The FootFall Project

"Tracing Attacks Through Non-Cooperative Networks and Stepping Stones with Timing-Based Watermarking"

Douglas Reeves & Peng Ning

Cyber Defense Laboratory
N.C. State University

IAIC Program Kickoff Meeting
November 17, 2003

The Name

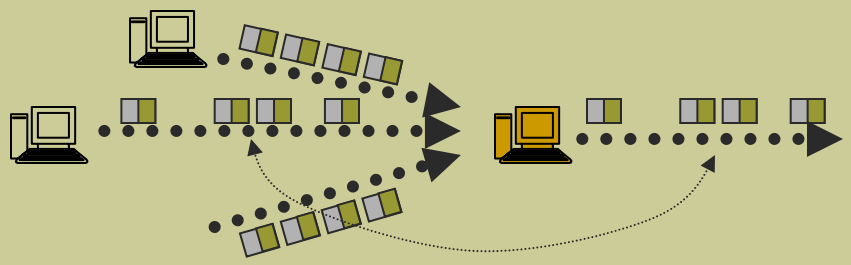
- "FootFall" = sound of a footstep
- We propose to record and analyze packet timing (traffic "footsteps") for tracing purposes

FootFall ≠ FootBall

"Tracing Attacks Through Non-Cooperative Networks and Stepping Stones with Timing-Based Watermarking"

Douglas S. Reeves and Peng Ning

Basic Idea



Correlation of packet timing across stepping stones allows traffic to be traced almost unnoticeably

Features

- Uses packet timing analysis for attack attribution
- Actively watermarks traffic to assist tracing
- Can be made arbitrarily robust by increasing the level of redundancy
- Difficult for attacker to detect; uses random packet selection, delays

Objectives

- Identify source of attacks
- Overcome encryption, stepping stones, timing perturbation, other anonymizing techniques
- Work even across non-cooperating networks
- Be robust against active efforts to evade

Schedule

Quarter										Task
1	2	3	4	5	6	7	8	9	10	
*		Implement off-line watermarking method, deliver and demonstrate								
	*		Implement on-line (real-time) kernel-level watermarking method, deliver and demo							
		*	*		Create solutions to enhanced anonymity techniques, report					
				*		Implement and test solutions to enhanced techniques, deliver and demo				
					*		Create solutions to system implementation issues, report			
						*		Implement solutions to system implementation issues, demo and deliver		
							*	*		Investigate solutions to advanced techniques, report
									*	Technology transfer, acceptance testing and training

ARDA

NC STATE UNIVERSITY

NEW!

Website

footfall.csc.ncsu.edu

ARDA

overview….

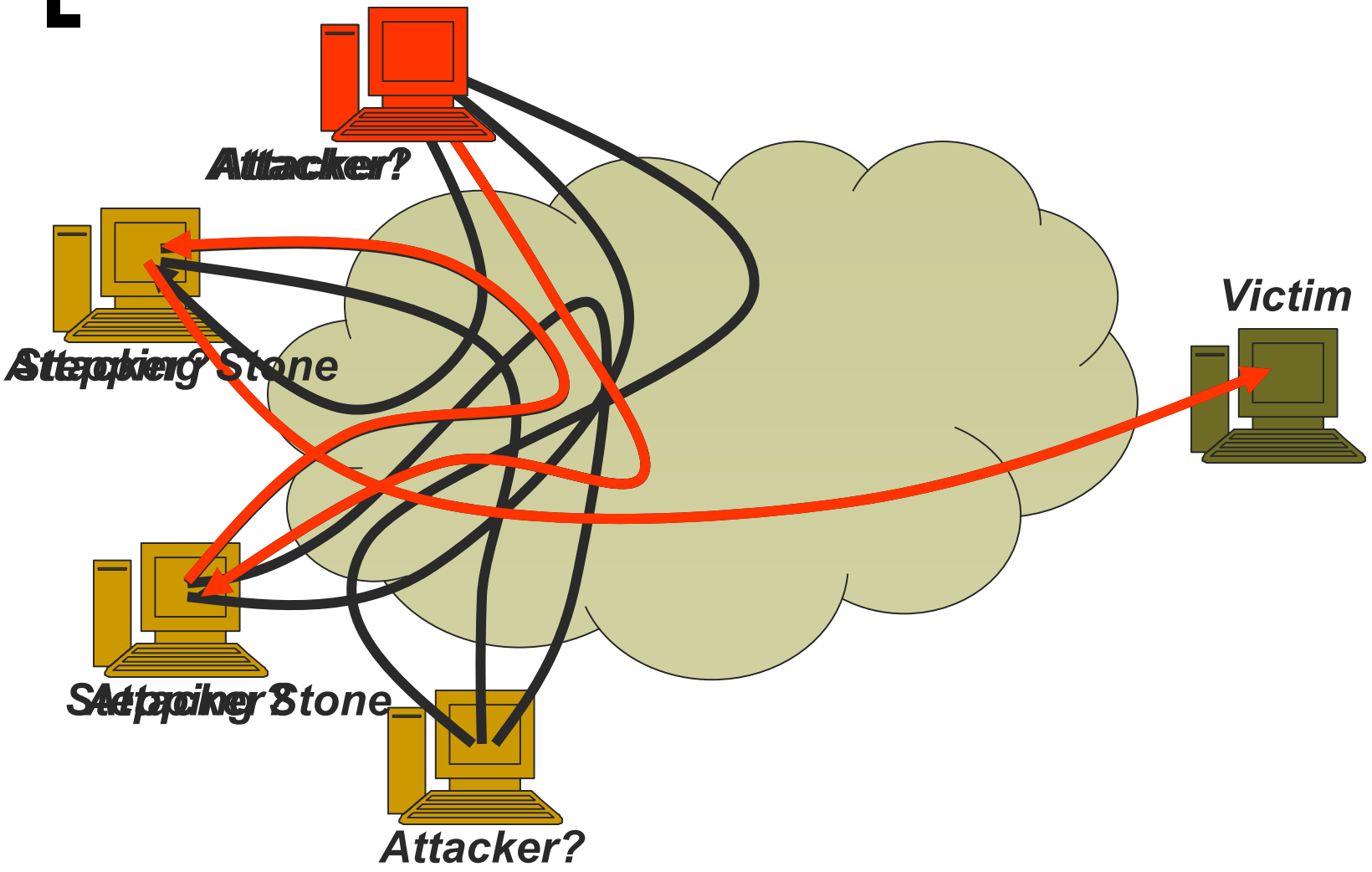
Anonymity and Attack Attribution

- Attackers have been successful at hiding their identity
  - greatly increases the likelihood of an attack
- Anonymizing techniques can be used…
  - benignly (preserve privacy of the individual)
  - maliciously (conceal identity of terrorists)

