RAINBOW: A Robust And Invisible Non-Blind Watermark for Network Flows

Amir Houmansadr

Negar Kiyavash

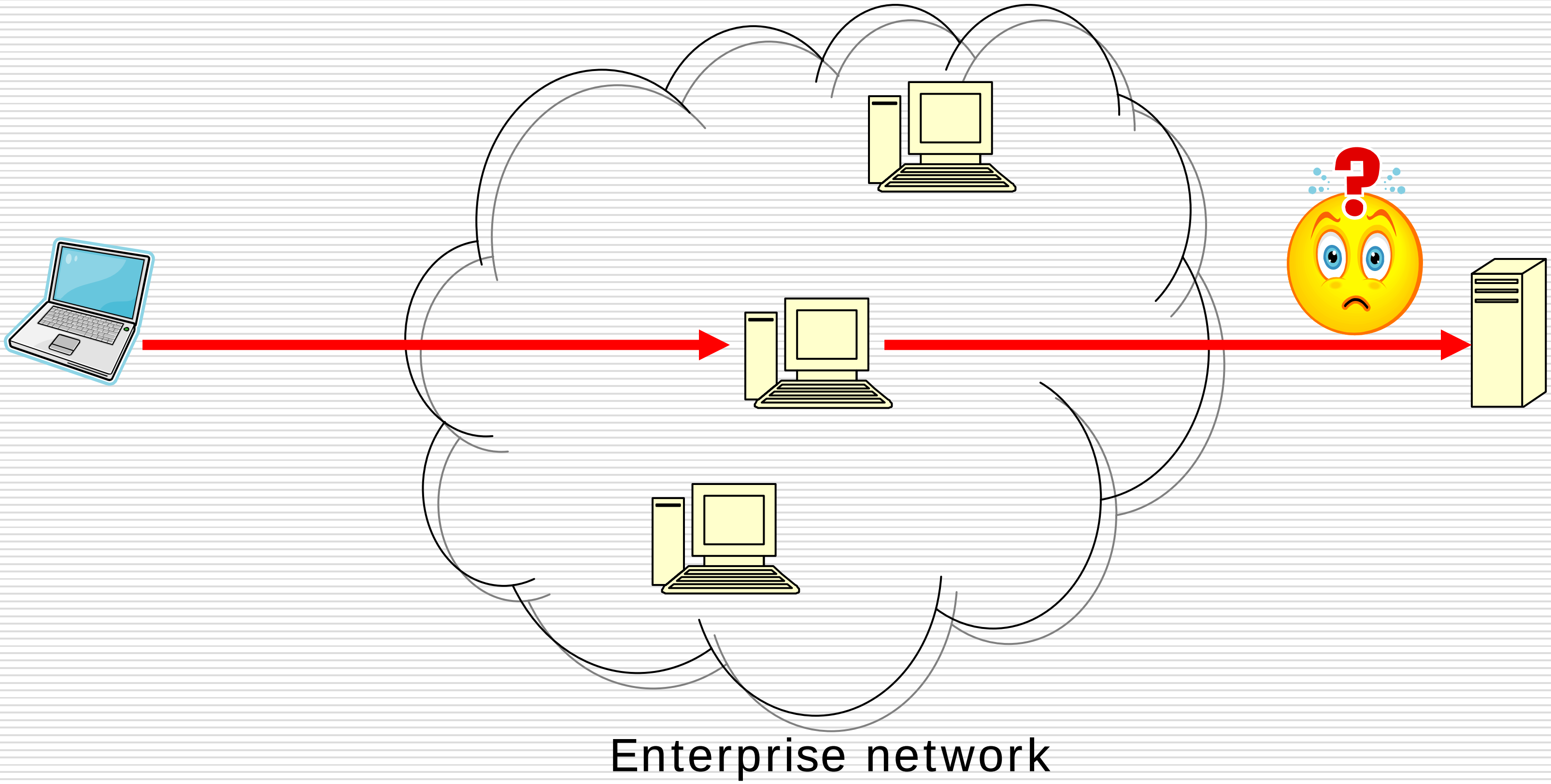
Nikita Borisov

University of Illinois at Urbana-Champaign

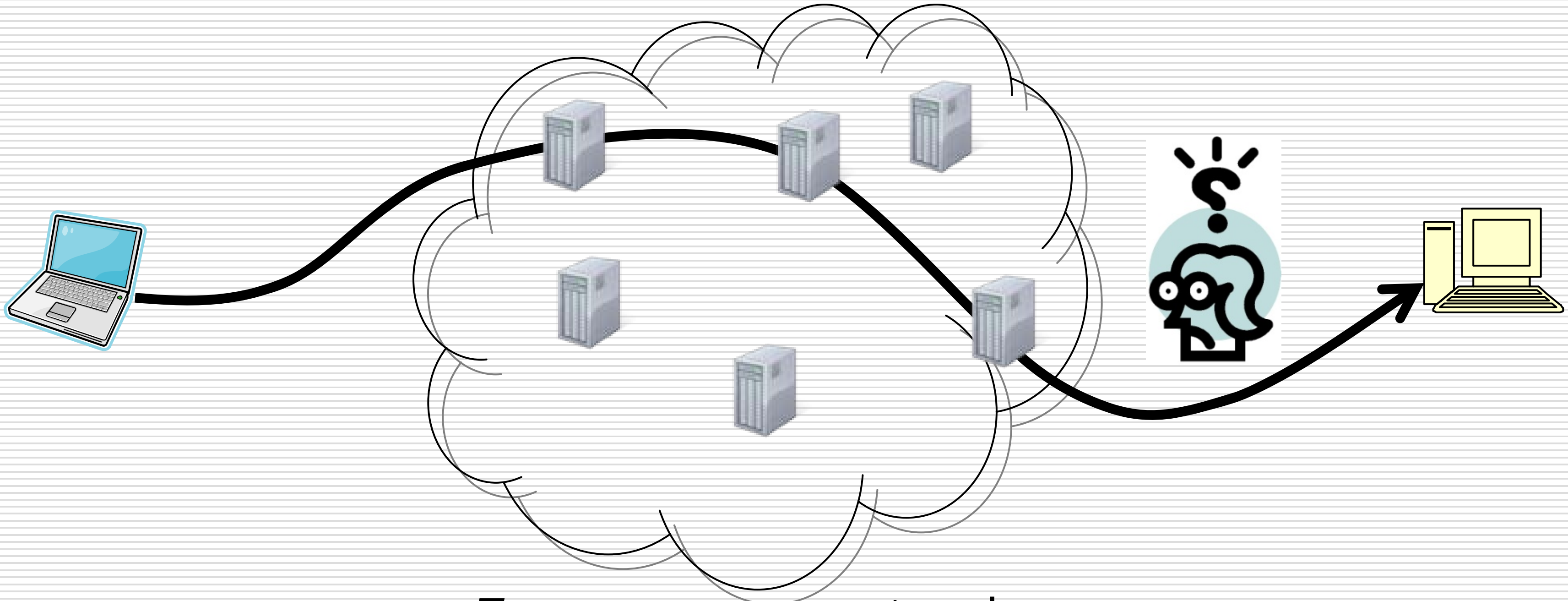
Traffic analysis

- Low-latency traffic analysis
  - Intrusion detection
  - Compromising anonymous networks

