Generalized tally-based decoders for traitor tracing and group testing

Boris Škorić and Wouter de Groot

Eindhoven University of Technology

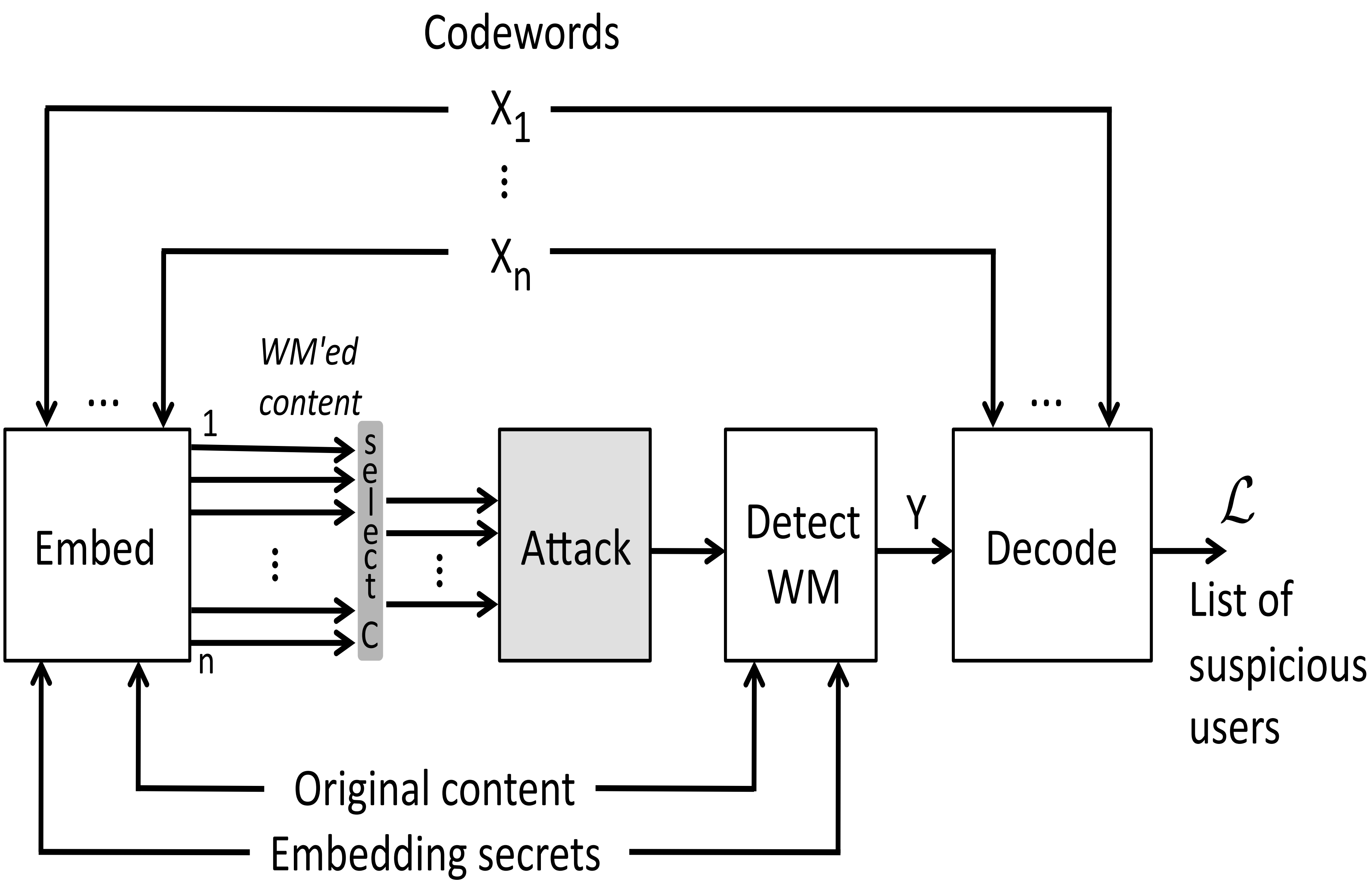
TU/e

WIFS 2015
November 19

Outline

- Collusion attacks on watermarks
- Tardos codes
- Attack vs. defense: game theory
- Decoders
  - Neyman-Pearson scores
  - composite symbols
- Group testing

Forensic watermarking



Collusion attacks

"Coalition of pirates"

- Attackers compare their content
- Differences point to watermark
- Try to remove watermark

Collusion-resistant watermarking

<u>Requirements</u>

- Resistance against c_0 attackers
- Low False Positive and False Negative error rate
- small watermark payload!