Anonymizing Techniques

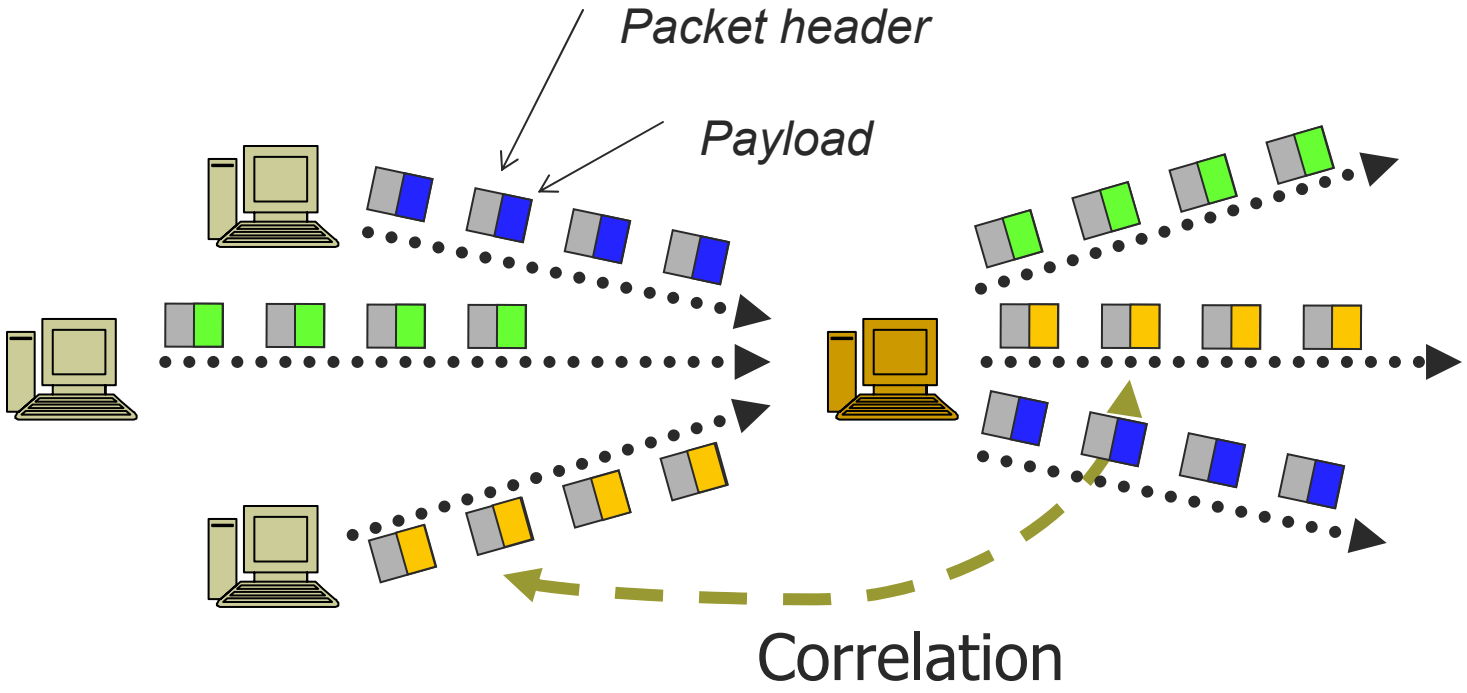
1. IP Address spoofing
2. Disguising of data (steganography)
3. Encryption (packet header and payload)
4. Use of surrogates or intermediaries (proxies, stepping stones, zombies)
5. Mixing with other traffic (camoflaging, mixing)
6. Randomized behavior (onion routing)

The Stepping Stone Problem



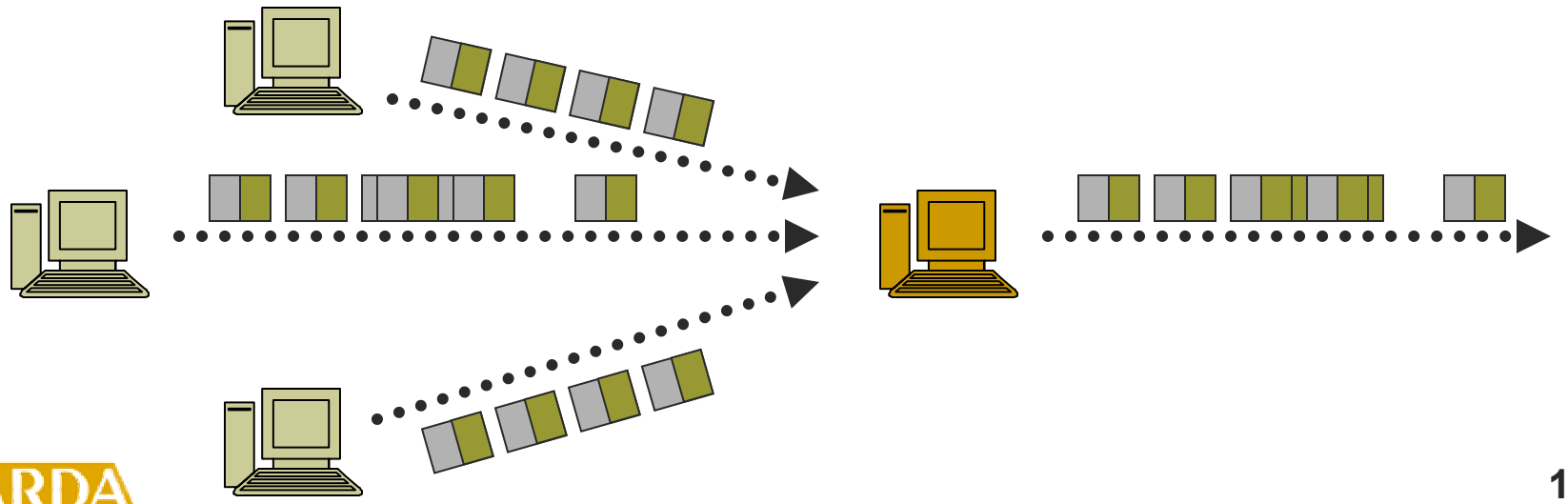
Correlation of flows

- Given an outgoing flow from a stepping stone, correlate it with an incoming flow to the stepping stone



Novel Ideas / Broad overview

1. Use packet timing for correlation
2. Watermark the timing to facilitate correlation
3. Use redundancy to make robust against attacks



technical approach….

