Tardos fingerprinting codes

Boris Škorić

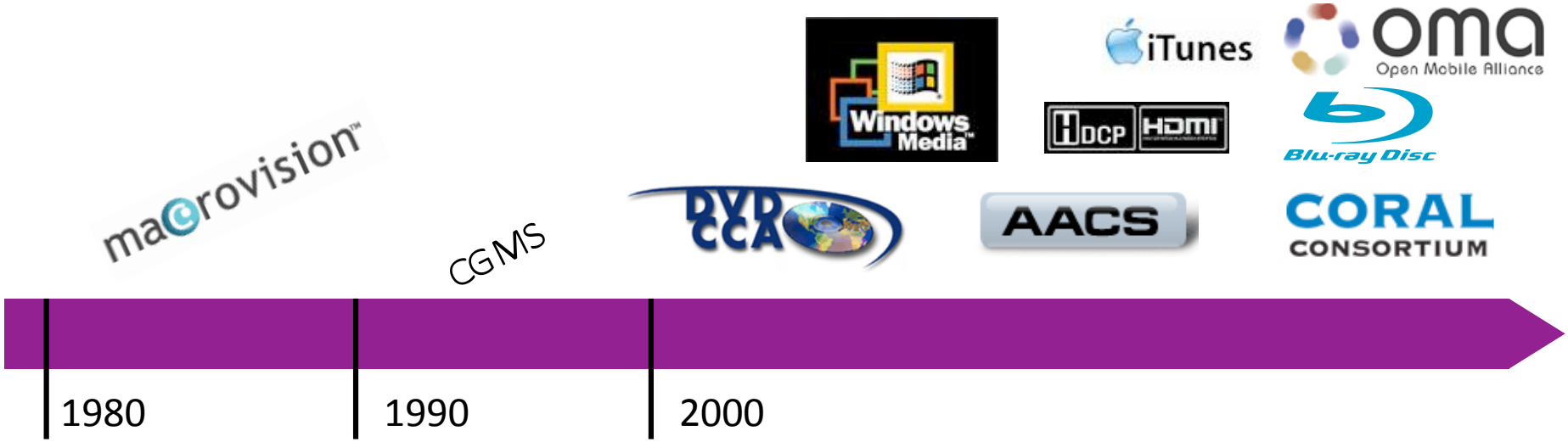
Seminar Univ. Darmstadt, 29 April 2009

joint work with Stefan Katzenbeisser and Mehmet Celik

Outline

- Trends in content protection
- Collusion-resistant watermarking
- The Tardos fingerprinting system
- Our improvements