Stepping stone detection



Compromising Anonymity

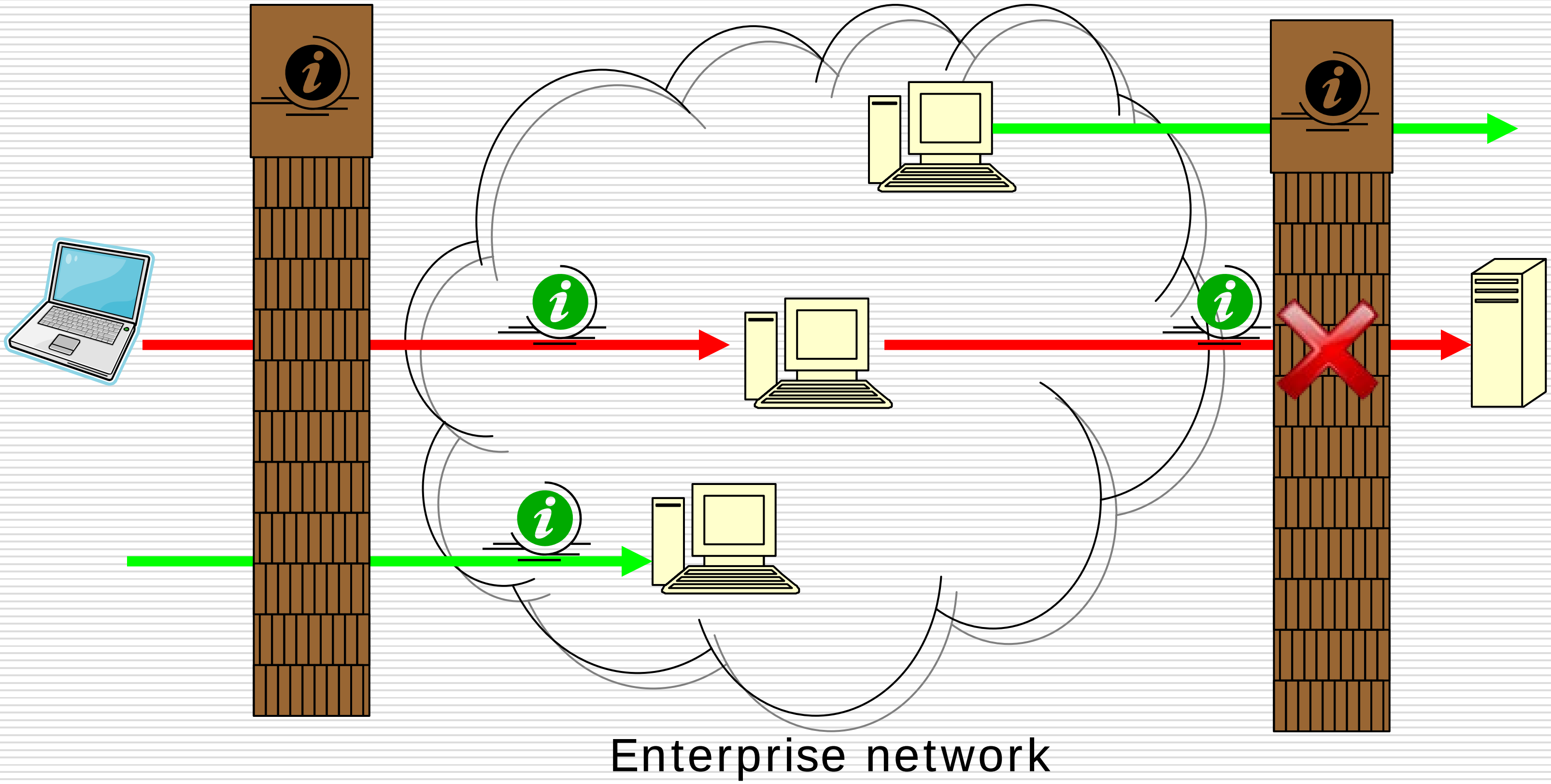


Tor anonymous network

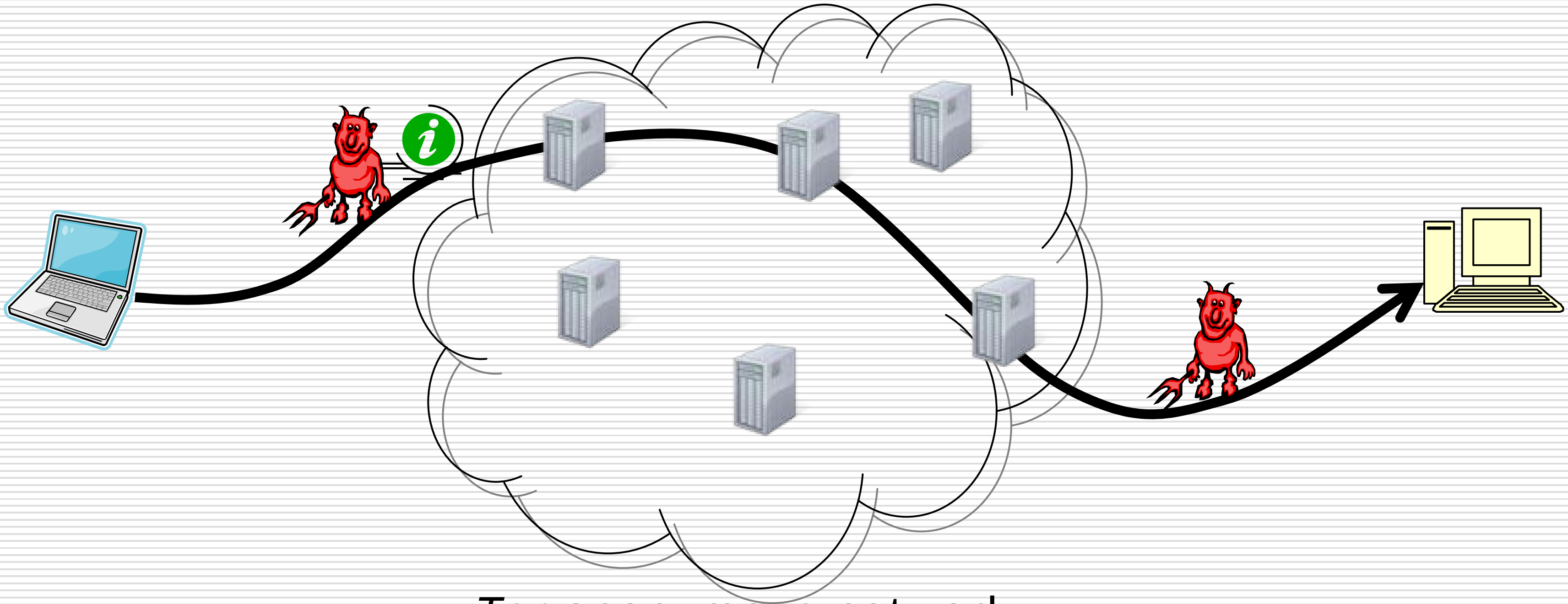
Traffic analysis

- Passive
  - Analyzing original packet counts, timing, ...
  - Common Problem: low efficiency
    - Slow decision (not real time) , high false errors, ...
- Active (watermarking)
  - Motivation: improve efficiency
  - Using modified packet timing, count, rate, ...
  - Multimedia watermarking: QIM, Patchwork, ...

