Traffic Analysis

"Privacy-Enhancing Technologies" Course Talk

Willard Rafnsson

Department of Computer Science and Engineering
Chalmers University of Technology
Gothenburg, Sweden

8. June, 2012.

Outline

- Introduction
  - History
  - Our Setting
- TA on Tor
  - Tor recap
  - Attacker Model
  - Attack
  - Cost
- Countermeasures?
  - Classifiers
  - Countermeasures
  - Efficiency

Military

Deducing information from communication patterns.

- Frequent communication: Planning
  - Between same points: Chain of command
- Morse “hand”
- WWII: HMS Glorious

Military

Deducing information from communication patterns.

- Frequent communication: Planning
  - Between same points: Chain of command
- Morse “hand”
- WWII: HMS Glorious

Low-quality compared to cryptanalysis, but easy/cheap to extract/process, and hard/expensive to counter.

- TA to select target for cryptanalysis.

Computer Security Setting

The use of traffic data, that is,

- transmission-time,
- lenght, and
- direction

of network packets to/from victim,
to extract information sensitive to the victim.
Note: not content of packets (encrypted?)

Computer Security Setting

The use of traffic data, that is,

- transmission-time,
- lenght, and
- direction

of network packets to/from victim,
to extract information sensitive to the victim.
Note: not content of packets (encrypted?)

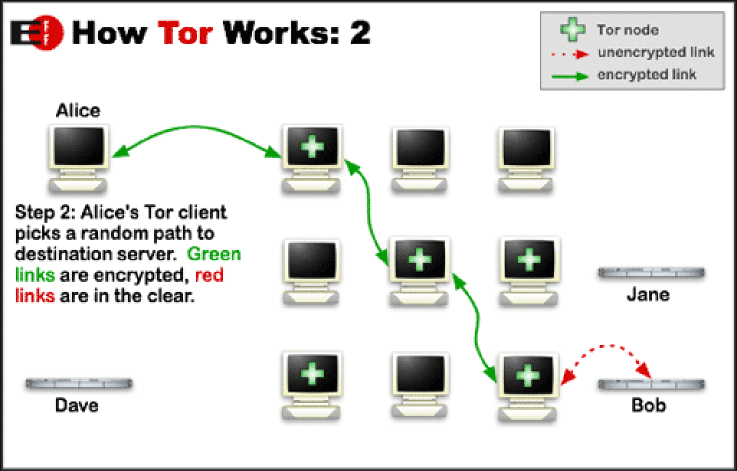
Typical objective: *Deanonymization*.

Example

"Timing Analysis of Keystrokes and Timing Attacks on SSH" (Song et al.)

Outline

Recall How Tor Works



Non-Global Attacker

Can

- observe,
- modify, and
- control

a fraction of the Tor network.

Non-Global Attacker

Can

- observe,
- modify, and
- control

a fraction of the Tor network.

Attacker can extract Tor connection path information.

Objective

Infer the nodes a stream goes through.

- know which OR stream begins at,
  - reduces anonymity.
- trace unrelated streams to same initiator.

Objective

Infer the nodes a stream goes through.

- know which OR stream begins at,
  - reduces anonymity.
- trace unrelated streams to same initiator.

Why possible:

- Each OR processes its streams in a round-robin fashion.
  - empty streams skipped to save time (low-latency demand)
- Adding a stream to a OR delays processing of existing streams at OR slightly.

The Attack

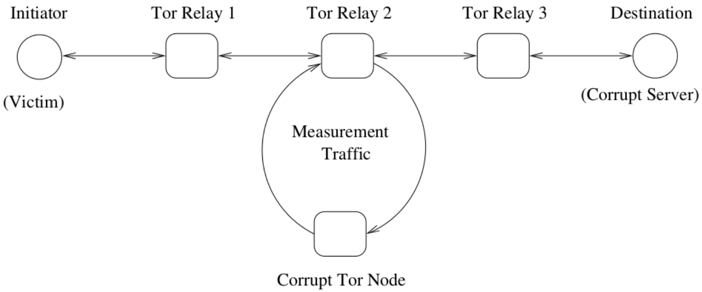


Figure 1. The attack setup

It's Cheap

To pull this off, you must be

- at the end of stream
  - compromised web server (trick victim)
  - man-in-the-middle (evil web hosting),
  - an exit node?
- able to probe all Tor nodes
  - $\geq$ 1 low-latency machine (same as end?)