<u>Attack model</u>

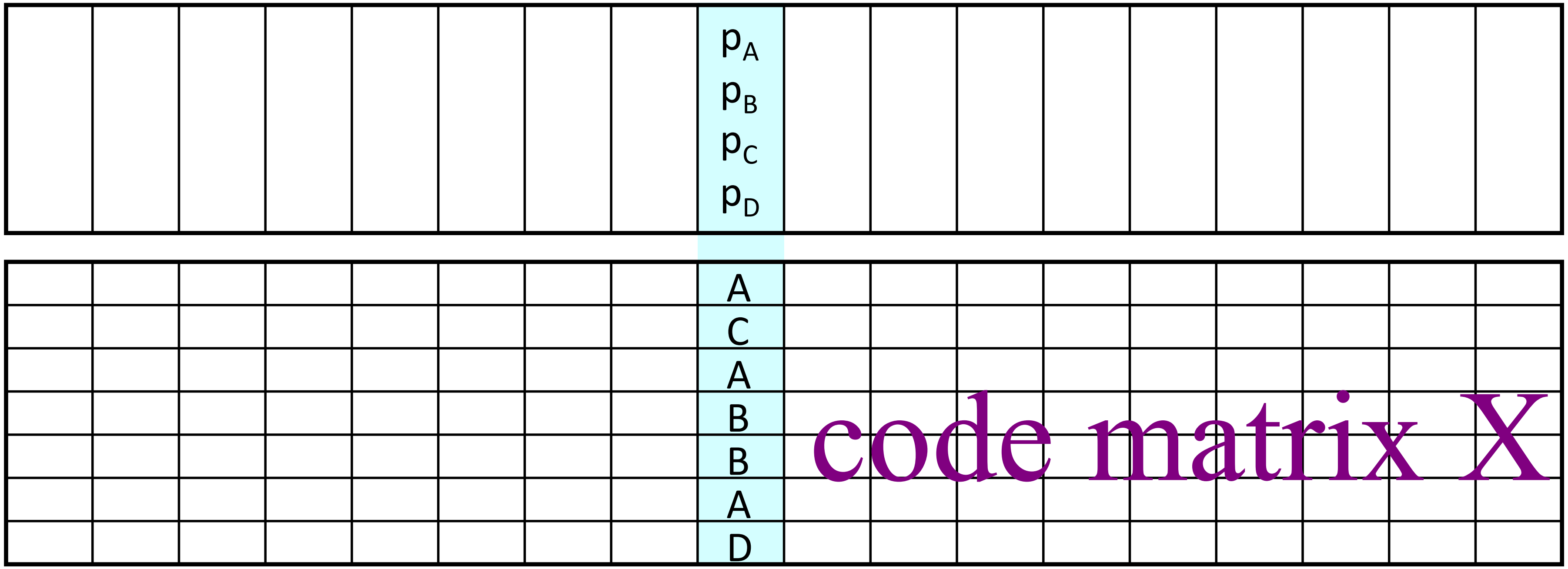
- Discrete positions with embedded symbols
- **Restricted digit model**: Choice from available symbols only

Bias-based code [Tardos 2003, ŠKC 2007]

Alphabet Q of size q

Step 1:
For each position, generate bias vector $\mathbf{p}=(p_\alpha)_{\alpha\in Q}$. $|\mathbf{p}|=1$ $\mathbf{p} \sim F$

Step 2:
For each position and user, draw watermark symbol: Pr[symbol α] = p_α.