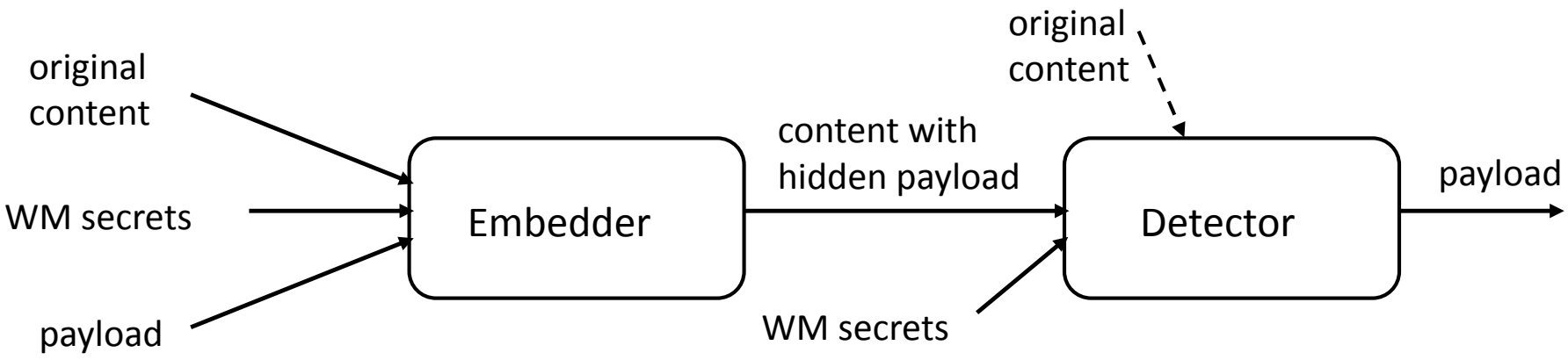
Trends in content protection



- Consumers increasingly dislike DRM
- Vista content protection spec

  "longest suicide note in history" (Gutmann 2006)
- "Disembodied" distribution ⇒ Hard to DRM-protect; Easy to watermark
- April 2007: EMI announces DRM-free music
- Gradual shift from copy prevention to distribution tracking

Watermarking (a.k.a. Fingerprinting)



<u>Forensic tracking</u>

- Payload = unique identifier of recipient
- Redistribution traced back to source

<u>Examples</u>

- Jan.2004: Man arrested for distributing oscar screeners.
- Digital cinema

Collusion attacks

"Coalition of pirates"

- Users pool their content
- Differences point to watermark
- Attackers remove watermark

■ = "detectable positions"

pirate #1	1	1	1	0	1	0	1	0	0	0	0	1
#2	1	0	1	0	1	0	1	0	1	0	1	1
#3	1	0	1	0	1	0	1	0	0	0	1	1
#4	1	1	1	0	0	0	1	1	0	0	0	1
Attacked Content	1	0/1	1	0	0/1	0	1	0/1	0/1	0	0/1	1

Collusion-resistant watermarking

Requirements

- Resistance against $c \leq c_0$ attackers
- Low False Positive error rate
- Low False Negative error rate
- ... and all that with small watermark payload! (7bits/min. video)

Attack model

- "**Marking assumption**": Modification only at detectable positions
- Several options
  - Restricted digit model: Choice from available symbols only
  - Unreadable digit model: Erasure allowed
  - Arbitrary digit model: Arbitrary symbol (but not erasure)
  - General digit model

equivalent for binary symbols

An apology

I work on this subject just because I like the math.

History of collusion resistance: Code length

Construction

[Hollmann et al 1998: IPP codes, c_0=2]

Boneh and Shaw 1998: $m = \mathcal{O}(c_0^4 \ln[n/\eta] \ln[1/\eta]), \quad q = 2$

Tardos 2003: $m = 100 c_0^2 \ln(1/\varepsilon_1), \quad q = 2, \quad \varepsilon_2 = \varepsilon_1^{c_0/4}$

Chor et al 2000: $m = 4c_0^2 \log n, \quad q = 2c_0^2$

Staddon et al 2001: $m = c_0^2 \log_q(n), \quad q > m - 2$

Tardos 2003:

$$m = \Omega(c_0^2 \ln[1/\varepsilon_1]), \quad q \text{ arbitrary}$$

Boneh and Shaw 1998:

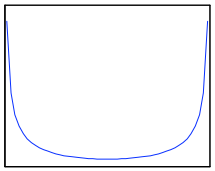
$$m = \Omega(c_0 \ln[1/c_0\eta]), \quad q = 2$$

Anthapadmanabhan et al. 2007: m = $\Omega(c_0)$

Proven lower bound

n = #users
m = code length in symbols
q = alphabet size
ε_1 = Prob[accuse specific innocent]
η = Prob[not all accused are guilty]
ε_2 = False Negative prob.