Timing Based Correlation

- Timing based approach works even with…
  - encrypted connections
  - padding of the packet payload
- But, vulnerable to timing perturbations by the attacker
  - make unrelated flows look similar (increase false positive rate)
  - make related flows look dissimilar (decrease true positive rate)

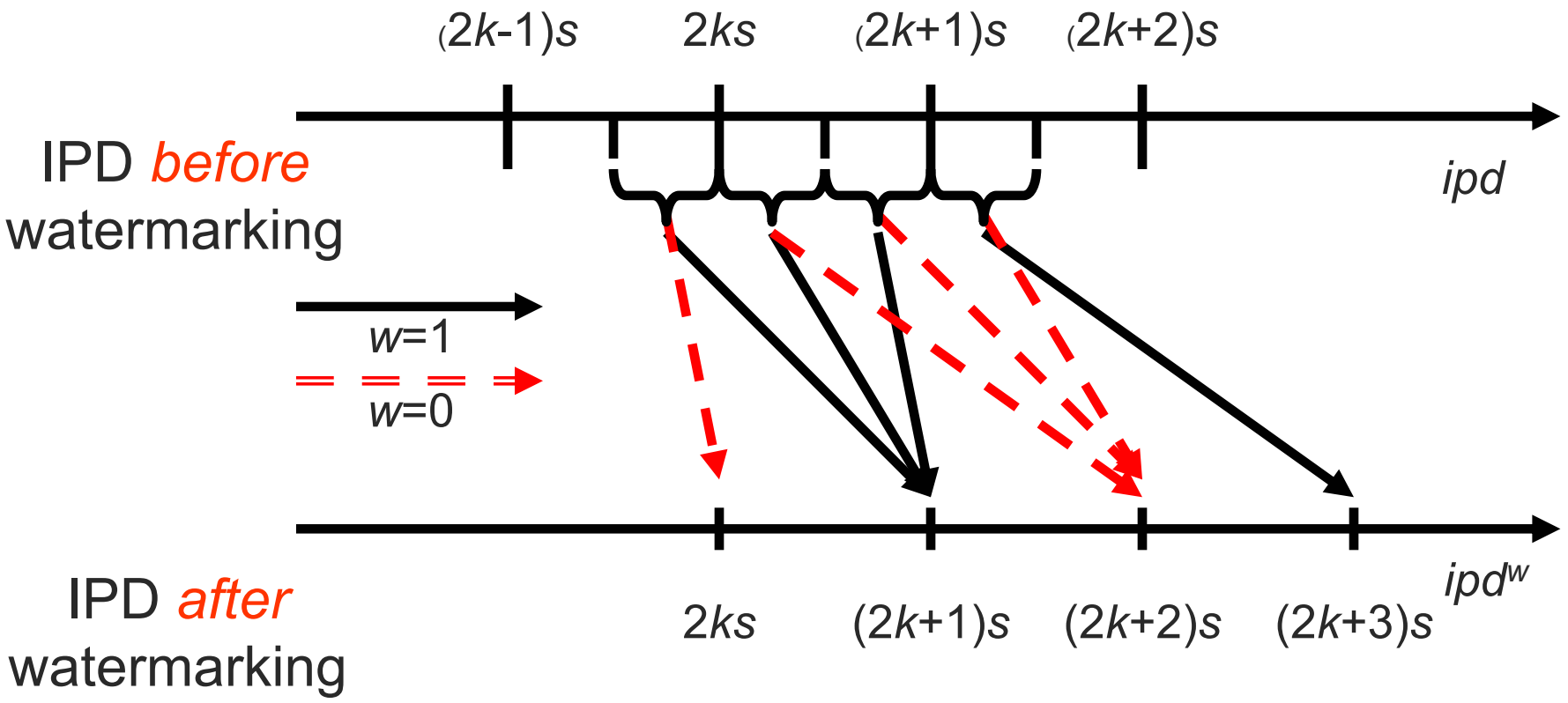
Limits of Timing Perturbation

- Donoho *et al.* investigated limits of timing perturbation
  - correlation possible for sufficiently long flows if timing perturbation has pareto distribution
- Fundamental questions
  - is correlation effective for flows if timing perturbation has arbitrary distribution?
  - what is the achievable tradeoff among
    (i) correlation true positive rate
    (ii) correlation false positive rate
    (iii) maximum amount of timing perturbation?

Correlation using Watermarking

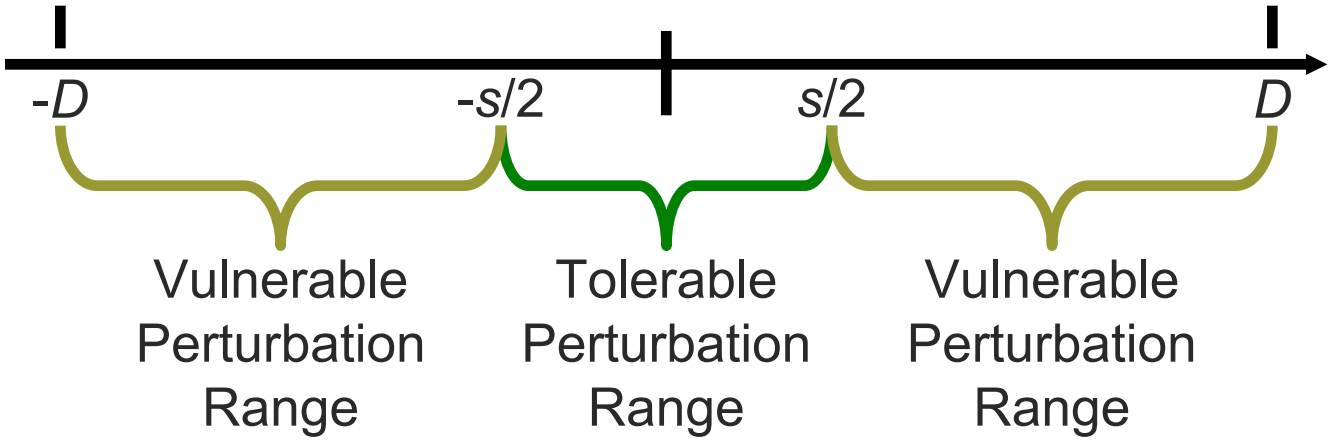
- Actively embed a unique watermark into the flow
  - by slightly adjusting the timing of selected packets
- Effective correlation is achieved if the embedded watermark is…
  - unique enough
  - robust enough

Embedding A Single-Bit Watermark



Robustness of the Watermark Bit

- Tolerable perturbation range
  - embedded watermark bit is guaranteed to be recoverable
- Vulnerable perturbation range
  - embedded watermark bit may not be recoverable

