	0	1	0	1	0	0 ...
Gábor	0	1	1	1	0	1	1	1	1	0	1	1	0	0	1	0 ...
Henry	0	1	0	1	0	1	1	1	1	0	0	1	0	1	1	0 ...

Introduction

Solution: Embed fingerprints

User	Copyrighted content (fingerprinted)															
Antonino	0	1	1	1	0	0	1	1	1	0	1	1	0	1	0	0	...
Boris	0	1	1	1	0	1	0	1	1	0	1	1	1	1	1	0	...
Caroline	0	1	0	1	0	1	0	1	1	0	0	1	1	0	1	0	...
David	0	1	1	1	0	0	0	1	1	0	1	1	0	0	0	0	...
Eve	0	1	0	1	0	1	0	1	1	0	1	1	1	0	0	0	...
Fred	0	1	0	1	0	0	1	1	1	0	0	1	0	1	0	0	...
Gábor	0	1	1	1	0	1	1	1	1	0	1	1	0	0	1	0	...
Henry	0	1	0	1	0	1	1	1	1	0	0	1	0	1	1	0	...
Copy	0	1	0	1	0	1	0	1	1	0	1	1	1	0	0	0	...

TU/e

Introduction

Solution: Embed fingerprints

User	Copyrighted content (fingerprinted)															
Antonino	0	1	**1**	1	0	**0**	**1**	1	1	0	**1**	1	**0**	**1**	**0**	0	...
Boris	0	1	**1**	1	0	**1**	**0**	1	1	0	**1**	1	**1**	**1**	**1**	0	...
Caroline	0	1	**0**	1	0	**1**	**0**	1	1	0	**0**	1	**1**	**0**	**1**	0	...
David	0	1	**1**	1	0	**0**	**0**	1	1	0	**1**	1	**0**	**0**	**0**	0	...
Eve	0	1	**0**	1	0	**1**	**0**	1	1	0	**1**	1	**1**	**0**	**0**	0	...
Fred	0	1	**0**	1	0	**0**	**1**	1	1	0	**0**	1	**0**	**1**	**0**	0	...
Gábor	0	1	**1**	1	0	**1**	**1**	1	1	0	**1**	1	**0**	**0**	**1**	0	...
Henry	0	1	**0**	1	0	**1**	**1**	1	1	0	**0**	1	**0**	**1**	**1**	0	...
Copy	0	1	**0**	1	0	**1**	**0**	1	1	0	**1**	1	**1**	**0**	**0**	0	...

TU/e

Introduction

Solution: Embed fingerprints

User	Copyrighted content (fingerprinted)															
Antonino	0	1	**1**	1	0	**0**	**1**	1	1	0	**1**	1	**0**	**1**	**0**	0	...
Boris	0	1	**1**	1	0	**1**	**0**	1	1	0	**1**	1	**1**	**1**	**1**	0	...
Caroline	0	1	**0**	1	0	**1**	**0**	1	1	0	**0**	1	**1**	**0**	**1**	0	...
David	0	1	**1**	1	0	**0**	**0**	1	1	0	**1**	1	**0**	**0**	**0**	0	...
Eve	0	1	**0**	1	0	**1**	**0**	1	1	0	**1**	1	**1**	**0**	**0**	0	...
Fred	0	1	**0**	1	0	**0**	**1**	1	1	0	**0**	1	**0**	**1**	**0**	0	...
Gábor	0	1	**1**	1	0	**1**	**1**	1	1	0	**1**	1	**0**	**0**	**1**	0	...
Henry	0	1	**0**	1	0	**1**	**1**	1	1	0	**0**	1	**0**	**1**	**1**	0	...
Copy	0	1	**0**	1	0	**1**	**0**	1	1	0	**1**	1	**1**	**0**	**0**	0	...

Introduction

Problem: Collusion attacks

User	Copyrighted content (fingerprinted)															
Antonino	0	1	1	1	0	0	1	1	1	0	1	1	0	1	0	0 ...
Boris	0	1	1	1	0	1	0	1	1	0	1	1	1	1	1	0 ...
Caroline	0	1	0	1	0	1	0	1	1	0	0	1	1	0	1	0 ...
David	0	1	1	1	0	0	0	1	1	0	1	1	0	0	0	0 ...
Eve	0	1	0	1	0	1	0	1	1	0	1	1	1	0	0	0 ...
Fred	0	1	0	1	0	0	1	1	1	0	0	1	0	1	0	0 ...
Gábor	0	1	1	1	0	1	1	1	1	0	1	1	0	0	1	0 ...
Henry	0	1	0	1	0	1	1	1	1	0	0	1	0	1	1	0 ...