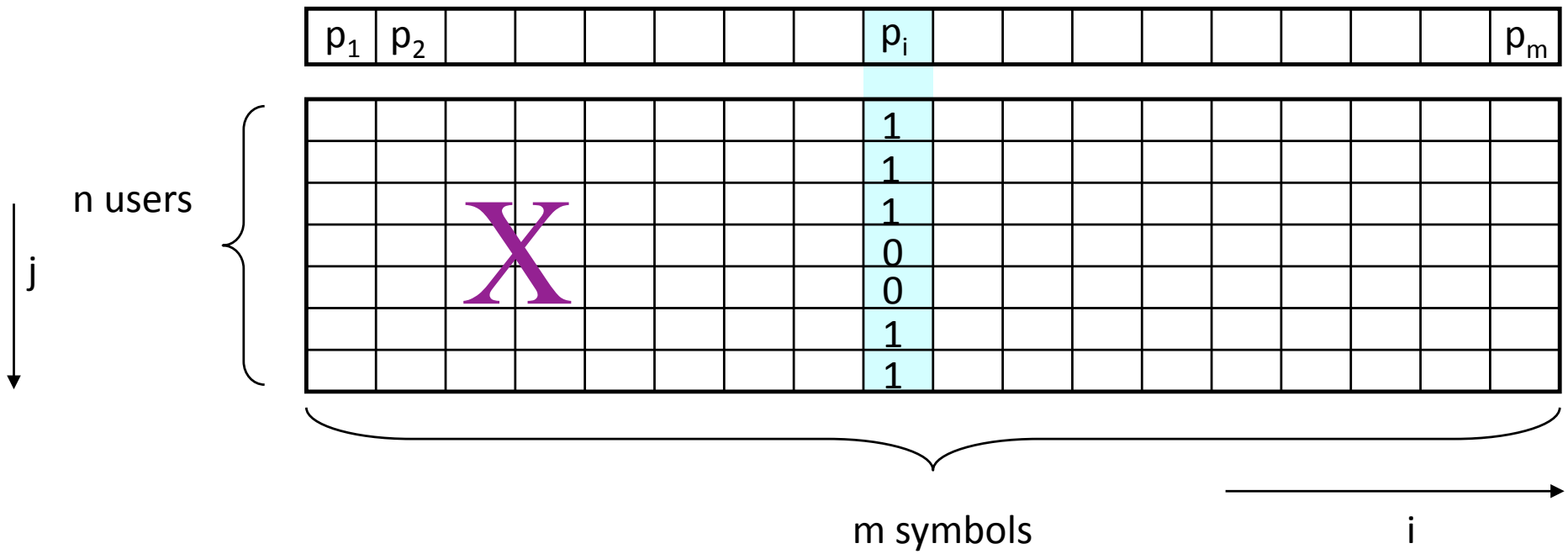
The Tardos scheme (2003)



Step 1: Generate iid $p_1,\ldots, p_m \in (0,1)$; Prob. density

$$f(p) \propto \frac{1}{\sqrt{p(1-p)}}$$

Step 2: Generate columns of X; **Prob[X_{ji}=1] = p_i**.



The Tardos scheme (continued)

3. Embed X into the content for users 1...n.
   Store p and X.
4. Coalition publishes content with payload y = {y_1,..., y_m}.
5. Content owner computes accusation values:

$$S_j = \sum_{i=1}^{m} y_i \, g(X_{ji}, p_i) \qquad g(1,p) = \sqrt{\frac{1-p}{p}} \quad g(0,p) = -\sqrt{\frac{p}{1-p}}$$

User #j gets accused if S_j > threshold

Proven properties:

- Resistance against ≤c_0 colluders with $m = 100c_0^2 \ln(1/\varepsilon_1), \quad \varepsilon_2 = \varepsilon_1^{c_0/4}$ independent of coalition strategy
- If c > c_0, then
  - **no false accusations!**
  - high prob. that nobody gets accused

Our contributions

Škorić, Katzenbeisser, Celik; 2007
Škorić, Vladimirova, Celik, Talstra; 2006

<u>New construction of X and new accusation</u>

- Arbitrary alphabet size q.
- Dirichlet distribution
- Symbol-symmetric accusation, even for q=2.