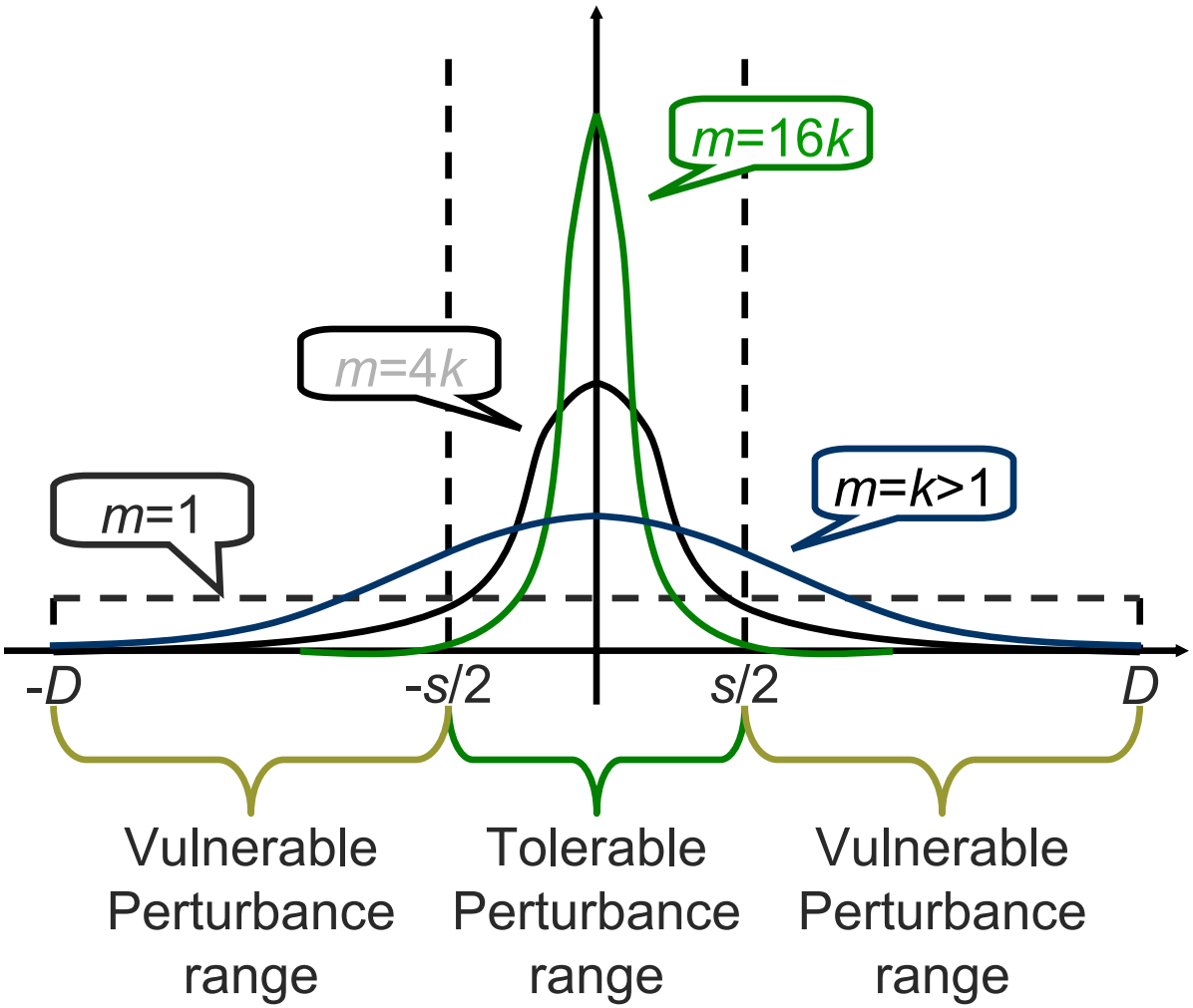


Probabilistically Robust Watermarks

- If the timing perturbation is outside the tolerable perturbation range, watermarking may fail
- How design the watermark so that the probability of this is small?
  - spread the watermark over a longer duration of the flow
  - embed the watermark bit over the average of *m* (multiple) IPDs

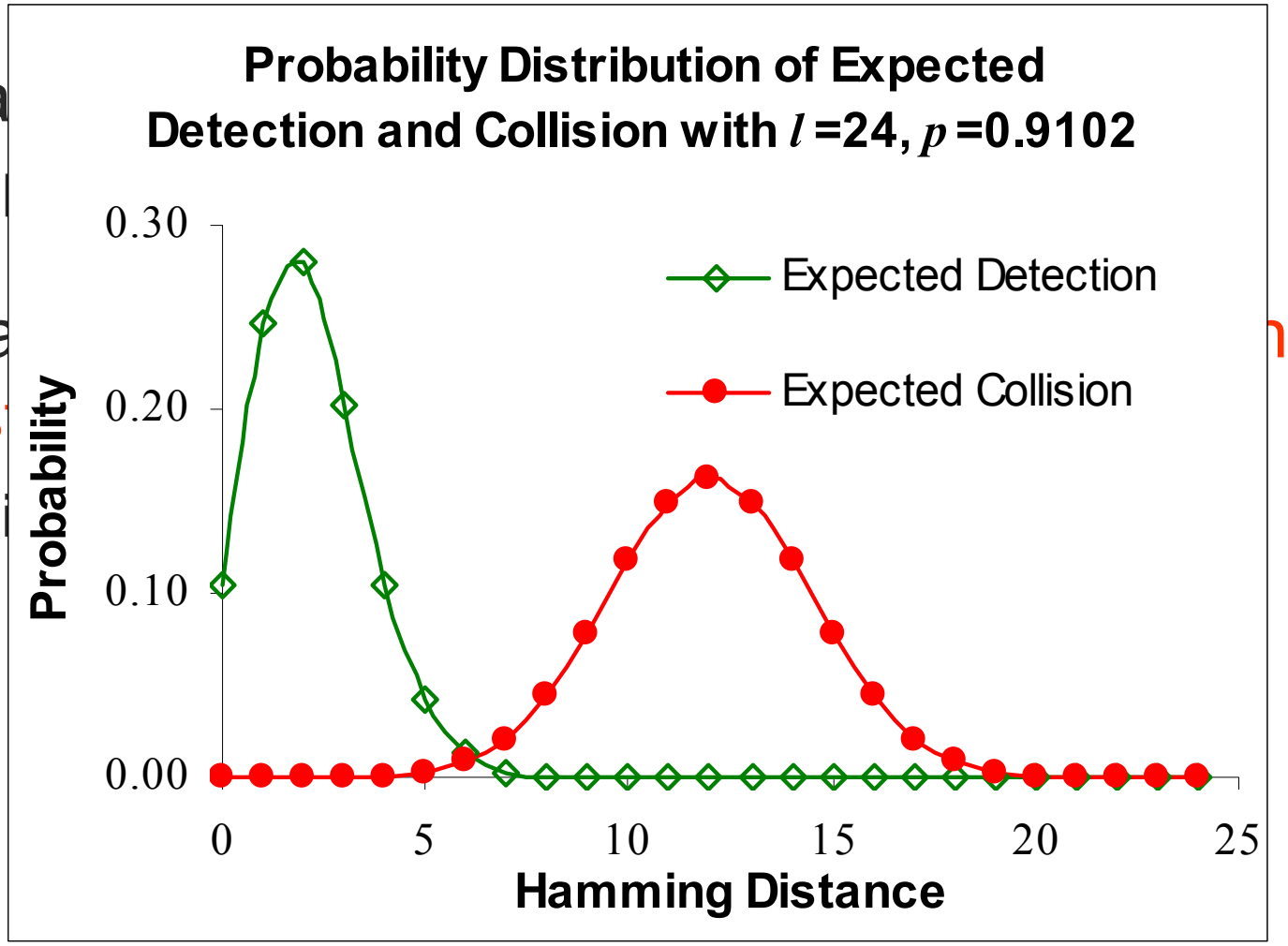
Reducing the Impact of Random Timing Perturbations