Stepping stone detection



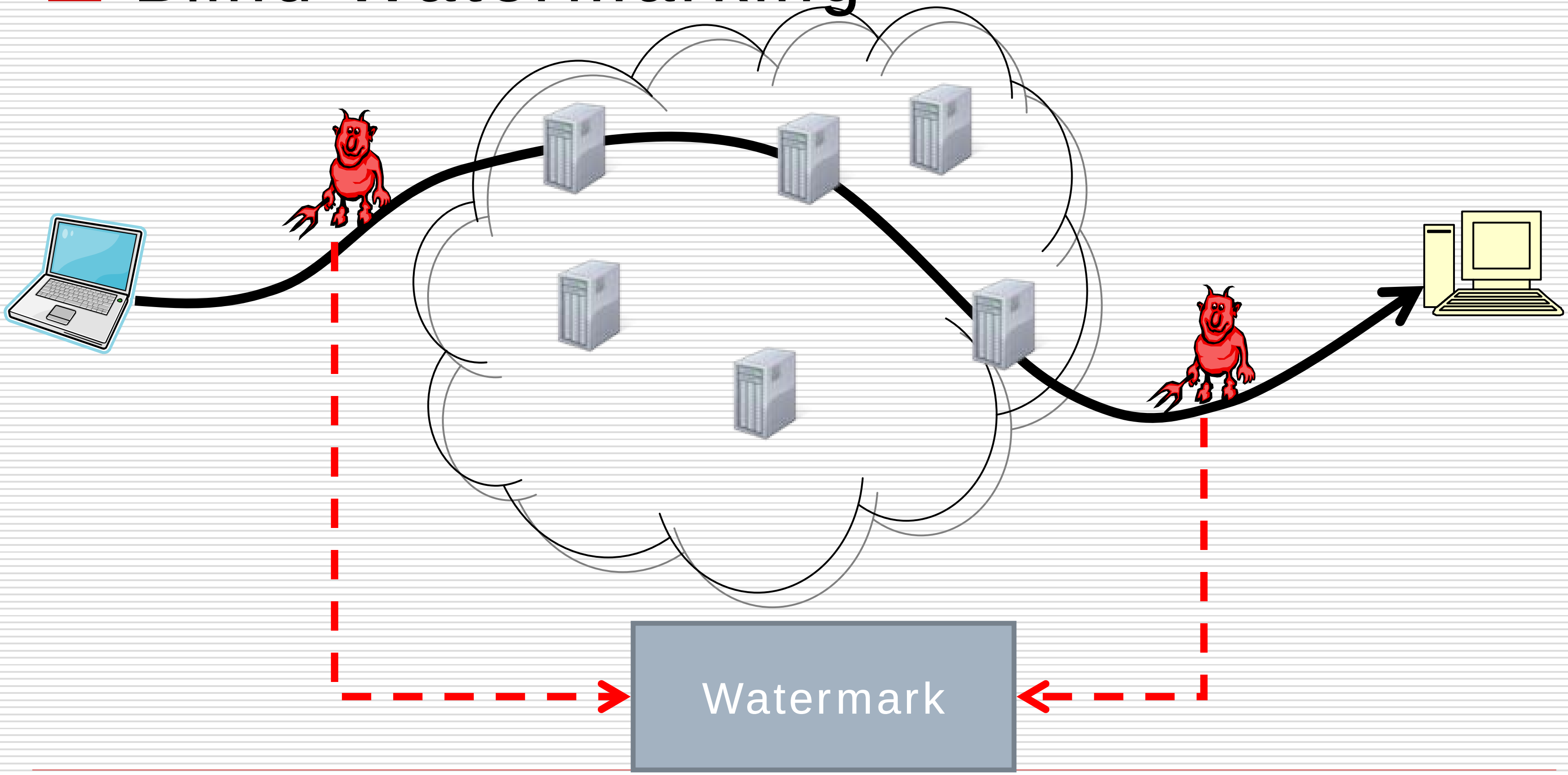
Compromising Anonymity



Tor anonymous network

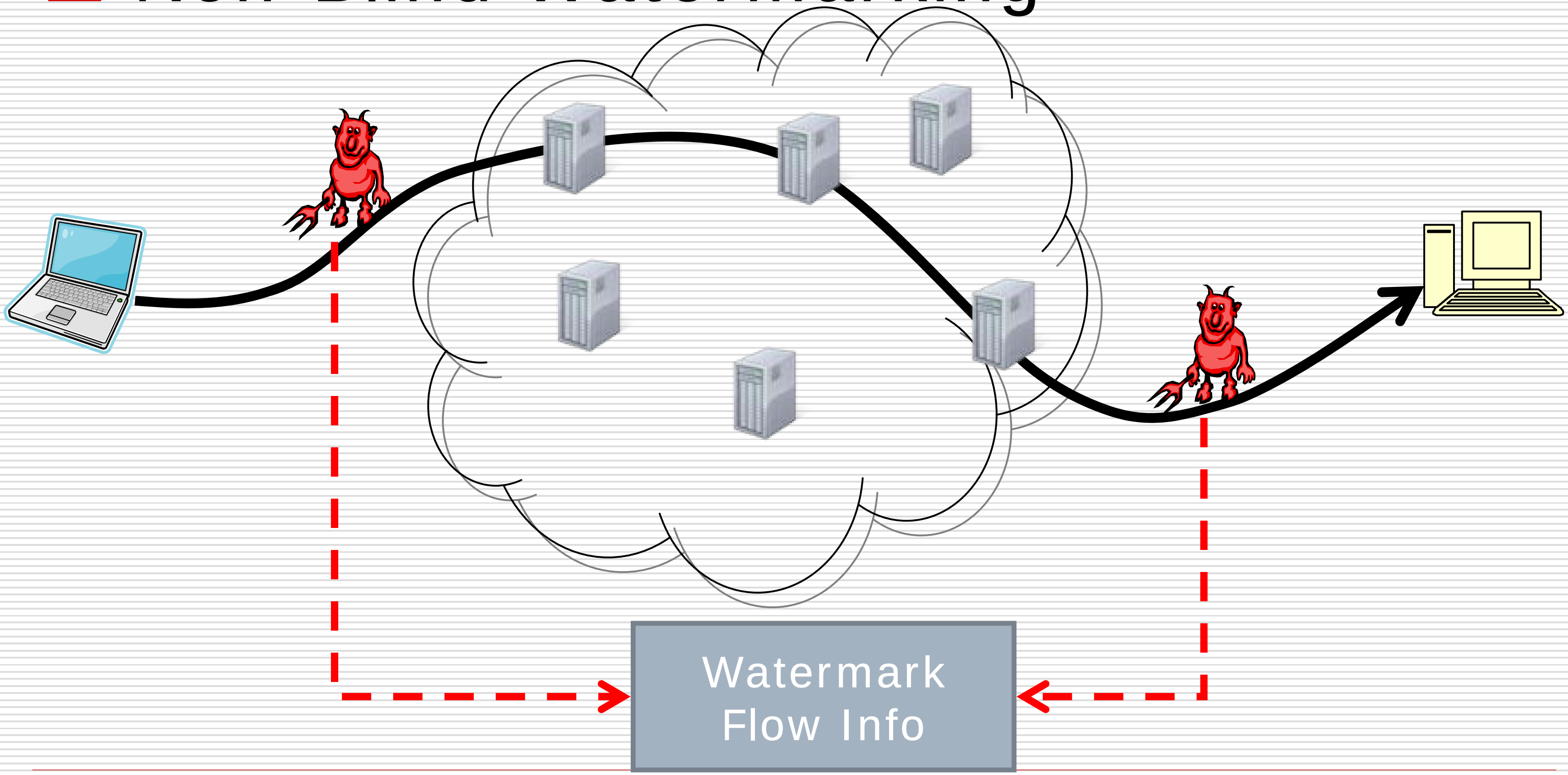
Terminology

- Blind Watermarking



Terminology

- Non-Blind Watermarking



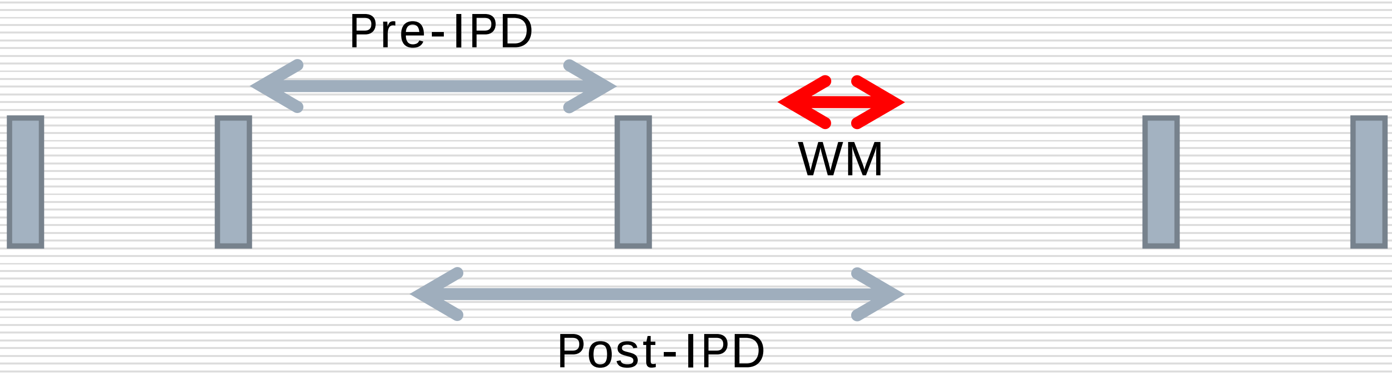
Motivation of RAINBOW

- **Watermarking**: efficient detection
- Common Problem with watermarking
  - Blind: Lack of **Invisibility**
    - Legitimate-user disturbance
    - Subject to attacks
- **Non-Blind**: in middle of passive schemes and active blind schemes
- **Robust** to network perturbations