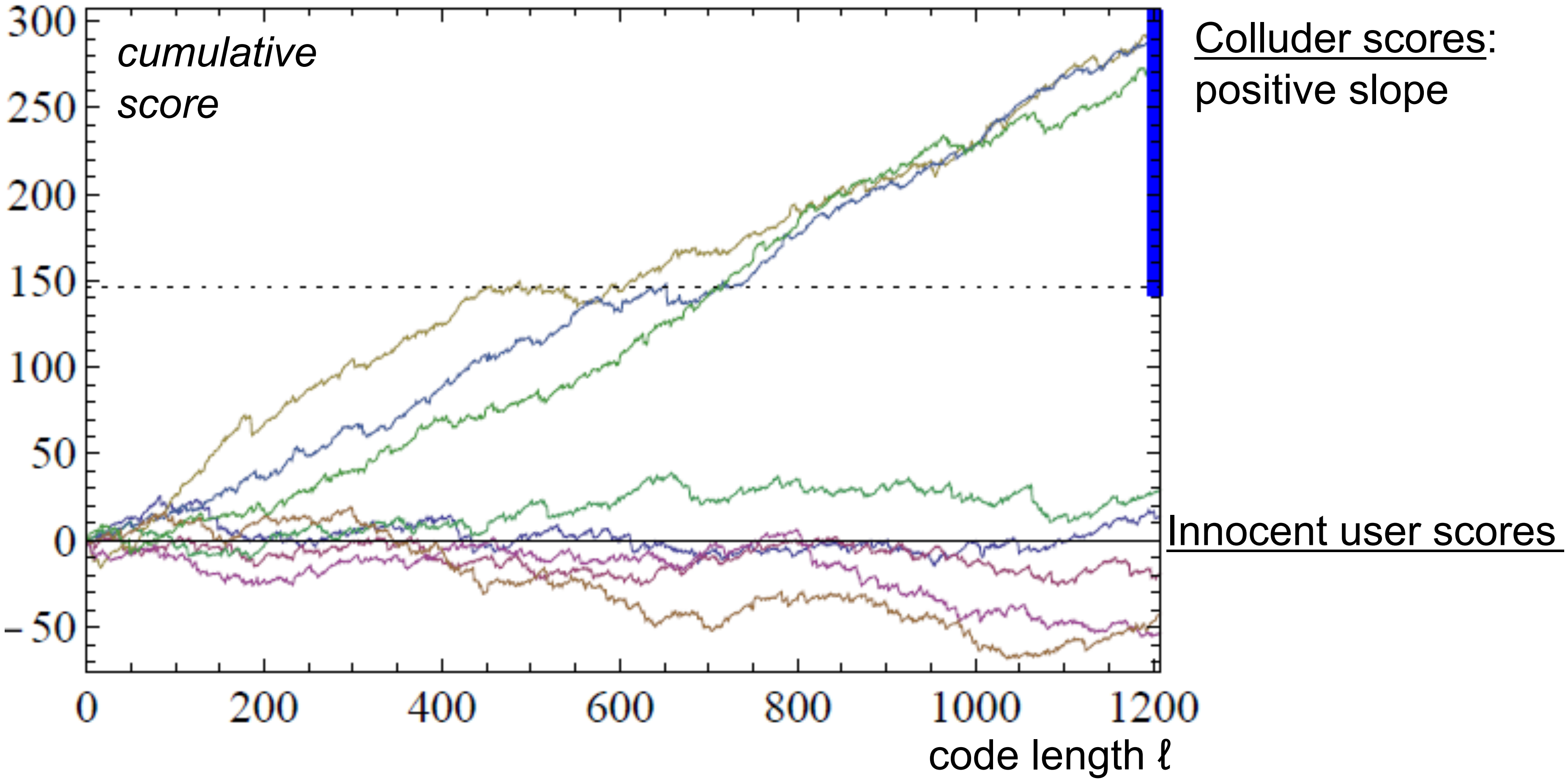


pirated copy carries watermark y

Step 3:
Find attackers based on X and y

Asymptotically optimal scaling: code length $\propto c_0^2$

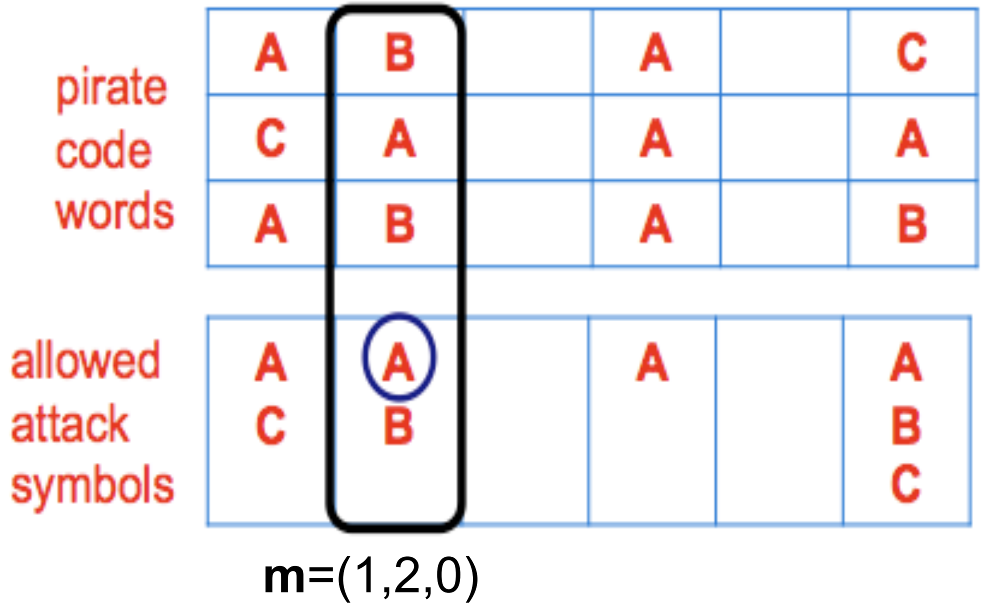
Separating the attackers from the innocents



Collusion channel (in Restricted Digit Model)

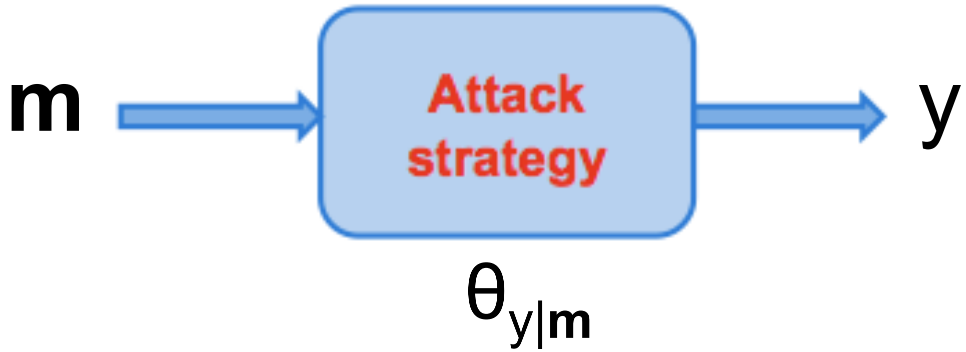
"Tally" vector **m**:

- #colluders = c
- m_α= #α received by colluders
- |**m**|=c



Attack:

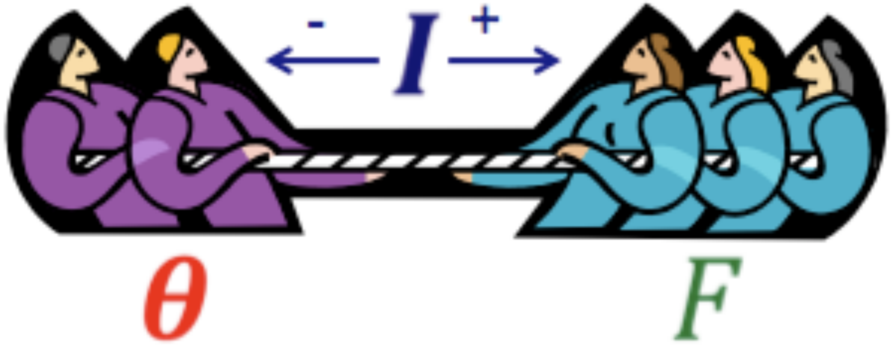
- Same strategy in each position (asymptotically strongest)
- Choose y as a function of **m**: $\theta_{y|\mathbf{m}}$ = Prob[output y given **m**]