Effect of the Central Limit Theorem of Statistics



Collision and Detection Rates

- Ma
  - I
- Wa ng
  dis
  - i

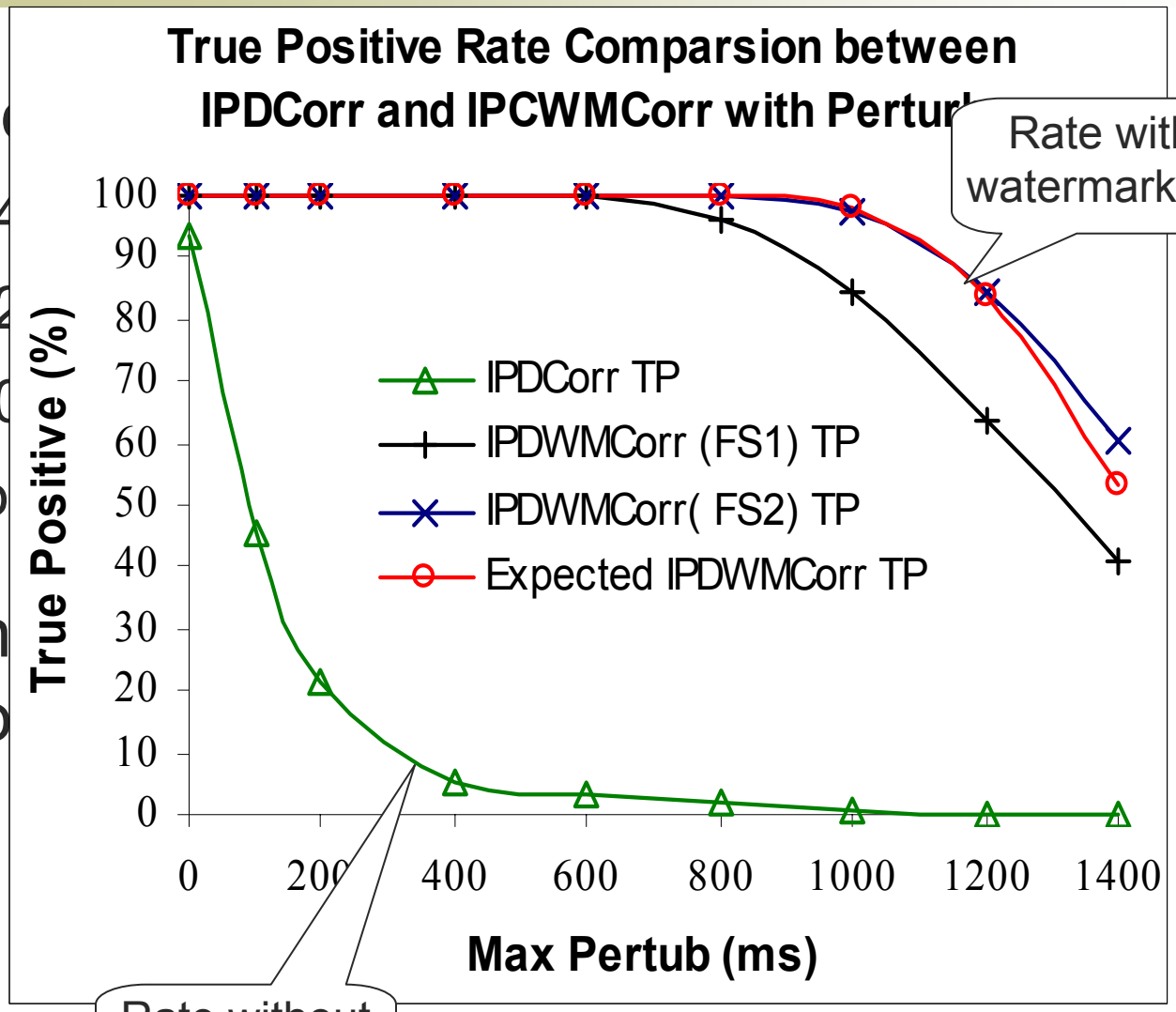


Results for IID Random Perturbations

- It is possible to achieve at the same time (!)
  (i) arbitrarily high watermark detection rate
  (ii) arbitrarily low watermark collision
  with
  - arbitrarily small average timing adjustment
  - arbitrarily large (but bounded) *iid* random timing perturbation of
  - arbitrary distribution
- As long as there are enough packets