Robust and Invisible Non-Blind Watermark
RAINBOW

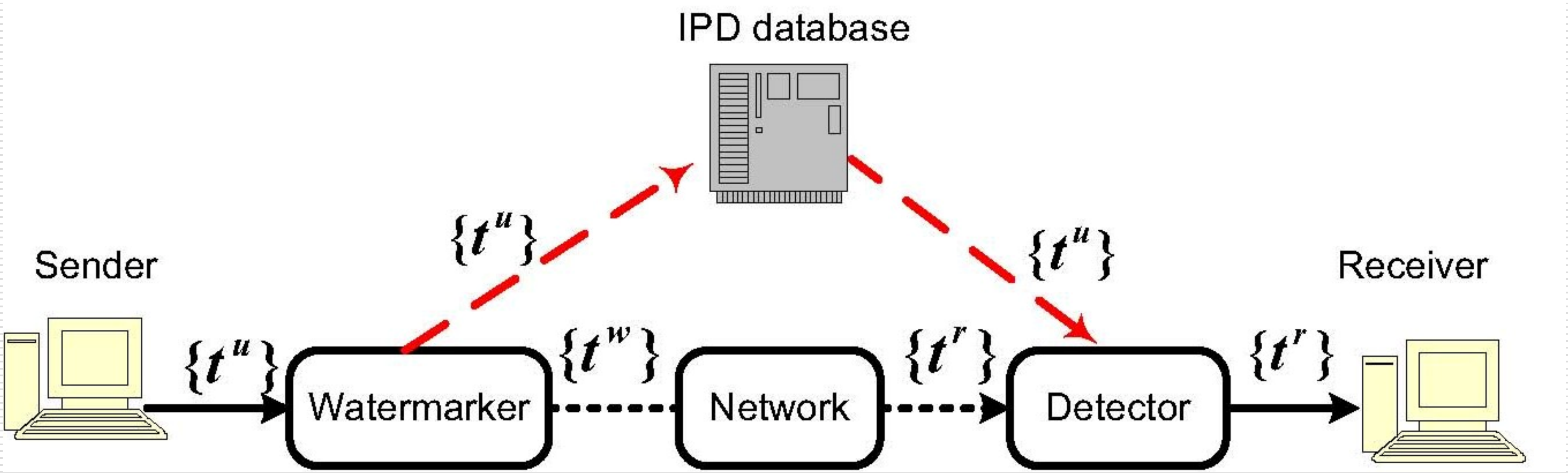
Watermark Insertion

- Uses Inter-Packet Delay (IPD) information for watermarking



- Based on spread spectrum multimedia watermarking

Insertion scheme



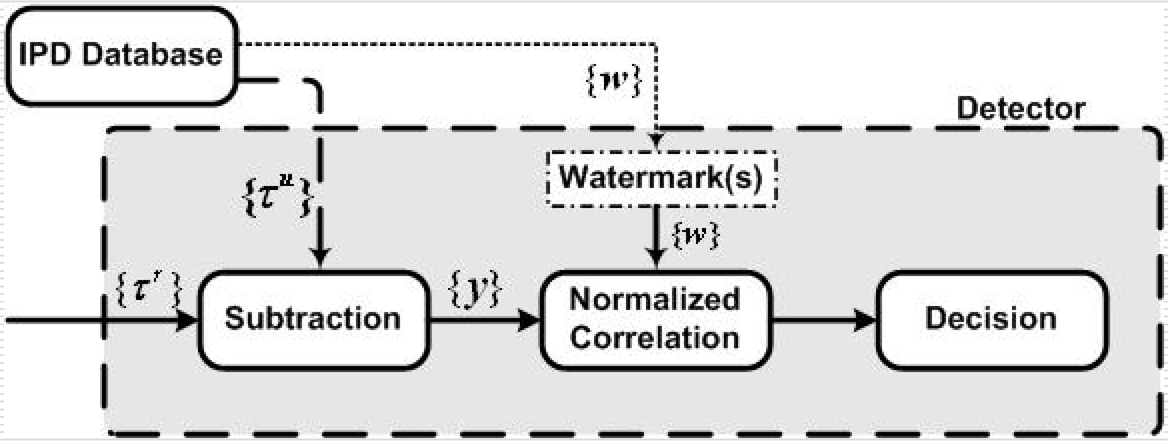
- $Post\_IPD(t^w) = Pre\_IPD(t^u) + Wm$
- $Recv\_IPD(t^r) - Pre\_IPD(t^u) = Wm + Jitter$

IPD database

- For new flows, watermarker creates an entry in database
  - Last N packets
  - Update during time
- Entry is removed from database, after connection ends
- Resources
  - Memory: 3.1 MB for an institution with 400 members

Detection scheme

- Use last N samples of received flow
- Recv_IPD – Pre_IPD = Wm + Net_Jitter
  - Detection of spread spectrum signal
- Network jitter model: Laplacian $Lap(0,b_\delta)$
  - Normalized Correlation is an efficient detection rule
- Decision based on threshold



System analysis

- Model system
  - Jitter $\delta \propto Lap(0, b_\delta)$
  - IPDs: exponential
- SNR $\gamma = \frac{a}{\sqrt{2}b_\delta}$
  - a : watermark amplitude
- Hypothesis testing
  - True detection $T_1 \propto Lap(\gamma, \frac{1}{\sqrt{2N}})$
  - False detection $T_0 \propto Lap(0, \frac{1}{\sqrt{2N}})$