It's Cheap

To pull this off, you must be

- at the end of stream
  - compromised web server (trick victim)
  - man-in-the-middle (evil web hosting),
  - an exit node?
- able to probe all Tor nodes
  - ≥ 1 low-latency machine (same as end?)

How it scales:

- N probe streams required.
- Attack cost: $O(N)$.

It's Cheap

To pull this off, you must be

- at the end of stream
  - compromised web server (trick victim)
  - man-in-the-middle (evil web hosting),
  - an exit node?
- able to probe all Tor nodes
  - ≥ 1 low-latency machine (same as end?)

How it scales:

- N probe streams required.
- Attack cost: $O(N)$.

For more:

"Low-Cost Traffic Analysis of Tor", S&P 2005

Outline

- Introduction
  - History
  - Our Setting
- TA on Tor
  - Tor recap
  - Attacker Model
  - Attack
  - Cost
- Countermeasures?
  - Classifiers
  - Countermeasures
  - Efficiency

Classifier

Recall what is observed:

- transmission-time,
- lenght, and
- direction

of network packets.

Classifier

Recall what is observed:

- transmission-time,
- lenght, and
- direction

of network packets.

Classifier: function which,
given (*length*, *direction*) (of a packet p),
returns a (guess of the) destination for p.

Classifier

Recall what is observed:

- transmission-time,
- lenght, and
- direction

of network packets.

Classifier: function which,
given $(length, direction)$ (of a packet p),
returns a (guess of the) destination for p.

Typically machine learning algorithms, trained on a large data set of traffic.

Observables

Recall what is observed:

- transmission-time, ← avoid change (latency)
- lenght, and
- direction ← cannot change

of network packets.

Padding seems like a good idea.

Padding

SSH/TLS/IPSec padding:

- Session Random 255 byte padding
- Packet Random 255 byte padding

Other:

- Linear (nearest $k128$)
- Exponential (nearest 2^k)
- Mice-Elephants
- MTU
- Packet Rnd. MTU (rnd $\{0, 8, \ldots, \mathtt{M} - l\}$)

C: Too much overhead ($\geq$ 40%) to be practical.

#Packets Obfuscation

Direct target sampling Define/derive distrib. D_{AB} over packets from A to B. When A sends a packet p of length i to B, instead,

- sample D_{AB} for smallest #lengths $i_1 \ldots i_k$ s.t. $\sum_{j=1}^{k-1} i_j < i \leq \sum_{j=1}^{k} i_j$,
- send first i_1 bytes of p as a packet, then next i_2, ... (pad last to i_k).

Traffic morphing (same idea, more complex to understand, more efficient in practise)

Too much overhead (40-80%) to be practical.

Transmission Time & Bandwidth

Not explored in literature:

- Total session time,
- Total bandwidth of data transmitted each direction,
- Transmission time of each packet
  - "burstiness" of packets.

Too much buffering / junk to be practical.

All Known Countermeasures Fail

Individual packet lengths need not be considered for high-accuracy classification

- padding & packet splitting ineffective countermeasure
  - fixed padding does not change bandwidth substantially
  - random padding "averages" out
  - burst information very informative
- best classifier still $> 80\%$ accuracy with privacy set size 128.

Simple Classifiers Are Accurate

Best classifier only marginally better than a naive Bayes classifier which *only* considers

- total session time,
- per-direction per-website bandwidth, and
- burst patterns.

VNG++ classifier developed.

VNG++ Counterm. Impractical

Buffered Fixed Length Obfuscator (BuFLO):

- fixed-interval send of
- fixed-length packets for a
- fixed minimum transmission time.

With well-configured parameters, best classifier down to 5.1% accuracy for privacy set size 128 (random guess is $\frac{1}{128} = 0.78\%$)

For more:

"Peek-a-Boo, I Still See You: Why Efficient Traffic Analysis Countermeasures Fail", S&P 201[12]