Introduction

Problem: Collusion attacks

User	Copyrighted content (fingerprinted)															
Antonino	0	1	1	1	0	0	1	1	1	0	1	1	0	1	0	0 ...
Boris	0	1	1	1	0	1	0	1	1	0	1	1	1	1	1	0 ...
Caroline	0	1	0	1	0	1	0	1	1	0	0	1	1	0	1	0 ...
David	0	1	1	1	0	0	0	1	1	0	1	1	0	0	0	0 ...
Eve	0	1	0	1	0	1	0	1	1	0	1	1	1	0	0	0 ...
Fred	0	1	0	1	0	0	1	1	1	0	0	1	0	1	0	0 ...
Gábor	0	1	1	1	0	1	1	1	1	0	1	1	0	0	1	0 ...
Henry	0	1	0	1	0	1	1	1	1	0	0	1	0	1	1	0 ...

Introduction

Problem: Collusion attacks

User	Copyrighted content (fingerprinted)															
Antonino	0	1	**1**	1	0	**0**	**1**	1	1	0	**1**	1	**0**	**1**	**0**	0	...
Boris	0	1	1	1	0	1	0	1	1	0	1	1	1	1	1	0	...
Caroline	0	1	0	1	0	1	0	1	1	0	0	1	1	0	1	0	...
David	0	1	1	1	0	0	0	1	1	0	1	1	0	0	0	0	...
Eve	0	1	**0**	1	0	**1**	**0**	1	1	0	**1**	1	**1**	**0**	**0**	0	...
Fred	0	1	0	1	0	0	1	1	1	0	0	1	0	1	0	0	...
Gábor	0	1	1	1	0	1	1	1	1	0	1	1	0	0	1	0	...
Henry	0	1	0	1	0	1	1	1	1	0	0	1	0	1	1	0	...
Copy	0	1	**1**	1	0	**1**	**0**	1	1	0	**1**	1	**0**	**1**	**0**	0	...

Introduction

Problem: Collusion attacks

User	Copyrighted content (fingerprinted)															
Antonino	0	1	**1**	1	0	**0**	**1**	1	1	0	**1**	1	**0**	**1**	**0**	0	...
Boris	0	1	**1**	1	0	**1**	**0**	1	1	0	**1**	1	**1**	**1**	**1**	0	...
Caroline	0	1	**0**	1	0	**1**	**0**	1	1	0	**0**	1	**1**	**0**	**1**	0	...
David	0	1	**1**	1	0	**0**	**0**	1	1	0	**1**	1	**0**	**0**	**0**	0	...
Eve	0	1	**0**	1	0	**1**	**0**	1	1	0	**1**	1	**1**	**0**	**0**	0	...
Fred	0	1	**0**	1	0	**0**	**1**	1	1	0	**0**	1	**0**	**1**	**0**	0	...
Gábor	0	1	**1**	1	0	**1**	**1**	1	1	0	**1**	1	**0**	**0**	**1**	0	...
Henry	0	1	**0**	1	0	**1**	**1**	1	1	0	**0**	1	**0**	**1**	**1**	0	...
Copy	0	1	**1**	1	0	**1**	**0**	1	1	0	**1**	1	**0**	**1**	**0**	0	...

Introduction

Solution: Collusion-resistant schemes

User	Copyrighted content (fingerprinted)																
Antonino	0	1	?	1	0	?	?	1	1	0	?	1	?	?	?	0	...
Boris	0	1	?	1	0	?	?	1	1	0	?	1	?	?	?	0	...
Caroline	0	1	?	1	0	?	?	1	1	0	?	1	?	?	?	0	...
David	0	1	?	1	0	?	?	1	1	0	?	1	?	?	?	0	...
Eve	0	1	?	1	0	?	?	1	1	0	?	1	?	?	?	0	...
Fred	0	1	?	1	0	?	?	1	1	0	?	1	?	?	?	0	...
Gábor	0	1	?	1	0	?	?	1	1	0	?	1	?	?	?	0	...
Henry	0	1	?	1	0	?	?	1	1	0	?	1	?	?	?	0	...
Copy	0	1	?	1	0	?	?	1	1	0	?	1	?	?	?	0	...

Introduction

Solution: Collusion-resistant schemes

User	Copyrighted content (fingerprinted)							
Antonino	?	?	?	?	?	?	?	...
Boris	?	?	?	?	?	?	?	...
Caroline	?	?	?	?	?	?	?	...
David	?	?	?	?	?	?	?	...
Eve	?	?	?	?	?	?	?	...
Fred	?	?	?	?	?	?	?	...
Gábor	?	?	?	?	?	?	?	...
Henry	?	?	?	?	?	?	?	...
Copy	?	?	?	?	?	?	?	...

Introduction
Some notation

- n: total number of users
- c: number of colluders/pirates ($c \ll n$)
- ℓ: code length, size of fingerprints
- X: code matrix, assigning fingerprints to users
- y: pirate output

Related work
Lower bounds

How many symbols ℓ are necessary for static fingerprinting?