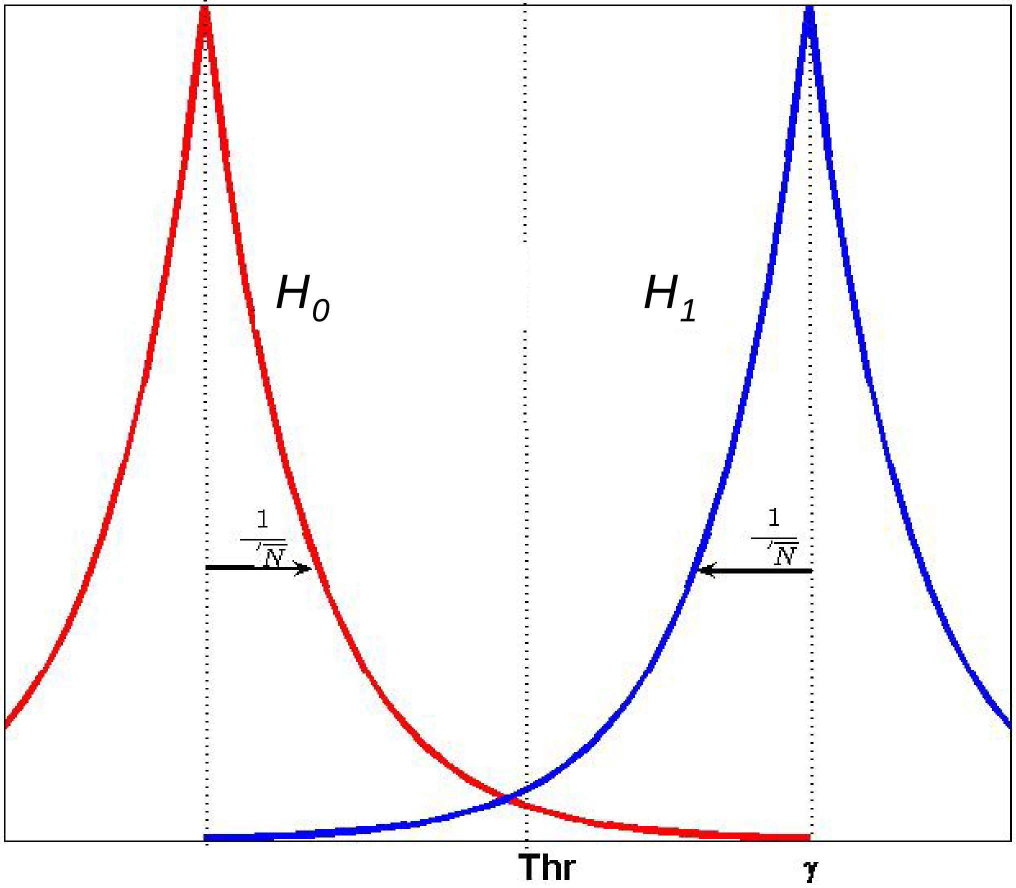
System analysis

- Detection threshold η

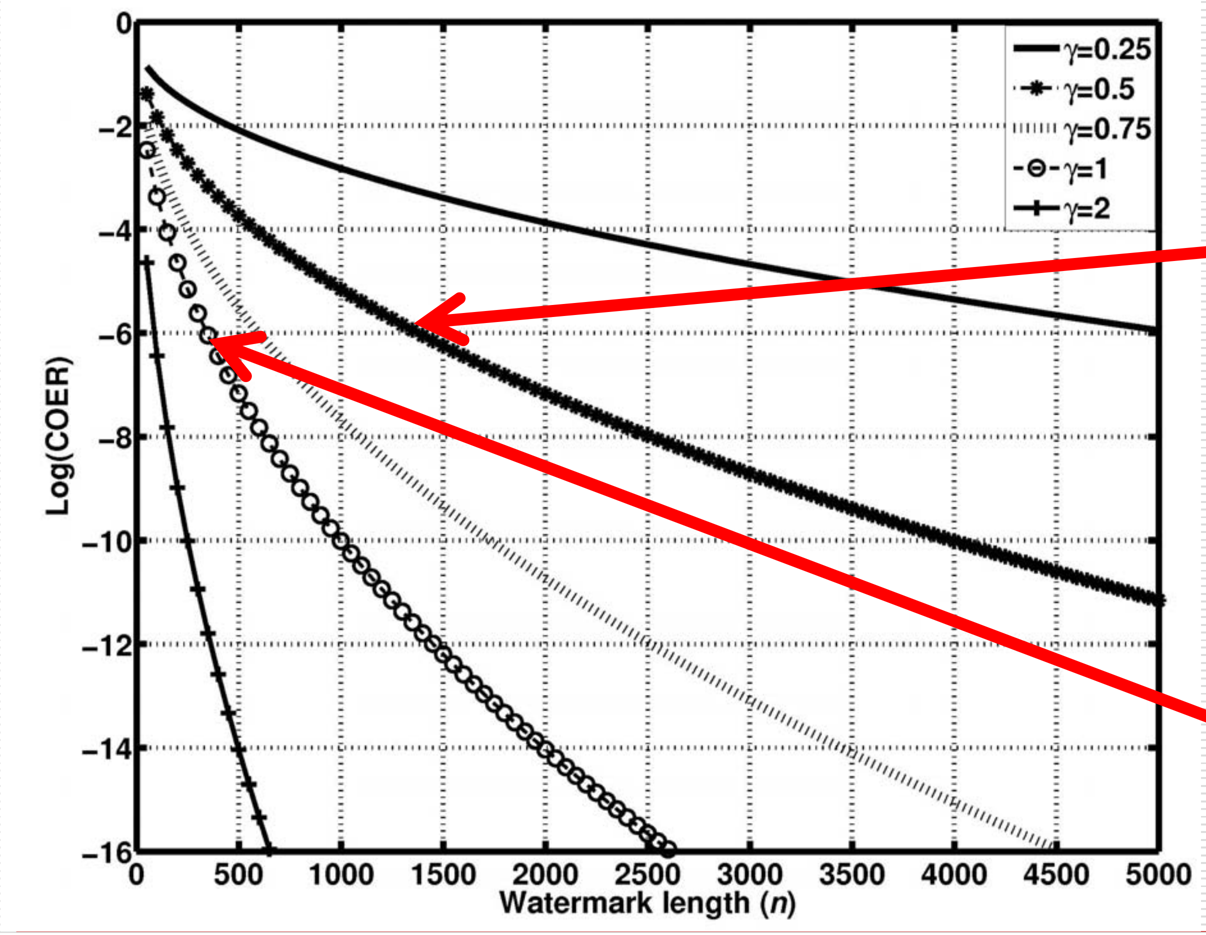
$$FP = \frac{1}{2} e^{-\eta\sqrt{2n}}$$

$$FN = \frac{1}{2} e^{-(\gamma-\eta)\sqrt{2n}}$$

- MinMax rule
- COER
- Neyman-Pearson