Information theory approach

[Moulin 2008]

- Collusion attack is "malicious noise".
- Use techniques from channel coding!
  - How much does Y reveal about **M**?
    (**M** is equivalent to colluder identities)
  - *Mutual information* I(**M**;Y)

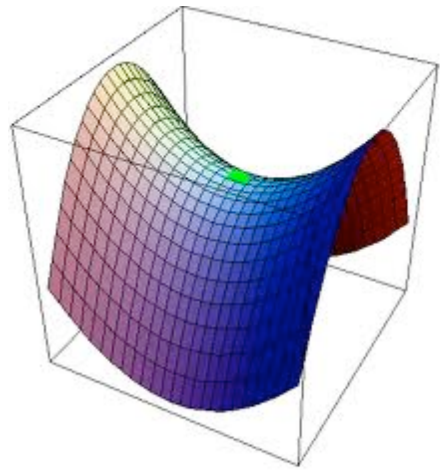


Game theory:

- Pay-off function I(**M**;Y|**P**)
- Tracer chooses bias distribution F(**p**)
- Colluders choose strategy θ

Fingerprinting capacity

$$C = \frac{1}{c} \max_F \min_\theta I(M;Y \mid P)$$



saddle point

Asymptotic saddlepoint

[Huang+Moulin 2012]

q-ary alphabet.
Pay-off function I(**M**;Y|**P**).

$$F(\boldsymbol{p}) \propto \prod_{\alpha \in Q} p_\alpha^{-1/2}$$

With increasing c,

- optimal bias distribution gets closer to **Jeffreys prior**.
- optimal attack gets closer to **Interleaving attack**.

$$\theta_{y|\boldsymbol{m}} = \frac{m_y}{c}$$

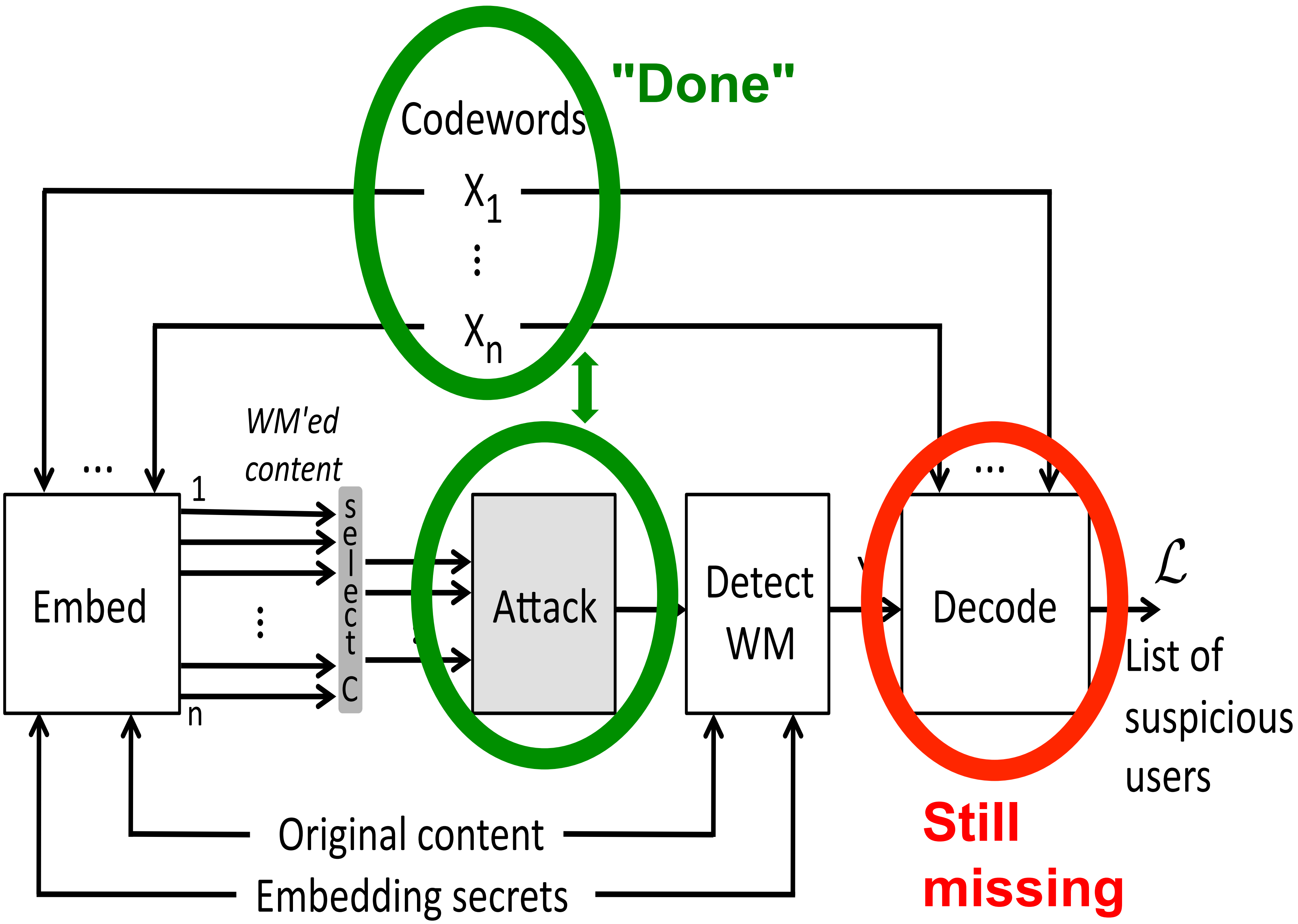
(pick random attacker)