- 1998: $\ell = \Omega(c \log n)$[1]
- 2003: $\ell = \Omega(c^2 \log \frac{n}{c})$[2]
- 2003: $\ell = \Omega(c^2 \log n)$[3]
- 2009: $\ell \overset{?}{\sim} 2c^2 \ln n$[4]
- 2012: $\ell \sim 2c^2 \ln n$[5]
    - asymptotic optimal attack is the interleaving attack

---

[1] D. Boneh and J. Shaw, "Collusion-secure fingerprinting for digital data," *IEEE Transactions on Information Theory*, vol. 44, no. 5, pp. 1897–1905, 1998.

[2] C. Peikert et al., "Lower bounds for collusion-secure fingerprinting," in *ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2003, pp. 472–479.

[3] G. Tardos, "Optimal probabilistic fingerprint codes," in *ACM Symposium on Theory of Computing (STOC)*, 2003, pp. 116–125.

[4] E. Amiri and G. Tardos, "High rate fingerprinting codes and the fingerprinting capacity," in *ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2009, pp. 336–345.

[5] Y.-W. Huang and P. Moulin, "On the saddle-point solution and the large-coalition asymptotics of fingerprinting games," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 1, pp. 160–175, 2012.

TU/e

Related work

Efficient decoders

How many symbols ℓ are sufficient for static fingerprinting?

- 1995: $\ell = O(c^4 \log n)$[1]
- 2003: $\ell = 100c^2 \ln n$[2] ("the Tardos scheme")
- 2006: $\ell \sim 4\pi^2 c^2 \ln n$[6]
- 2008: $\ell \sim \pi^2 c^2 \ln n$[7]
- 2008: $\ell \overset{?}{\sim} \frac{1}{2}\pi^2 c^2 \ln n$[7]
- 2009: $\ell \approx 5.35 c^2 \ln n$[8]
- 2011: $\ell \sim \frac{1}{2}\pi^2 c^2 \ln n$[9]

---

[6] B. Skoric et al., "Tardos fingerprinting is better than we thought," *IEEE Transactions on Information Theory*, vol. 54, no. 8, pp. 3663–3676, 2008.

[7] B. Skoric et al., "Symmetric Tardos fingerprinting codes for arbitrary alphabet sizes," *Designs, Codes and Cryptography*, vol. 46, no. 2, pp. 137–166, 2008.

[8] K. Nuida et al., "An improvement of discrete Tardos fingerprinting codes," *Designs, Codes and Cryptography*, vol. 52, no. 3, pp. 339–362, 2009.

[9] T. Laarhoven and B. de Weger, "Optimal symmetric Tardos traitor tracing schemes," *Designs, Codes and Cryptography*, vol. 71, no. 1, pp. 83–103, 2014.

TU/e

Previously on IH&MMSec 2013

Limitations of the symmetric Tardos scheme[10]

- Theorem: Using the symmetric score function, the current code length $\ell \sim \frac{1}{2}\pi^2 c^2 \ln n$ is asymptotically optimal
- Alternatively: Using the symmetric score function, it is impossible to achieve the fingerprinting capacity

---

[10] T. Laarhoven and B. de Weger, "Discrete distributions in the Tardos scheme, revisited," in *ACM Workshop on Information Hiding and Multimedia Security (IH&MMSec)*, 2013, pp. 13–18.

[11] J.-J. Oosterwijk et al., "Optimal suspicion functions for Tardos traitor tracing schemes," in *ACM Workshop on Information Hiding and Multimedia Security (IH&MMSec)*, 2013, pp. 19–28.

Previously on IH&MMSec 2013

Limitations of the symmetric Tardos scheme[10]

- Theorem: Using the symmetric score function, the current code length $\ell \sim \frac{1}{2}\pi^2 c^2 \ln n$ is asymptotically optimal
- Alternatively: Using the symmetric score function, it is impossible to achieve the fingerprinting capacity

[10] T. Laarhoven and B. de Weger, "Discrete distributions in the Tardos scheme, revisited," in *ACM Workshop on Information Hiding and Multimedia Security (IH&MMSec)*, 2013, pp. 13–18.

[11] J.-J. Oosterwijk et al., "Optimal suspicion functions for Tardos traitor tracing schemes," in *ACM Workshop on Information Hiding and Multimedia Security (IH&MMSec)*, 2013, pp. 19–28.