MinMax analysis



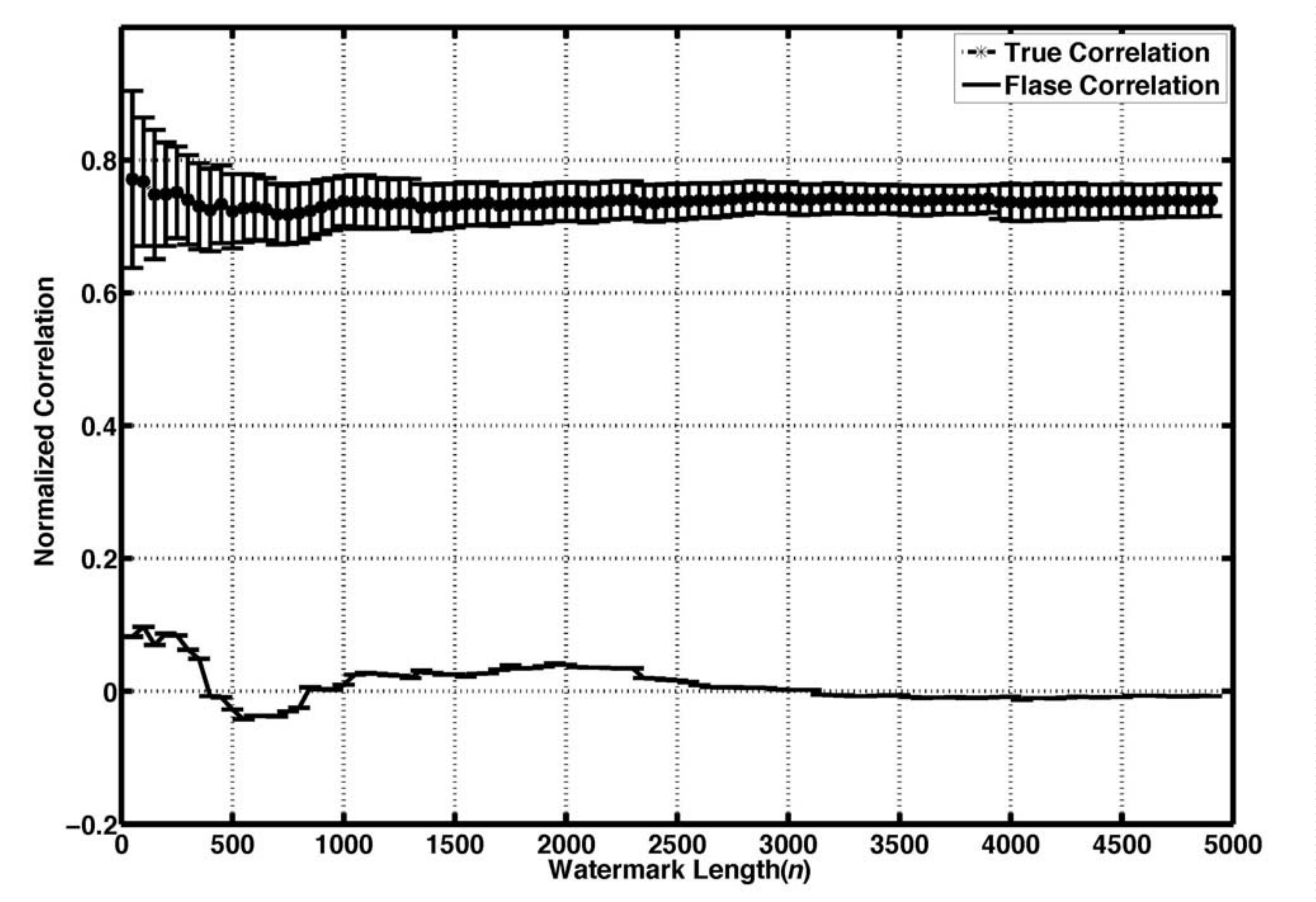
a= 5ms
n=1300
FN=10^{-6}
FP=10^{-6}

a= 10ms
n=400
FN=10^{-6}
FP=10^{-6}

Implementations

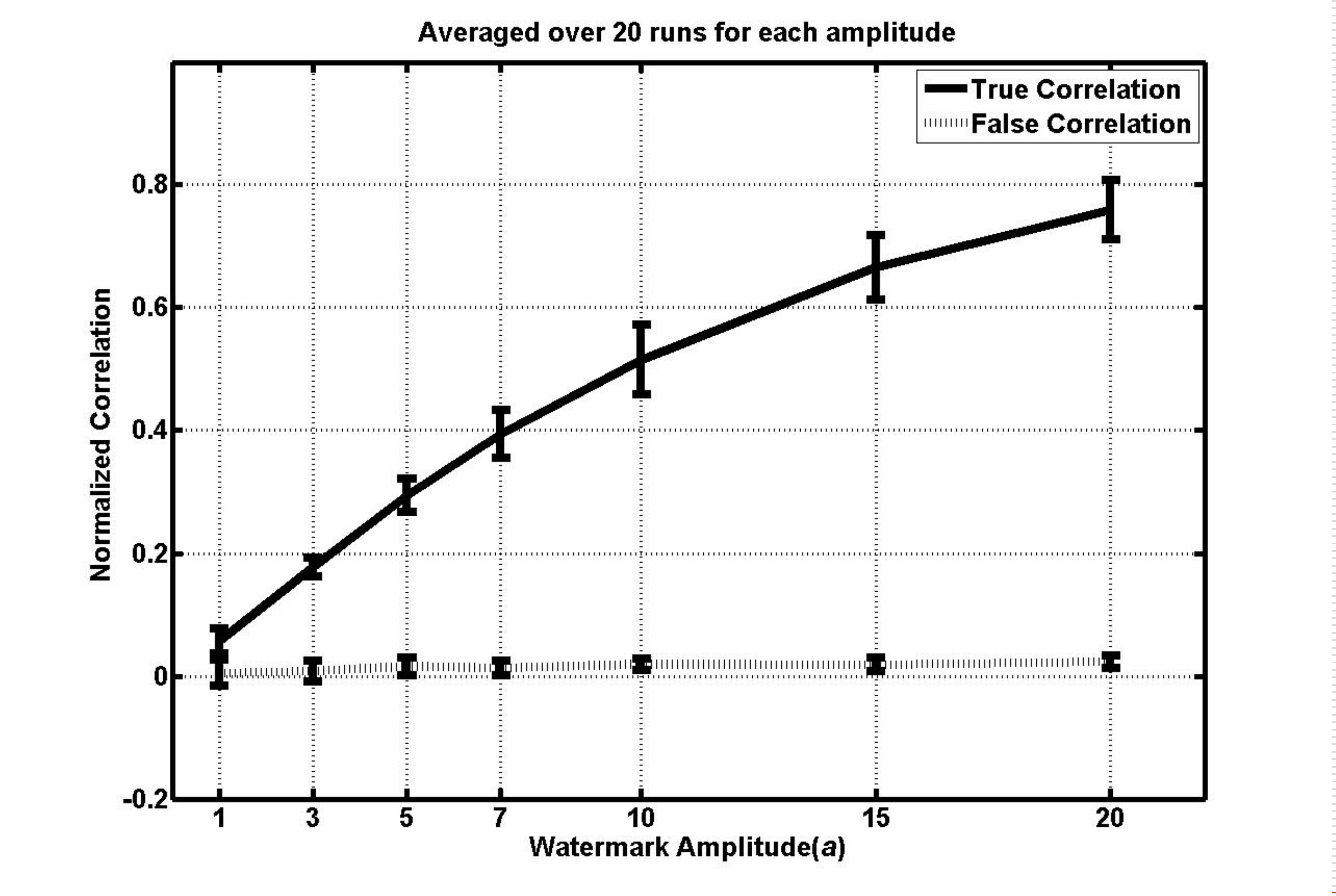
- PlanetLab infrastructure
  - Larger jitter than normal traffic
- SSH traffic