Asymptotic capacity result

$$\ell_{\text{sufficient}} = \frac{2c^2}{q-1} \ln \frac{n}{P_{\text{FP}}}$$

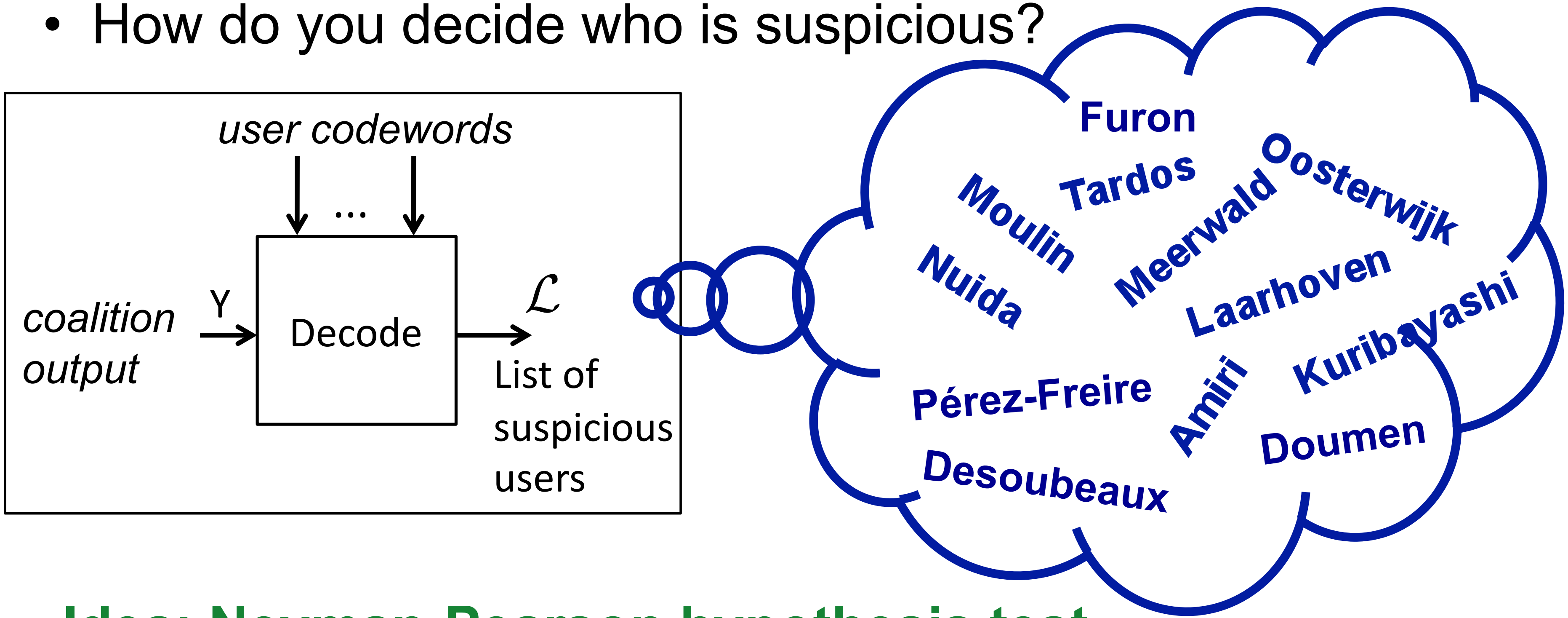
Larger q is better

[Boesten+Škorić 2011]


"Done"
Codewords
X_1
X_n
WM'ed content
1
n
select
Embed
Attack
Detect WM
Decode
$\mathcal{L}$
List of suspicious users
Original content
Embedding secrets
Still missing

Decoding

- Capacity analysis says nothing about the decoder!
- How do you decide who is suspicious?



Idea: Neyman-Pearson hypothesis test.

- best P_{FN} at given P_{FP}
- best P_{FP} at given P_{FN}

Neyman-Pearson scores

Hypothesis H_j : "j is part of the coalition".

Neyman-Pearson score:

$$S_j = \frac{\Pr[\ H_j \mid \text{evidence}]}{\Pr[\neg H_j \mid \text{evidence}]}$$

If S_j > threshold Z, then consider j to be guilty.

Assume colluder symmetry and position symmetry:

S_j equivalent to $$\ln \frac{\mathbb{E}_{\bar{\boldsymbol{M}}|x,j\in\mathcal{C}} \prod_{i\in[\ell]} \theta_{y_i|\boldsymbol{M}_i}}{\mathbb{E}_{\bar{\boldsymbol{M}}|x,j\notin\mathcal{C}} \prod_{i\in[\ell]} \theta_{y_i|\boldsymbol{M}_i}}$$

1. **Score depends on (unknown) strategy θ**.
2. Expectation $E_{\ldots}$ means: **sum over all possible coalitions of size c**.