TU/e

Previously on IH&MMSec 2013

Limitations of the symmetric Tardos scheme[10]

- Theorem: Using the symmetric score function, the current code length $\ell \sim \frac{1}{2}\pi^2 c^2 \ln n$ is asymptotically optimal
- Alternatively: Using the symmetric score function, it is impossible to achieve the fingerprinting capacity

Optimize the score functions for fixed attacks[11]

- If scores are Gaussian, these score functions are optimal
- The 'interleaving defense' works against arbitrary attacks
- Score functions for other attacks work well, too!

---

[10] T. Laarhoven and B. de Weger, "Discrete distributions in the Tardos scheme, revisited," in *ACM Workshop on Information Hiding and Multimedia Security (IH&MMSec)*, 2013, pp. 13–18.

[11] J.-J. Oosterwijk et al., "Optimal suspicion functions for Tardos traitor tracing schemes," in *ACM Workshop on Information Hiding and Multimedia Security (IH&MMSec)*, 2013, pp. 19–28.

Previously on IH&MMSec 2013

Optimize the score functions for fixed attacks[15]

- If scores are Gaussian, these score functions are optimal
- The 'interleaving defense' works against arbitrary attacks
- Score functions for other attacks work well, too!

Previously on IH&MMSec 2013

Optimize the score functions for fixed attacks[15]

- If scores are Gaussian, these score functions are optimal
- The 'interleaving defense' works against arbitrary attacks
- Score functions for other attacks work well, too!

Open questions (not open anymore)

- Lower bounds: Are these score functions optimal?

Previously on IH&MMSec 2013

Optimize the score functions for fixed attacks[15]

- If scores are Gaussian, these score functions are optimal
- The 'interleaving defense' works against arbitrary attacks
- Score functions for other attacks work well, too!

Open questions (not open anymore)

- Lower bounds: Are these score functions optimal?
- Efficient decoders: Can we do even better?

TU/e

Lower bounds

Optimize the score functions for fixed attacks[15]

- If scores are Gaussian, these score functions are optimal
- The 'interleaving defense' works against arbitrary attacks
- Score functions for other attacks work well, too!

Open questions (not open anymore)

- Lower bounds: Are these score functions optimal?
- Efficient decoders: Can we do even better?

Lower bounds
Randomized construction

Assigning fingerprints to users, generating the code X

- Choose a parameter $p \in (0, 1)$
- For every segment i and user j: $\mathbb{P}(X_{j,i} = 1) = p$

Lower bounds

Randomized construction

Assigning fingerprints to users, generating the code X

- Choose a parameter $p \in (0, 1)$
- For every segment i and user j: $\mathbb{P}(X_{j,i} = 1) = p$

Finding the coalition $\mathcal{C} \subseteq \{1, \ldots, n\}$

- Simple decoding: Decide whether $j \in \mathcal{C}$ based on...
  - X: The information $X_{j,i}$ for all i
  - Y: The pirate output bits y
  - P: The parameter p
- Joint decoding: Decide whether $j \in \mathcal{C}$ based on...
  - X': The information $X_{k,i}$ for all i and all k
  - Y: The pirate output bits y
  - P: The parameter p

Lower bounds
Simple decoding

Finding the coalition $\mathcal{C} \subseteq \{1, \ldots, n\}$

- Simple decoding: Decide whether $j \in \mathcal{C}$ based on...
  - X: The information $X_{j,i}$ for all i
  - Y: The pirate output bits y
  - P: The parameter p

Lower bounds
Simple decoding

Finding the coalition $\mathcal{C} \subseteq \{1, \ldots, n\}$

- Simple decoding: Decide whether $j \in \mathcal{C}$ based on...
  - X: The information $X_{j,i}$ for all i
  - Y: The pirate output bits y
  - P: The parameter p

For fixed pirate strategies, the simple capacity is given by[5]

$$C^{\text{simple}} = \max_{p \in (0,1)} I(X; Y | P = p).$$

TU/e

Lower bounds
Simple decoding

Finding the coalition $\mathcal{C} \subseteq \{1, \ldots, n\}$

- Simple decoding: Decide whether $j \in \mathcal{C}$ based on...
  - X: The information $X_{j,i}$ for all i
  - Y: The pirate output bits y
  - P: The parameter p

For fixed pirate strategies, the simple capacity is given by[5]

$$C^{\text{simple}} = \max_{p \in (0,1)} I(X; Y | P = p).$$

$I(X; Y | P = p)$ is an explicit function of the strategy and p.