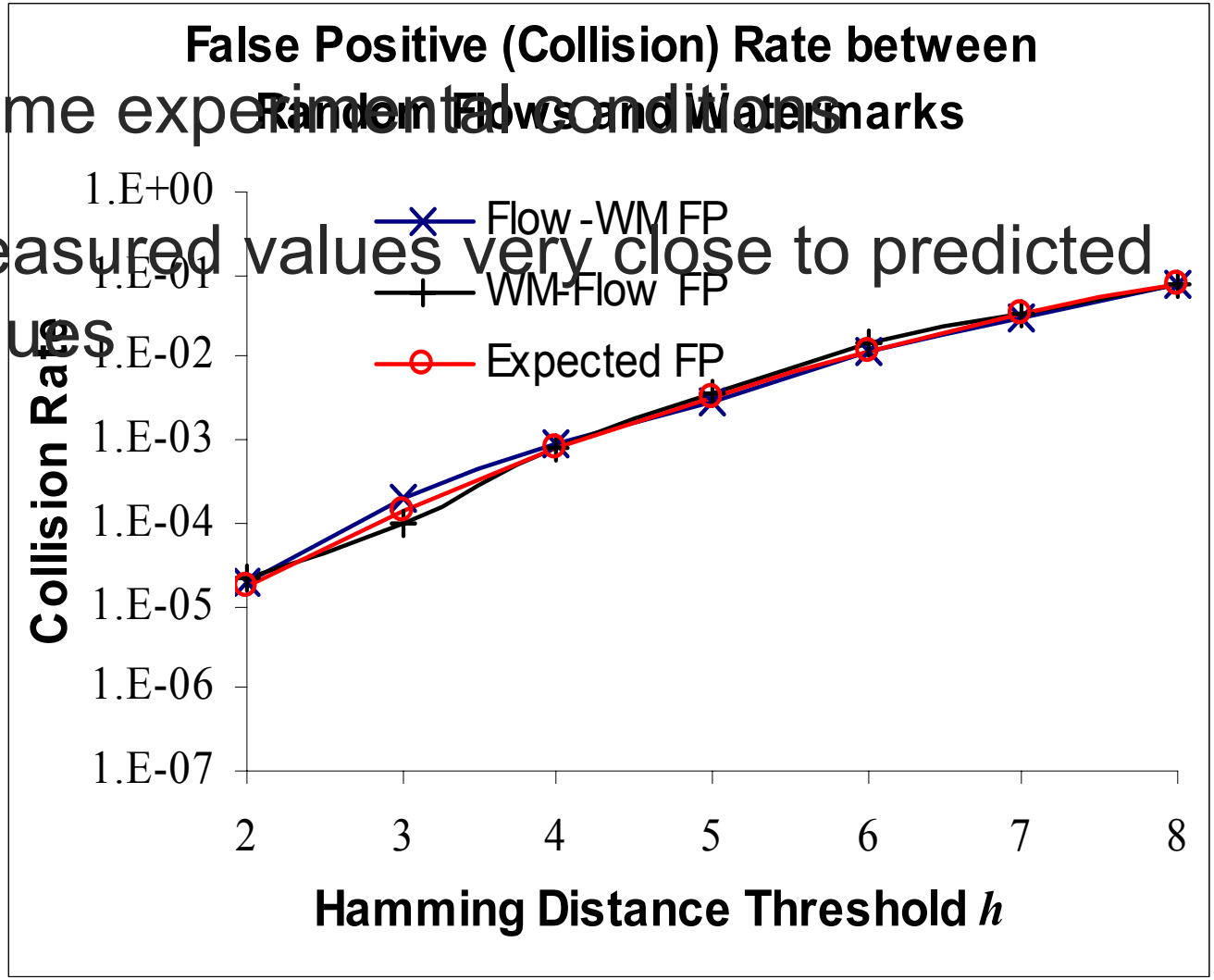
Experiment: True Positive Rate

- Con
  - 24
  - 12
  - 40
  - to
- Usin
  repo

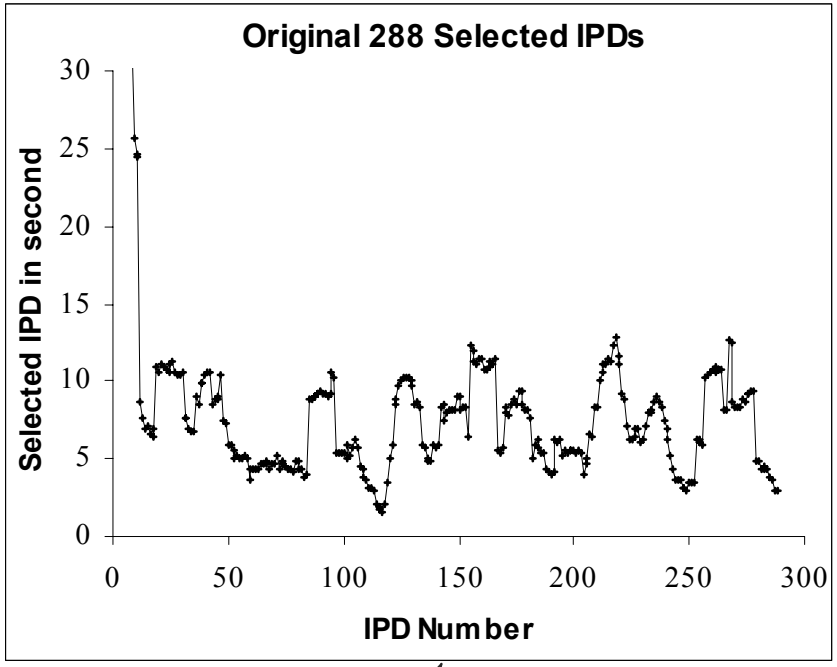


Experiment: False Positive Rate

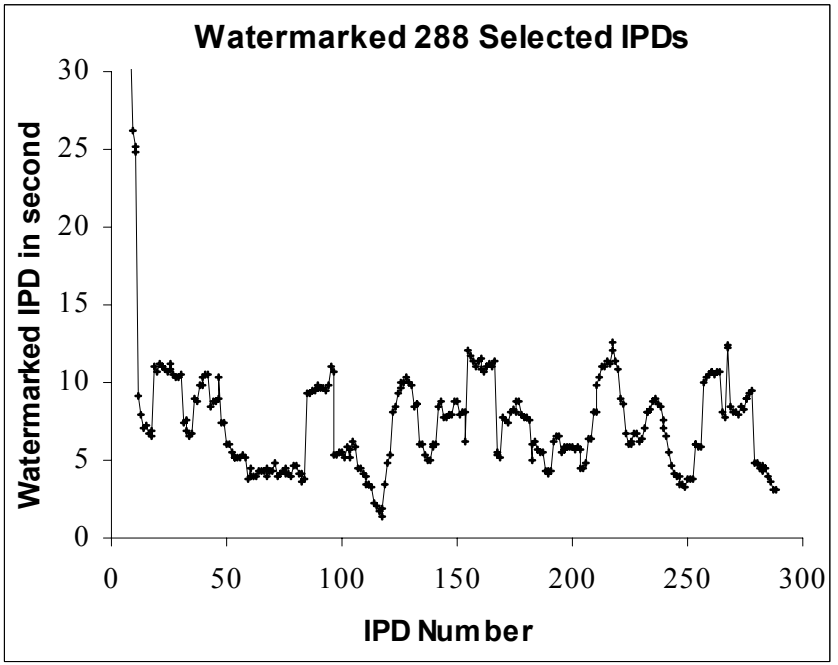
- Same experimental conditions
- Measured values very close to predicted values



Watermarking Obvious to Attacker?



IPDs **Before** Watermarking



IPDs **After** Watermarking

assessment….