Implementation results



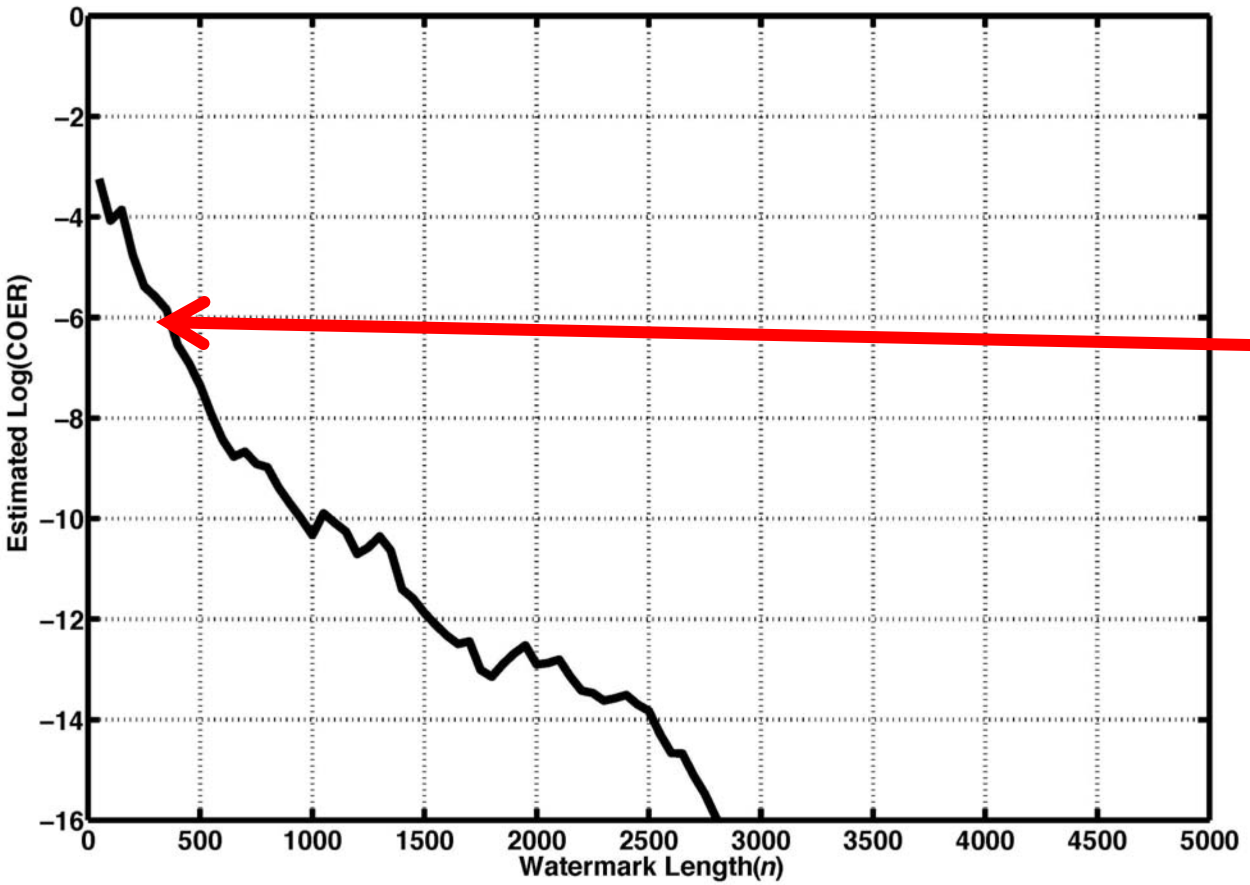
a=10 ms
100 flows

Implementation results



n=500
jitter=10ms

Practical COER



$\gamma = 1$

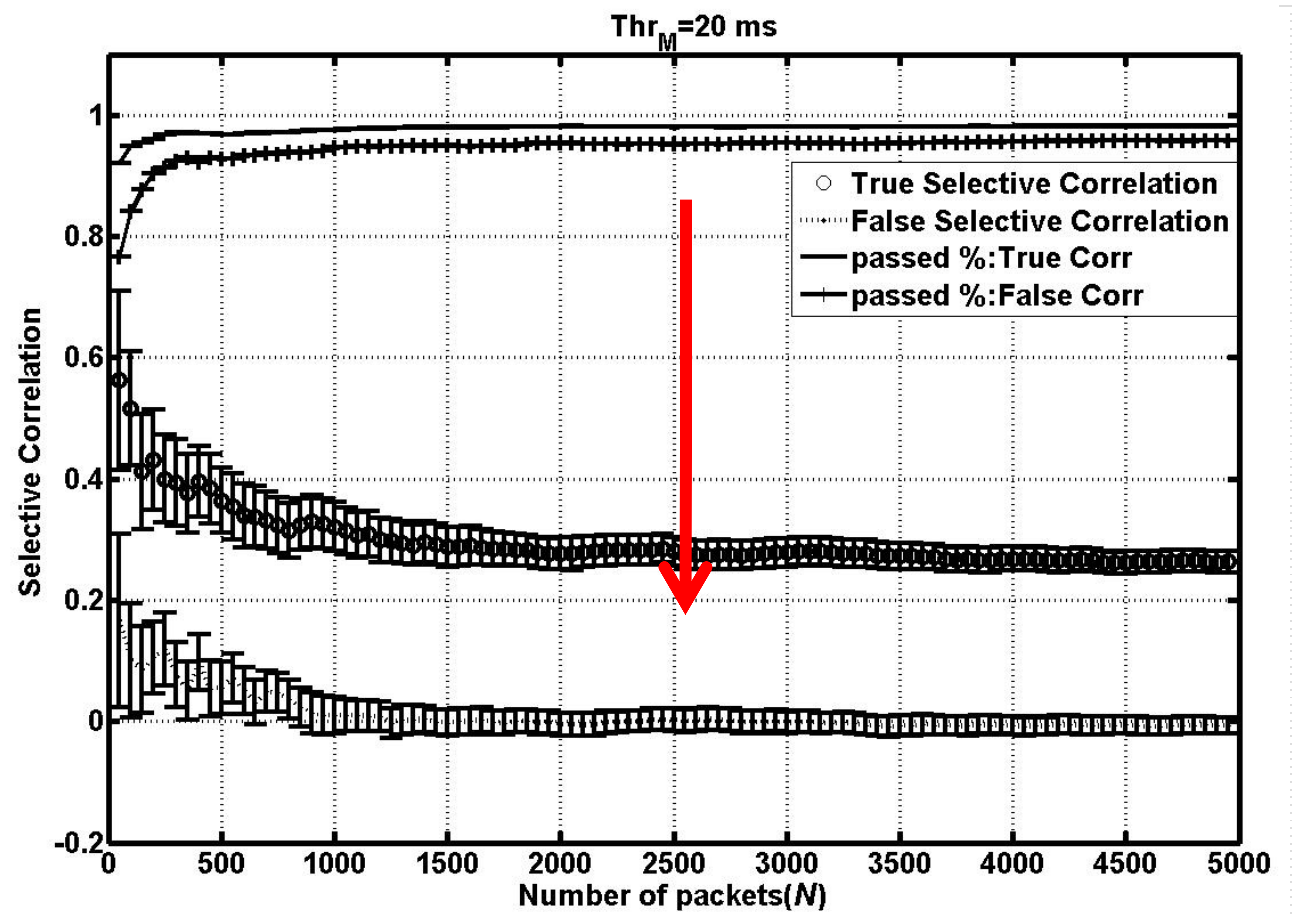
a = 10ms
n = 400
COER = 10^{-6}

Selective correlation

- Sources of flow modification
  - Protocol specific causes: duplicated, retransmitted, re-packetized, ...
  - Protocol specific packets: TCP ACK/SYN, SSH initial packets, ...
  - Initial delay
- Matching block
  - Sliding windows

Implementation

$r = 20\%$



Invisibility

- Using Non-blind spread spectrum watermark we expect high invisibility
- Confirmed through information-theoretic tools:
  - Kolmogorov-Smirnov test
    - 98% confidence
  - Entropy-based tools of Giavencchio for covert channels (CCS'07)

Performance comparison

- Run time: 0.4 microsec for 400 connections with 5000 packets
- Detection time: about 3 min (400 packets)
- False errors of order 10^{-6}
  - Passive schemes: 10^{-2}
  - Blind watermarks: at most 10^{-5}
- Invisibility