Lower bounds

Joint decoding

Finding the coalition $\mathcal{C} \subseteq \{1, \ldots, n\}$

- Joint decoding: Decide whether $j \in \mathcal{C}$ based on...
  - X': The information $X_{k,i}$ for all i and all k
  - Y: The pirate output bits $y_{\mathcal{S}}$
  - P: The parameter p

Lower bounds
Joint decoding

Finding the coalition $\mathcal{C} \subseteq \{1, \ldots, n\}$

- Joint decoding: Decide whether $j \in \mathcal{C}$ based on...
  - X': The information $X_{k,i}$ for all i and all k
  - Y: The pirate output bits $y_{\mathcal{S}}$
  - P: The parameter p

For fixed pirate strategies, the joint capacity is given by[5]

$$C^{\text{joint}} = \max_{p \in (0,1)} I(X'; Y | P = p).$$

TU/e

Lower bounds

Joint decoding

Finding the coalition $\mathcal{C} \subseteq \{1, \ldots, n\}$

- Joint decoding: Decide whether $j \in \mathcal{C}$ based on...
  - X': The information $X_{k,i}$ for all i and all k
  - Y: The pirate output bits $y_{\mathcal{S}}$
  - P: The parameter p

For fixed pirate strategies, the joint capacity is given by[5]

$$C^{\text{joint}} = \max_{p \in (0,1)} I(X'; Y | P = p).$$

$I(X'; Y | P = p)$ is an explicit function of the strategy and p.

Lower bounds
Pirate strategies

Common pirate strategies:

- Interleaving atk: Randomly choose a pirate, output his symbol
- All-1 attack: Always output a 1 if possible
- Majority voting: Always output the most received symbol
- Minority voting: Always output the least received symbol
- Coin-flip attack: Flip a fair coin to choose the output
- ...

Lower bounds

Results

Pirate strategy	C^{simple}	C^{joint}
(Unknown attacks)	$1/(2c^2 \ln 2)$[5]	$1/(2c^2 \ln 2)$[5]
Interleaving attack	$1/(2c^2 \ln 2)$[5]	$1/(2c^2 \ln 2)$[5]
All-1 attack	$\ln 2/c$	$1/c$
Majority voting	$1/(\pi c \ln 2)$	$1/c$
Minority voting	$\ln 2/c$	$1/c$
Coin-flip attack	$\ln 2/(4c)$	$\log_2(\frac{5}{4})/c$
...	...	...

Lower bounds

Results

Pirate strategy	C^{simple}	C^{joint}
(Unknown attacks)	$0.72/c^2$ [5]	$0.72/c^2$ [5]
Interleaving attack	$0.72/c^2$ [5]	$0.72/c^2$ [5]
All-1 attack	$0.69/c$	$1.00/c$
Majority voting	$0.46/c$	$1.00/c$
Minority voting	$0.69/c$	$1.00/c$
Coin-flip attack	$0.17/c$	$0.32/c$
...	...	...

Lower bounds
Results

Pirate strategy	C^{simple}	C^{joint}	Results
(Unknown attacks)	$0.72/c^2$ [5]	$0.72/c^2$ [5]	$0.72/c^2$
Interleaving attack	$0.72/c^2$ [5]	$0.72/c^2$ [5]	$0.72/c^2$
All-1 attack	$0.69/c$	$1.00/c$	$0.72/c$
Majority voting	$0.46/c$	$1.00/c$	$0.46/c$
Minority voting	$0.69/c$	$1.00/c$	$0.72/c$
Coin-flip attack	$0.17/c$	$0.32/c$	$0.36/c$
...	...	...	

TU/e

Lower bounds

Results

Pirate strategy	C^{simple}	C^{joint}	Results
(Unknown attacks)	$0.72/c^2$ [5]	$0.72/c^2$ [5]	$0.72/c^2$
Interleaving attack	$0.72/c^2$ [5]	$0.72/c^2$ [5]	$0.72/c^2$
All-1 attack	$0.69/c$	$1.00/c$	$0.72/c$
Majority voting	$0.46/c$	$1.00/c$	$0.46/c$
Minority voting	$0.69/c$	$1.00/c$	$0.72/c$
Coin-flip attack	$0.17/c$	$0.32/c$	$0.36/c$
...	...	...	

Under the Gaussian assumption, the score functions of Oosterwijk et al. perform better than what is theoretically possible!

Lower bounds

Results

Pirate strategy	C^{simple}	C^{joint}	Results
(Unknown attacks)	$0.72/c^2$ [5]	$0.72/c^2$ [5]	$0.72/c^2$
Interleaving attack	$0.72/c^2$ [5]	$0.72/c^2$ [5]	$0.72/c^2$
All-1 attack	$0.69/c$	$1.00/c$	$0.72/c$
Majority voting	$0.46/c$	$1.00/c$	$0.46/c$
Minority voting	$0.69/c$	$1.00/c$	$0.72/c$
Coin-flip attack	$0.17/c$	$0.32/c$	$0.36/c$
...	...	...	

Under the Gaussian assumption, the score functions of Oosterwijk et al. perform better than what is theoretically possible!

- Optimist: Those are great results!