Previous Work

- Correlation based on users' login activity
  - caller ID
  - DIDS
- Correlation based on packet contents (payload)
  - thumbprinting
  - SWT (Sleepy Watermark Tracing) ← our work
- Correlation based on inter-packet delays (IPDs)
  - on/off based
  - deviation based
  - IPD based ← our work
  - wavelet based

ARDA

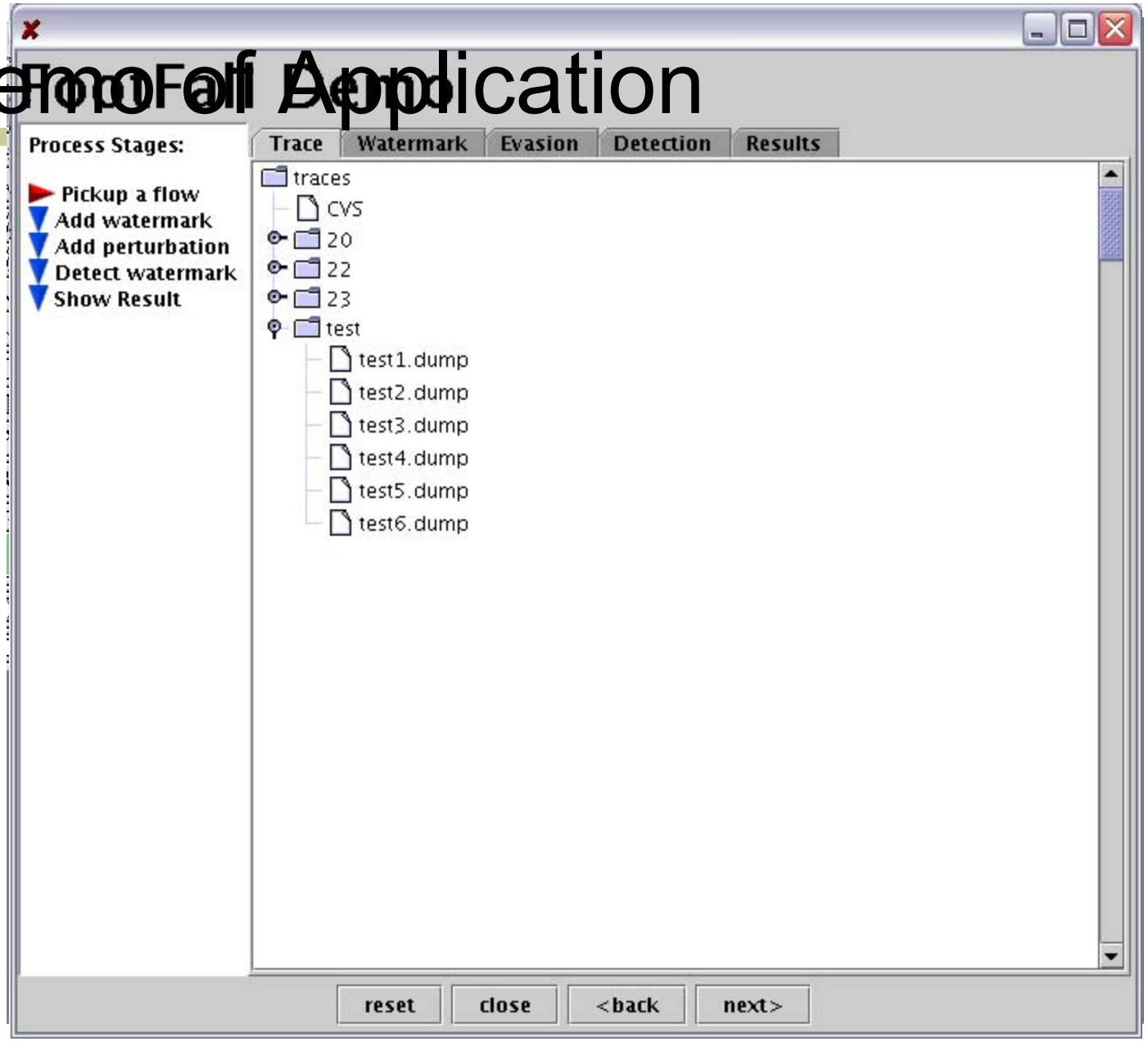
NEW!

Software Implementation

- We have implemented watermarking as an application program
- Watermarks recorded traces of network traffic
  - NLANR repository of traffic traces
- Calculates
  - false and true positives with watermarking
  - false and true positives without watermarking
- Convenient GUI

Demo of Application

NEW!



NEW!

Real-Time Execution

- Acquiring and setting up testbed
- Evaluated platforms, selected Linux
- Experiments measuring timing accuracy
- Research into timing system – interrupts, timers, system calls, etc.
- In progress: implementing watermarking (and analysis) at the kernel level

Teaming

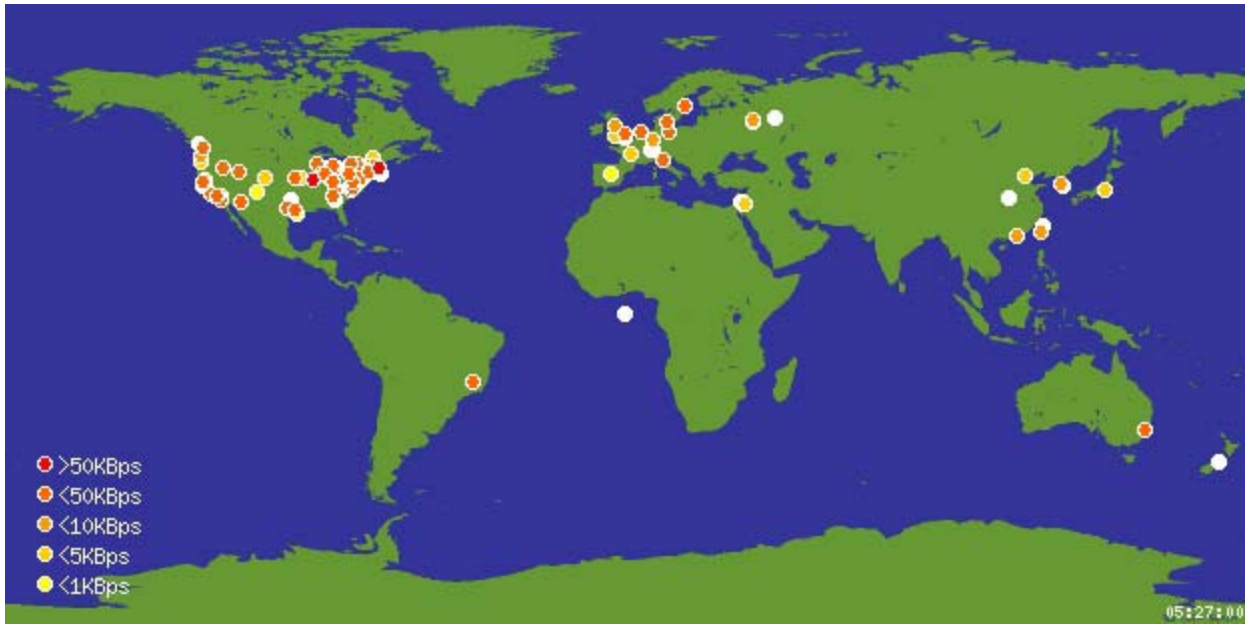
- All NC State University
  - Faculty (Reeves, Ning)
  - Consultant (Xinyuan Wang)
  - 4 Ph.D. students

Facilities

NEW! Internal testbed

- Campus network
- Planet Lab

NEW!