Neyman-Pearson scores (2)

Problems:

1. **Score depends on (unknown) strategy θ.**
2. Expectation $E_{...}$: **sum over all possible coalitions of size c.**

Solutions:

1. Theorem by Abbe and Zheng (2010): $\boldsymbol{\theta}_{\text{saddlepoint}}$ gives *Universal Decoder*.

   - insert the Interleaving attack

2. "Forget" part of the evidence. "Remember" only x_j and
   - biases **p** (Laarhoven 2014)
   - symbol tallies (Škorić 2014)
   - composite-symbol tallies. **NEW!**

Neyman-Pearson scores (3)

Laarhoven score:

$$\delta_{xy} \ln(1 + \frac{1}{c-1} \cdot \frac{1}{p_y})$$

Škorić 2014:

$$\delta_{xy} \ln(1 + \frac{1}{c-1} \cdot \frac{n-1}{t_y - 1})$$

(s=1)

Global tally t_y = #users who received symbol y

p	p	p	p	p	p
A	A	A	C	D	A
C	C	B	B	B	C
A	D	B	A	B	C
B	C	A	B	A	D
B	D	B	B	C	C
A	C	B	A	C	B
D	B	C	C	B	C

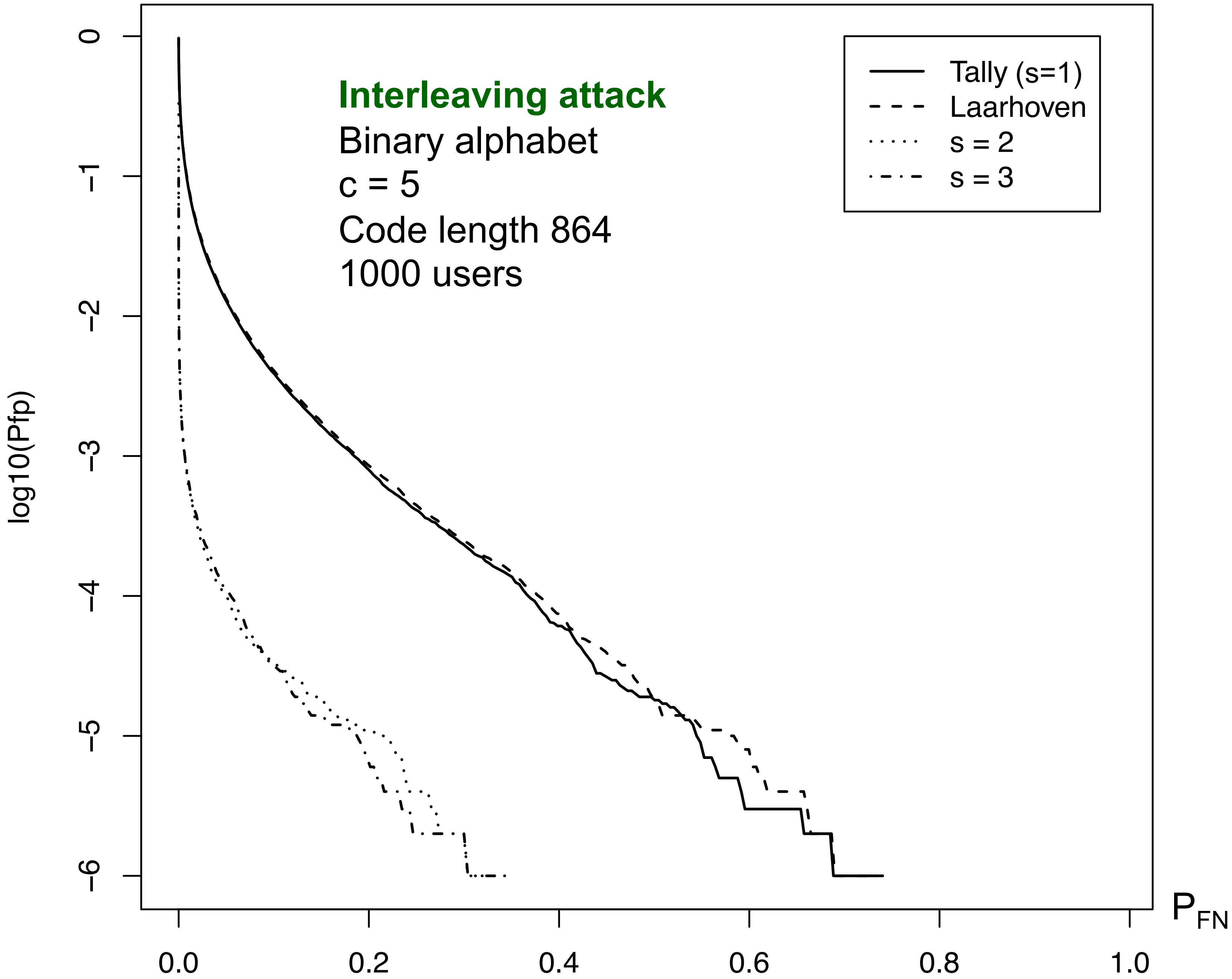
→ *composite symbols "DBC", "CBC"*

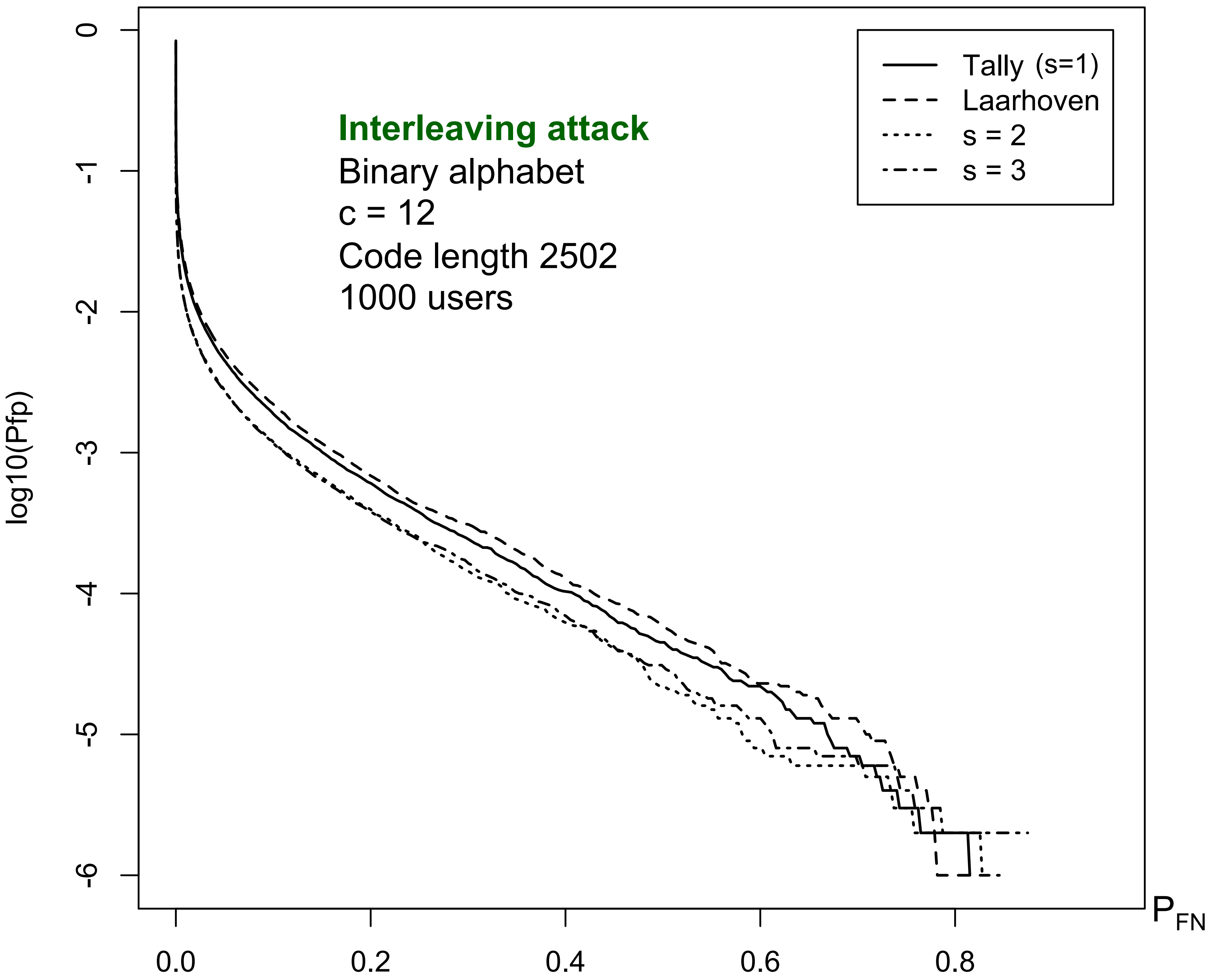
s columns combined

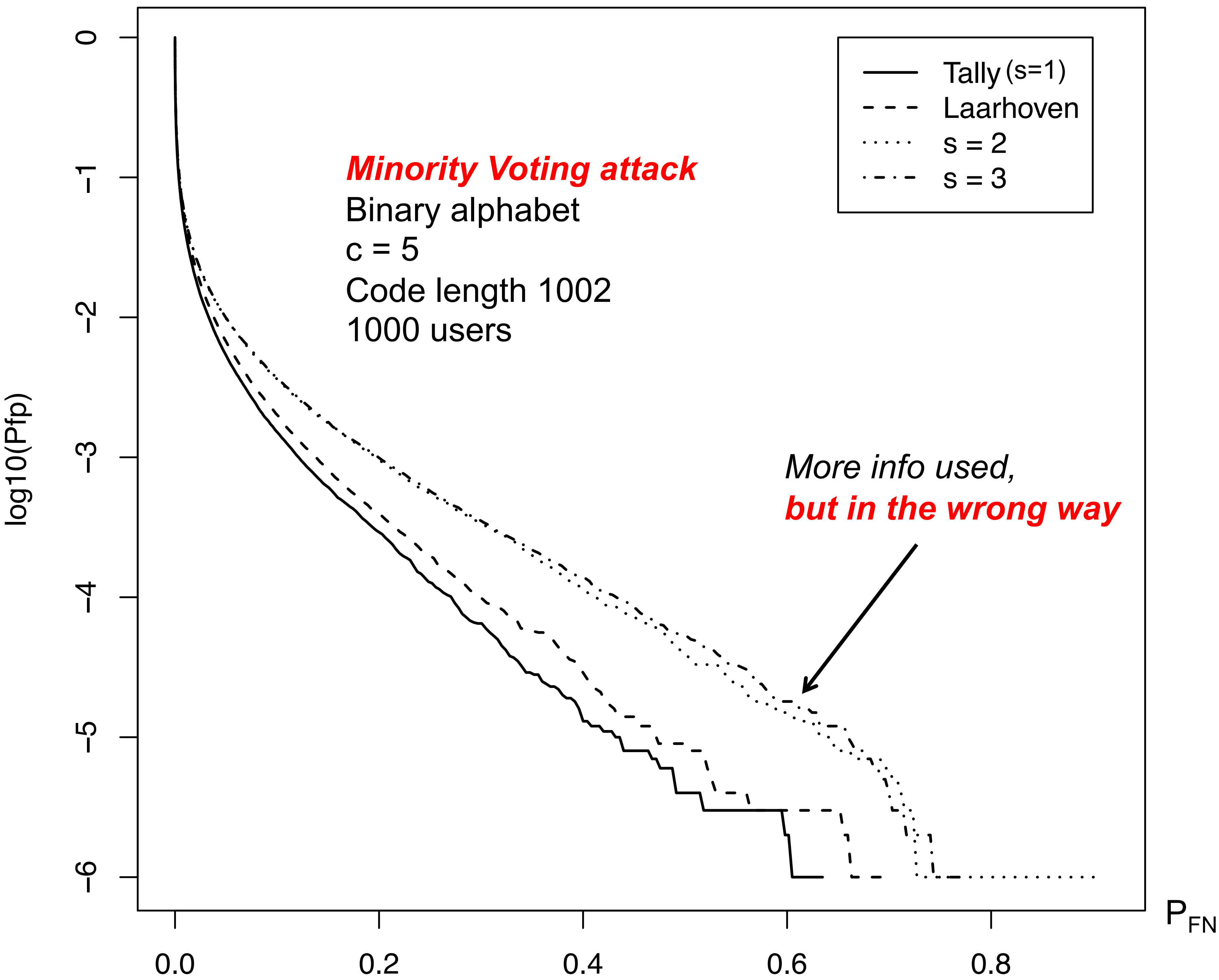
NEW IDEA:

Use more info by combining columns

simulation software: Wouter de Groot




Interleaving attack