Lower bounds

Results

Pirate strategy	C^{simple}	C^{joint}	Results
(Unknown attacks)	$0.72/c^2$ [5]	$0.72/c^2$ [5]	$0.72/c^2$
Interleaving attack	$0.72/c^2$ [5]	$0.72/c^2$ [5]	$0.72/c^2$
All-1 attack	$0.69/c$	$1.00/c$	$0.72/c$
Majority voting	$0.46/c$	$1.00/c$	$0.46/c$
Minority voting	$0.69/c$	$1.00/c$	$0.72/c$
Coin-flip attack	$0.17/c$	$0.32/c$	$0.36/c$
...	...	...	

Under the Gaussian assumption, the score functions of Oosterwijk et al. perform better than what is theoretically possible!

- Optimist: Those are great results!
- Realist: The Gaussian assumption may be wrong...

Lower bounds

Conclusion

Optimize the score functions for fixed attacks[15]

- If scores are Gaussian, these score functions are optimal
- The 'interleaving defense' works against arbitrary attacks
- Score functions for other attacks work well, too!

Open questions (not open anymore)

- Lower bounds: Are these score functions optimal?
- Efficient decoders: Can we do even better?

TU/e

Lower bounds
Conclusion

Optimize the score functions for fixed attacks[15]

- If scores are Gaussian, these score functions are optimal
- The 'interleaving defense' works against arbitrary attacks
- Score functions for other attacks work well, too!

Open questions (not open anymore)

- Lower bounds: Are these score functions optimal?
- Efficient decoders: Can we do even better?

Lower bounds
Conclusion

Optimize the score functions for fixed attacks[15]

- If scores are Gaussian, these score functions are optimal
- The 'interleaving defense' works against arbitrary attacks
- Score functions for other attacks work well, too!

Open questions (not open anymore)

- Lower bounds: Are these score functions optimal? **No.**
- Efficient decoders: Can we do even better?

Efficient decoders

Introduction

Optimize the score functions for fixed attacks[15]

- If scores are Gaussian, these score functions are optimal
- The 'interleaving defense' works against arbitrary attacks
- Score functions for other attacks work well, too!

Open questions (not open anymore)

- Lower bounds: Are these score functions optimal? **No.**
- Efficient decoders: Can we do even better?

Efficient decoders

Fixed attacks

Neyman-Pearson lemma[12]:

Given some data $\mathcal{D}$, the most powerful test (of size α) to distinguish between two hypotheses H_0 and H_1 is to test if, for some constant η_α,

$$\Lambda(\mathcal{D}) = \frac{\mathbb{P}(\mathcal{D} \mid H_0)}{\mathbb{P}(\mathcal{D} \mid H_1)} \leq \eta_\alpha. \tag{1}$$

[12] J. Neyman and E. S. Pearson, "On the problem of the most efficient tests of statistical hypotheses," *Philosophical Transactions of the Royal Society of London. Series A*, vol. 231, no. 694-706, pp. 289–337, 1933.

TU/e

Efficient decoders

Fixed attacks

Neyman-Pearson lemma[12]:

Given some data $\mathcal{D} = \{X, Y\}$, the most powerful test (of size α) to distinguish between two hypotheses H_0 and H_1 is to test if, for some constant η_α,

$$\Lambda(\mathcal{D}) = \frac{\mathbb{P}(\mathcal{D} \mid H_0)}{\mathbb{P}(\mathcal{D} \mid H_1)} \leq \eta_\alpha. \tag{1}$$

[12] J. Neyman and E. S. Pearson, "On the problem of the most efficient tests of statistical hypotheses," *Philosophical Transactions of the Royal Society of London. Series A*, vol. 231, no. 694-706, pp. 289–337, 1933.

Efficient decoders

Fixed attacks

Neyman-Pearson lemma[12]:
Given some data $\mathcal{D} = \{X, Y\}$, the most powerful test (of size α) to distinguish between two hypotheses $H_0 : j \in \mathcal{C}$ and H_1 is to test if, for some constant η_α,

$$\Lambda(\mathcal{D}) = \frac{\mathbb{P}(\mathcal{D} \mid H_0)}{\mathbb{P}(\mathcal{D} \mid H_1)} \leq \eta_\alpha. \quad (1)$$

[12] J. Neyman and E. S. Pearson, "On the problem of the most efficient tests of statistical hypotheses," *Philosophical Transactions of the Royal Society of London. Series A*, vol. 231, no. 694-706, pp. 289–337, 1933.

Efficient decoders

Fixed attacks

Neyman-Pearson lemma[12]:
Given some data $\mathcal{D} = \{X, Y\}$, the most powerful test (of size α) to distinguish between two hypotheses $H_0 : j \in \mathcal{C}$ and $H_1 : j \notin \mathcal{C}$ is to test if, for some constant η_α,

$$\Lambda(\mathcal{D}) = \frac{\mathbb{P}(\mathcal{D} \mid H_0)}{\mathbb{P}(\mathcal{D} \mid H_1)} \leq \eta_\alpha. \quad (1)$$

[12] J. Neyman and E. S. Pearson, "On the problem of the most efficient tests of statistical hypotheses," *Philosophical Transactions of the Royal Society of London. Series A*, vol. 231, no. 694-706, pp. 289–337, 1933.