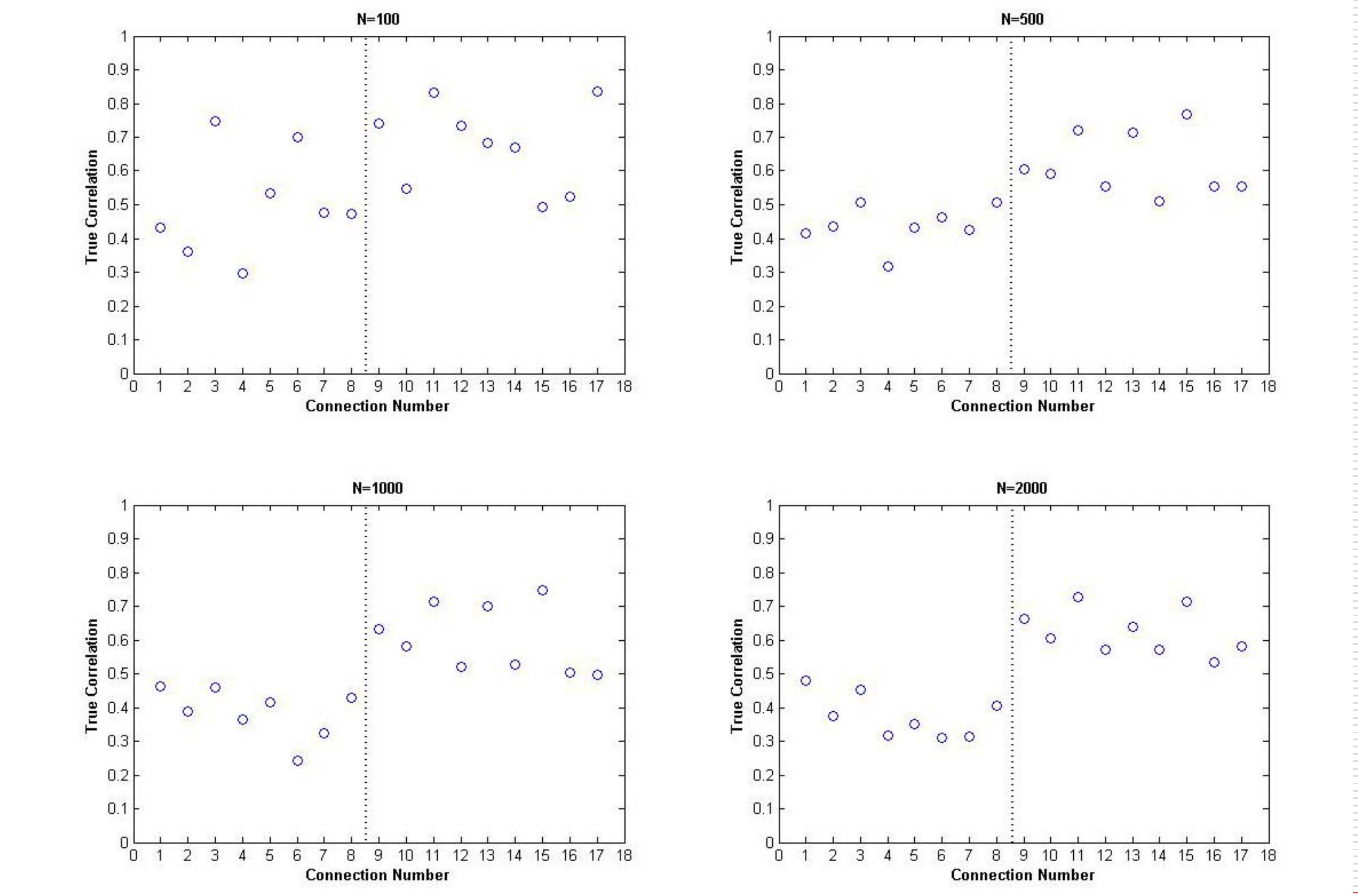
Conclusions

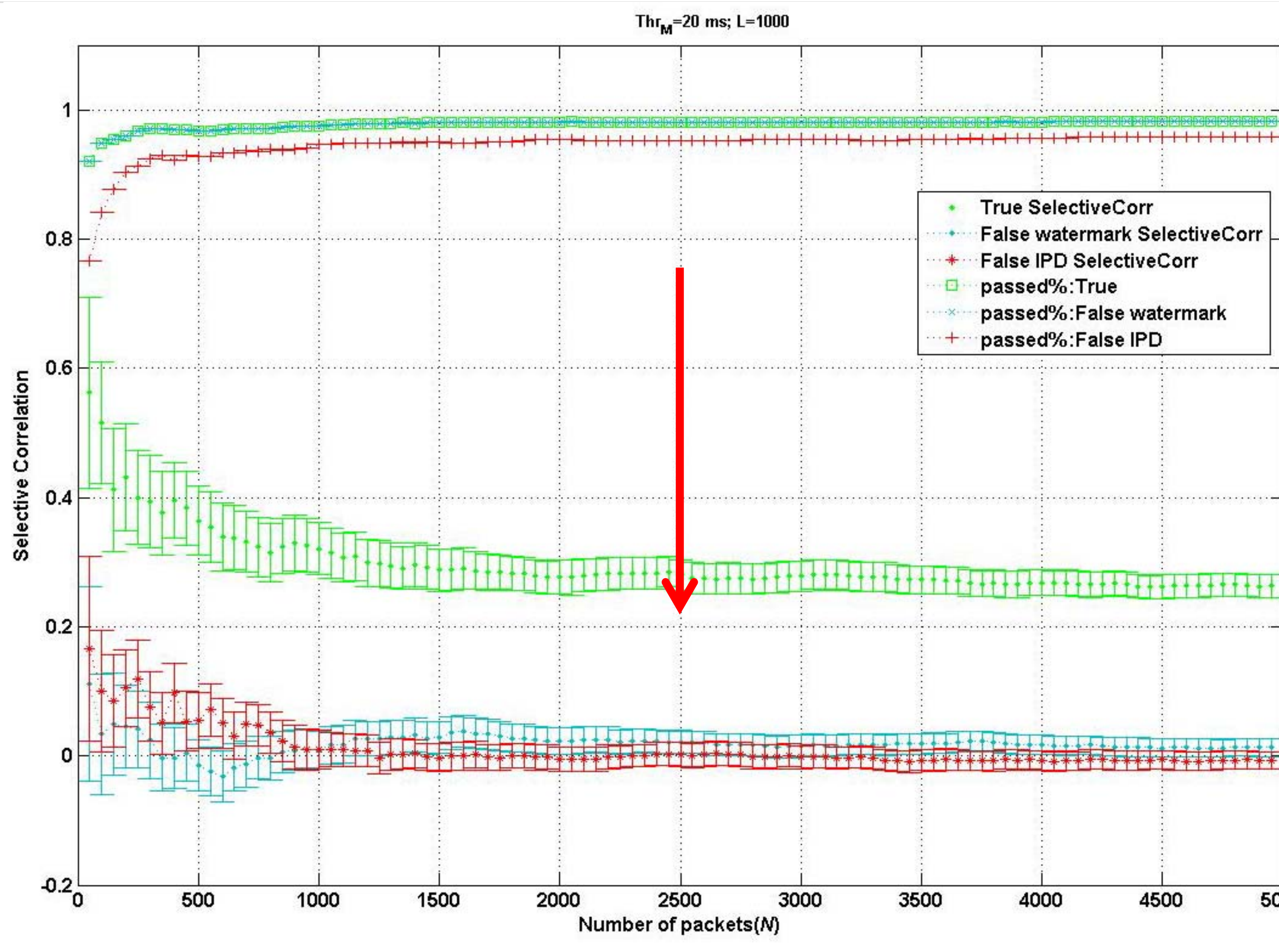
- RAINBOW: A novel traffic analysis
  - In between of passive and blind active
- High Detection Efficiency
- Invisibility
- Robustness to flow modifications

- Future work: Use fast coding tools to insert watermarks more efficiently
  - Effective semi-blind or blind schemes

Thanks

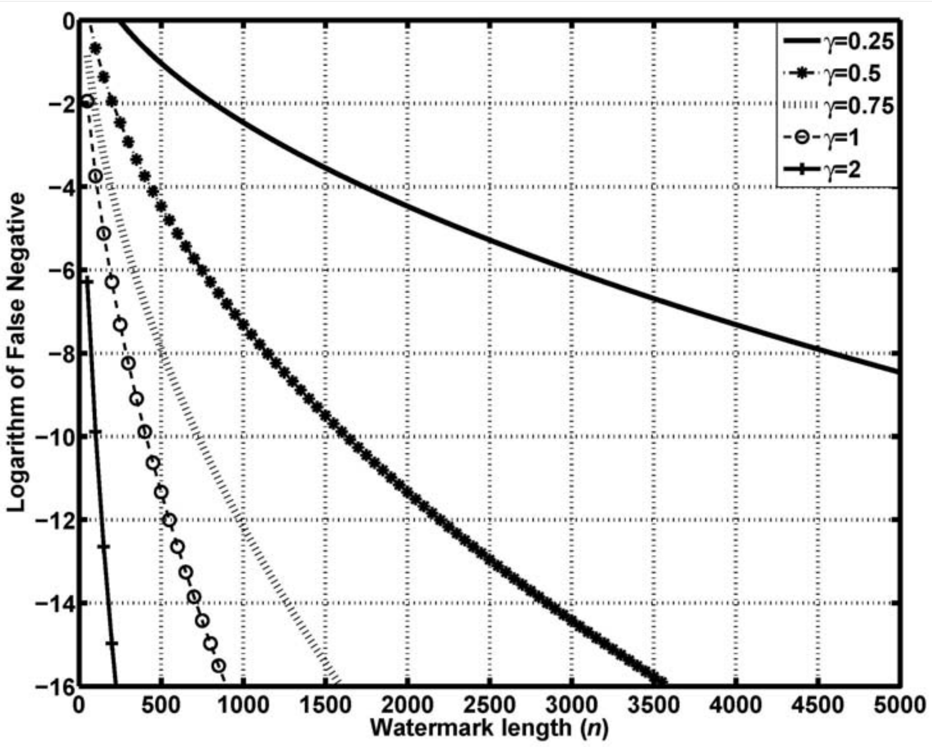
Implementation results



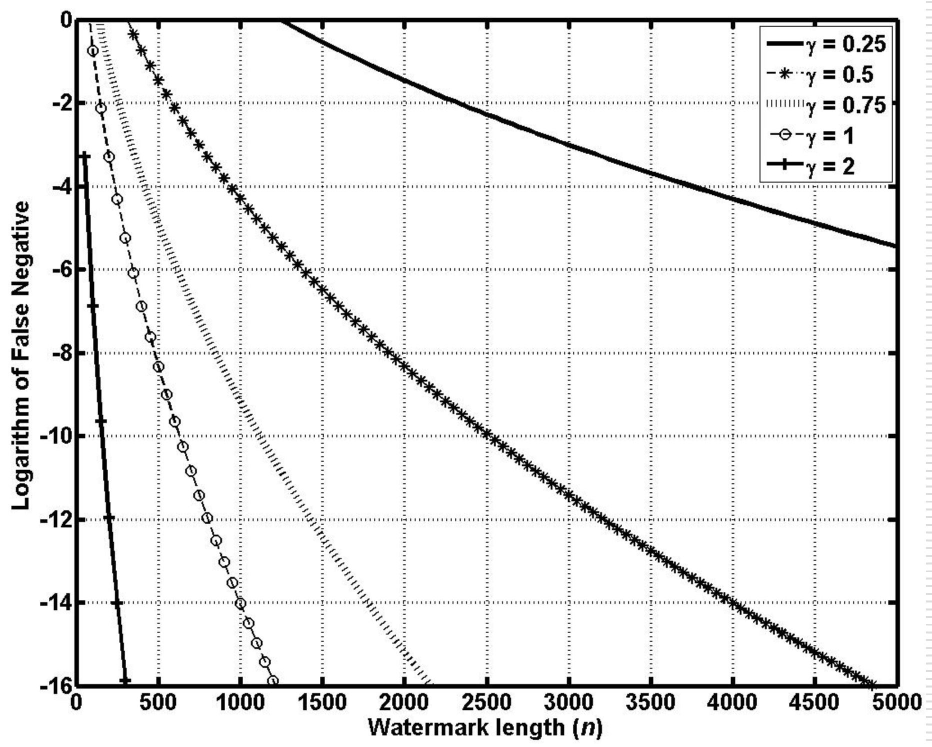


Thr_M=20 ms; L=1000
Selective Correlation
Number of packets(N)
True SelectiveCorr
False watermark SelectiveCorr
False IPD SelectiveCorr
passed%:True
passed%:False watermark
passed%:False IPD

Neyman-Pearson analysis



FP=10^{-3}



FP=10^{-6}