Binary alphabet
c = 12
Code length 2502
1000 users
Tally (s=1)
Laarhoven
s = 2
s = 3
log10(Pfp)
0
-1
-2
-3
-4
-5
-6
0.0
0.2
0.4
0.6
0.8
P FN



Minority Voting attack
Binary alphabet
c = 5
Code length 1002
1000 users
Tally (s=1)
Laarhoven
s = 2
s = 3
More info used,
but in the wrong way
log10(Pfp)
0
−1
−2
−3
−4
−5
−6
0.0
0.2
0.4
0.6
0.8
P FN

How to combine score functions

Battery of score functions

- The bad decoders cause False Negative, **not False Positive**!
- The good decoders catch the colluders

Group testing

Real-life problem in epidemology:

- Blood samples from n people
- Expensive test => too few tests
- Long duration => tests in parallel
- Combine blood samples

Traitor Tracing	***Group Testing***
colluder	infected
symbol 0/1	1 = included in test 0 = not included
code length	number of tests
arbitrary attack θ	θ = All1 attack

Fixed "attack"

⇓

The Neyman-Pearson approach to construct score functions is particularly well suited to Group Testing.

Summary

<u>Composite symbol tally</u>:

- Improved Traitor Tracing at "small" c
- Improved Group Testing

<u>Still to be done</u>:

- Further validation
  - simulations, provable bounds, etc.
  - q>2
  - Group Testing numerics etc.
- *Dynamic* scenarios
  - different conditions, different solutions?
- More realistic attack models
  - Combined Digit Model, noisy medical tests, ...

$$g_2(\xi,\lambda,\boldsymbol{t}) = \ln[-1+\frac{n-2}{n-c}\cdot \tag{16}$$

$$\frac{(c-1)t^{\{1\}}_{\lambda[1]}t^{\{2\}}_{\lambda[2]}+(n-c)t_\lambda}{(c-1)(t^{\{1\}}_{\lambda[1]}-\delta_{\xi[1]\lambda[1]})(t^{\{2\}}_{\lambda[2]}-\delta_{\xi[2]\lambda[2]})+(n-1-c)(t_\lambda-\delta_{\xi\lambda})}].$$

$$g_3(\xi,\lambda,\boldsymbol{t}) = \ln[-1+\frac{n-3}{n-c}\cdot\frac{A_3}{B_3}],\ \text{with} \tag{18}$$

$$\begin{aligned}
A_3 &= c^{(3)}t^{\{1\}}_{\lambda[1]}t^{\{2\}}_{\lambda[2]}t^{\{3\}}_{\lambda[3]} \\
&\quad +c^{(2)}(n-c)(t^{\{1,2\}}_{\lambda[12]}t^{\{3\}}_{\lambda[3]}+t^{\{1,3\}}_{\lambda[13]}t^{\{2\}}_{\lambda[2]}+t^{\{2,3\}}_{\lambda[23]}t^{\{1\}}_{\lambda[1]}) \\
&\quad +c(n-c)(n-2c)t_\lambda \qquad (19) \\
B_3 &= A_3 \text{ with } \boldsymbol{t}\to\boldsymbol{t}-\boldsymbol{e}_\xi,\ \ n\to n-1 \qquad (20)
\end{aligned}$$