Efficient decoders

Fixed attacks

Neyman-Pearson lemma[12]:
Given some data $\mathcal{D} = \{X, Y\}$, the most powerful test (of size α) to distinguish between two hypotheses $H_0 : j \in \mathcal{C}$ and $H_1 : j \notin \mathcal{C}$ is to test if, for some constant η_α,

$$\Lambda(\mathcal{D}) = \frac{\mathbb{P}(X, Y \mid j \in \mathcal{C})}{\mathbb{P}(X, Y \mid j \notin \mathcal{C})} \leq \eta_\alpha. \quad (1)$$

[12] J. Neyman and E. S. Pearson, "On the problem of the most efficient tests of statistical hypotheses," *Philosophical Transactions of the Royal Society of London. Series A*, vol. 231, no. 694-706, pp. 289–337, 1933.

Efficient decoders

Fixed attacks

Neyman-Pearson lemma[12]:
Given some data $\mathcal{D} = \{X, Y\}$, the most powerful test (of size α) to distinguish between two hypotheses $H_0 : j \in \mathcal{C}$ and $H_1 : j \notin \mathcal{C}$ is to test if, for some constant η_α,

$$\Lambda(\mathcal{D}) = \frac{\mathbb{P}(X, Y \mid j \in \mathcal{C})}{\mathbb{P}(X, Y \mid j \notin \mathcal{C})} \leq \eta_\alpha. \tag{1}$$

Likelihood ratio $\Lambda(\mathcal{D})$ corresponds to the 'score function' and *provably* achieves capacity for fixed attacks.

[12] J. Neyman and E. S. Pearson, "On the problem of the most efficient tests of statistical hypotheses," *Philosophical Transactions of the Royal Society of London. Series A*, vol. 231, no. 694-706, pp. 289–337, 1933.

Efficient decoders

Arbitrary attacks

Results of Abbe and Zheng[13][14]:
Given some data $\mathcal{D}$, the best test to distinguish between two hypotheses H_0 and $\mathcal{H}_a = \{H_1, H_2, \ldots\}$ is to test H_0 against the worst-case attack $H_a^* \in \mathcal{H}_a$ using likelihood ratios.

[13] E. Abbe and L. Zheng, "Linear universal decoding for compound channels," *IEEE Transactions on Information Theory*, vol. 56, no. 12, pp. 5999–6013, 2012.

[14] P. Meerwald and T. Furon, "Toward practical joint decoding of binary Tardos fingerprinting codes," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 4, pp. 1168–1180, 2012.

Efficient decoders

Arbitrary attacks

Results of Abbe and Zheng[13][14]:

Given some data $\mathcal{D}$, the best test to distinguish between two hypotheses H_0 and $\mathcal{H}_a = \{H_1, H_2, \dots\}$ is to test H_0 against the worst-case attack $H_a^* \in \mathcal{H}_a$ using likelihood ratios.

- $H_1, H_2, \dots$ correspond to different pirate strategies

[13] E. Abbe and L. Zheng, "Linear universal decoding for compound channels," *IEEE Transactions on Information Theory*, vol. 56, no. 12, pp. 5999–6013, 2012.

[14] P. Meerwald and T. Furon, "Toward practical joint decoding of binary Tardos fingerprinting codes," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 4, pp. 1168–1180, 2012.

TU/e

Efficient decoders

Arbitrary attacks

Results of Abbe and Zheng[13][14]:

Given some data $\mathcal{D}$, the best test to distinguish between two hypotheses H_0 and $\mathcal{H}_a = \{H_1, H_2, \ldots\}$ is to test H_0 against the worst-case attack $H_a^* \in \mathcal{H}_a$ using likelihood ratios.

- $H_1, H_2, \ldots$ correspond to different pirate strategies
- Worst-case attack is typically quite complicated

---

[13] E. Abbe and L. Zheng, "Linear universal decoding for compound channels," *IEEE Transactions on Information Theory*, vol. 56, no. 12, pp. 5999–6013, 2012.

[14] P. Meerwald and T. Furon, "Toward practical joint decoding of binary Tardos fingerprinting codes," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 4, pp. 1168–1180, 2012.

Efficient decoders

Arbitrary attacks

Results of Abbe and Zheng[13][14]:
Given some data $\mathcal{D}$, the best test to distinguish between two hypotheses H_0 and $\mathcal{H}_a = \{H_1, H_2, \ldots\}$ is to test H_0 against the worst-case attack $H_a^* \in \mathcal{H}_a$ using likelihood ratios.