$$p_i \to \{p_{i\alpha}\}_{\alpha=0}^{q-1}$$

$$f(p) \to f(\{p_\alpha\}) \propto \prod_\alpha p_\alpha^{-1+\kappa}$$

$$S_j \to \sum_i g(X_{ji} == y_i,\ p_{i,y_i})$$

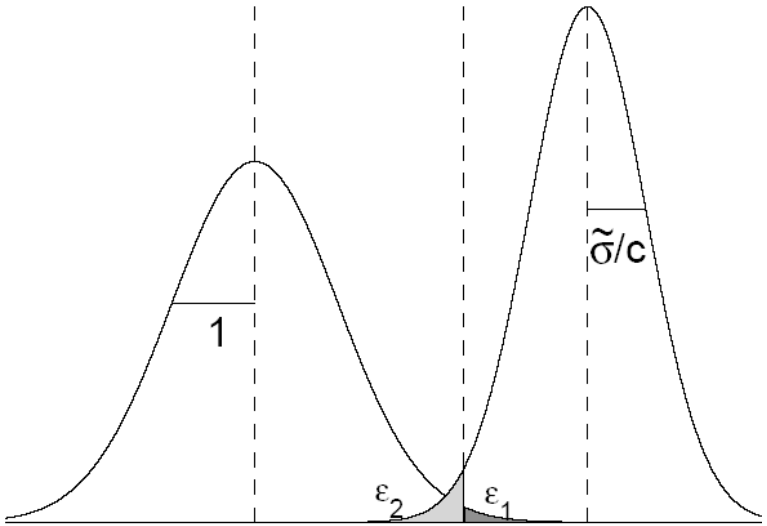
								p_{i0}									
								p_{i1}									
								p_{i2}									
								p_{i3}									

								2									
								0									
								2									
				X				2									
								3									
								2									
								2									

Our contributions (continued)

<u>Decouple ε_2 from ε_1.</u>

- Desired FP and FN properties are independent
- Usually $\varepsilon_1 << \varepsilon_2$ instead of vice versa



<u>New analysis method based on statistics</u>

- Compute FN and FP error rate from $E[S_j]$ and $E[S_j^2]$
- Central Limit Theorem:
  For large c_0, accusations have Gaussian distribution

Assumptions:

- Column-symmetric coalition strategy
- Restricted digit model when q>2.

Results

Tardos 2003

$$\varepsilon_2 = \varepsilon_1^{c_0/4}$$

Binary alphabet,

$$m = 100 c_0^2 \ln \frac{1}{\varepsilon_1}$$

Symmetric scheme
& statistical analysis for large coalitions

ε_1, ε_2 independent (usually $\varepsilon_1 << \varepsilon_2$)

Binary alphabet:

$$m = \frac{\pi^2}{2} c_0^2 \ln \frac{1}{\varepsilon_1 \sqrt{2\pi}} \cdot \left\{ 1 + \mathcal{O}\left(\sqrt{\frac{q \ln \varepsilon_2}{c_0 \ln \varepsilon_1}}\right) \right\}$$

q-ary alphabet: further reduction of m

q= 3: 35%

q=10: 80%

Summary

- Tardos scheme is asymptotically best known
- Two improvements
  1. Symbol-symmetric binary scheme
  2. Statistical analysis method

  Together: code length reduced by factor ≈ 20 for large c_0
- Scheme for arbitrary alphabet size.

  In restricted digit model $\Rightarrow$ further reduction of code length.

Open questions / future work

- How much can be gained from tweaking the f function?
  - Nuida et al: Gauss-Legendre quadrature for q=2
  - how much more can be achieved?
  - equivalent construction for q>2?
- How Gaussian are the accusations for small c_0?
  - CLT convergence; work in progress with Antonino Simone
- Is there a construction that has the $(c_0)^1$ asymptotics?
- More general attack models
  - "combined digit model",
    work in progress with Katzenbeisser, Celik, Schaathun