ARDA

challenges….

Harder Problems

1. Basic anonymizing techniques
   - ✓ spoofing IP source addresses
   - ✓ encryption of packet payload
   - ✓ use of stepping stones (intermediate hosts)
   - ✓ perturbation of packet timing characteristics
2. Enhanced techniques
   - dropping, retransmitting, or reordering packets in a flow
   - re-packetization of data in a flow
   - tunneling (VPNs, IP within IP, encryption of header and payload)

(cont'd)

…

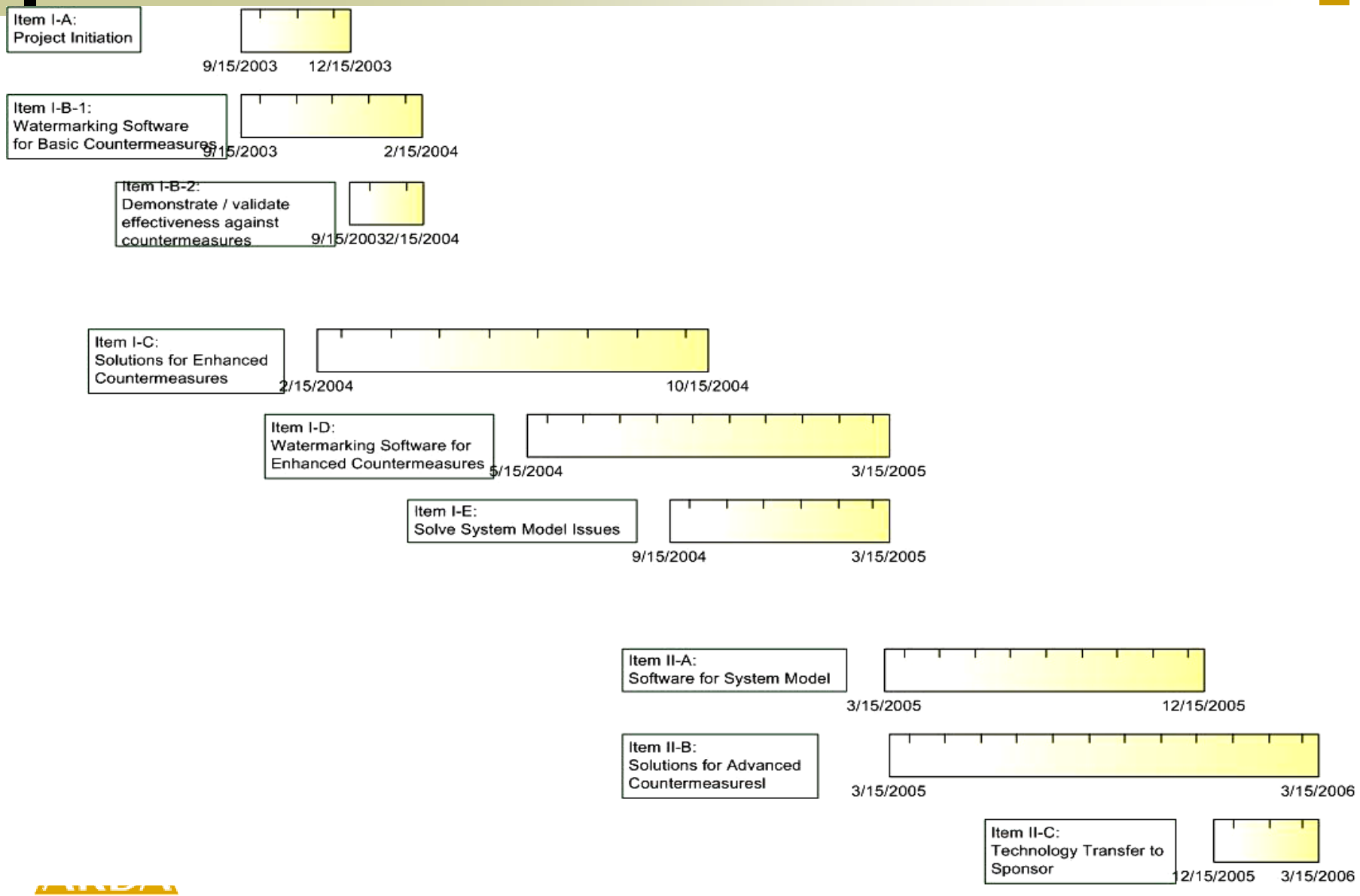
- mounting slow, long-timescale attacks
- splitting one flow into multiple flows, and merging them back together at another point in the network

3. Advanced techniques
   - addition of padding traffic (chaff, camouflaging)
   - use of mixers or anonymizing proxies
   - onion routing

High-Risk / High Payoff?

- The bad news
  - the attackers have a lot of traffic to hide in
  - they can exploit sophisticated anonymizing techniques
  - only some network providers will assist with tracing
- The good news
  - we have discovered and demonstrated a powerful, robust traffic watermarking technique
  - the technique is lightweight and cheap to implement
  - many anonymizing techniques make traffic *easier* to trace
- What are the basic limits of traceability?

Schedule


Item I-A:
Project Initiation
9/15/2003
12/15/2003
Item I-B-1:
Watermarking Software
for Basic Countermeasures
9/15/2003
2/15/2004
Item I-B-2:
Demonstrate / validate
effectiveness against
countermeasures
9/15/2003
2/15/2004
Item I-C:
Solutions for Enhanced
Countermeasures
2/15/2004
10/15/2004
Item I-D:
Watermarking Software for
Enhanced Countermeasures
5/15/2004
3/15/2005
Item I-E:
Solve System Model Issues
9/15/2004
3/15/2005
Item II-A:
Software for System Model
3/15/2005
12/15/2005
Item II-B:
Solutions for Advanced
Countermeasuresl
3/15/2005
3/15/2006
Item II-C:
Technology Transfer to
Sponsor
12/15/2005
3/15/2006

Technology Transition Plans

- Reports and presentations to ARDA, and to intelligence community representatives
- Demonstrations and data
- Software
- Publications
- Students
- Commercialization? Not unless…

Deliverables (Already!)

- ✓ Website
- ✓ Presentation at ACM CCS 2003 (2 weeks ago)
- ✓ Application software with GUI

Questions or Comments?