- $H_1, H_2, \ldots$ correspond to different pirate strategies
- Worst-case attack is typically quite complicated
- Replace 'worst-case attack' with 'asympt. worst-case attack'

---

[13] E. Abbe and L. Zheng, "Linear universal decoding for compound channels," *IEEE Transactions on Information Theory*, vol. 56, no. 12, pp. 5999–6013, 2012.

[14] P. Meerwald and T. Furon, "Toward practical joint decoding of binary Tardos fingerprinting codes," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 4, pp. 1168–1180, 2012.

TU/e

Efficient decoders

Arbitrary attacks

Results of Abbe and Zheng[13][14]:

Given some data $\mathcal{D}$, the best test to distinguish between two hypotheses H_0 and $\mathcal{H}_a = \{H_1, H_2, \dots\}$ is to test H_0 against the worst-case attack $H_a^* \in \mathcal{H}_a$ using likelihood ratios.

- $H_1, H_2, \dots$ correspond to different pirate strategies
- Worst-case attack is typically quite complicated
- Replace 'worst-case attack' with 'asympt. worst-case attack'
  - Asymptotic worst-case attack is the interleaving attack

[13] E. Abbe and L. Zheng, "Linear universal decoding for compound channels," *IEEE Transactions on Information Theory*, vol. 56, no. 12, pp. 5999–6013, 2012.

[14] P. Meerwald and T. Furon, "Toward practical joint decoding of binary Tardos fingerprinting codes," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 4, pp. 1168–1180, 2012.

TU/e

Efficient decoders

Arbitrary attacks

Results of Abbe and Zheng[13][14]:

Given some data $\mathcal{D}$, the best test to distinguish between two hypotheses H_0 and $\mathcal{H}_a = \{H_1, H_2, \dots\}$ is to test H_0 against the worst-case attack $H_a^* \in \mathcal{H}_a$ using likelihood ratios.

- $H_1, H_2, \dots$ correspond to different pirate strategies
- Worst-case attack is typically quite complicated
- Replace 'worst-case attack' with 'asympt. worst-case attack'
  - Asymptotic worst-case attack is the interleaving attack
  - Leads to simple expressions and asymptotic optimal decoder

---

[13] E. Abbe and L. Zheng, "Linear universal decoding for compound channels," *IEEE Transactions on Information Theory*, vol. 56, no. 12, pp. 5999–6013, 2012.

[14] P. Meerwald and T. Furon, "Toward practical joint decoding of binary Tardos fingerprinting codes," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 4, pp. 1168–1180, 2012.

TU/e

Efficient decoders

Optimized decoders for fixed attacks

- Decoders provably achieve capacity for given attacks
- Motivated by the Neyman-Pearson lemma
- No (incorrect) Gaussian assumption needed

Universal decoder for arbitrary attacks

- Log-likelihood decoder for the interleaving attack is optimal
- Motivated by results of Abbe and Zheng
- No Gaussian assumption needed (but scores are Gaussian)
- No more cut-offs on the distribution function!

Efficient decoders

Optimize the score functions for fixed attacks[15]

- If scores are Gaussian, these score functions are optimal
- The 'interleaving defense' works against arbitrary attacks
- Score functions for other attacks work well, too!

Open questions (not open anymore)

- Lower bounds: Are these score functions optimal? **No.**
- Efficient decoders: Can we do even better?

Efficient decoders

Optimize the score functions for fixed attacks[15]

- If scores are Gaussian, these score functions are optimal
- The 'interleaving defense' works against arbitrary attacks
- Score functions for other attacks work well, too!

Open questions (not open anymore)

- Lower bounds: Are these score functions optimal? **No.**
- Efficient decoders: Can we do even better? **Yes, we can!**

Conclusion

Explicit asymptotics of the capacities of various models[15]

- Information-theoretic approach: Mutual information game
- Both simple (efficient) and joint (optimal) decoding
- Can be applied to arbitrary pirate strategies

Capacity-achieving decoders for arbitrary models[16]

- Statistical approach: Neyman-Pearson hypothesis testing
- Both simple and joint decoding
- Asymptotically optimal regardless of the pirate attack
- 'Interleaving decoder' is an improved universal decoder

---

[15] T. Laarhoven, "Asymptotics of fingerprinting and group testing: tight bounds from channel capacities," *submitted to IEEE Transactions on Information Theory*, pp. 1–14, 2014.

[16] T. Laarhoven, "Asymptotics of fingerprinting and group testing: capacity-achieving log-likelihood decoders," *submitted to IEEE Transactions on Information Theory*, pp. 1–13, 2014.

Questions?