Towards Predicting Efficient and Anonymous Tor Circuits

Armon Barton\textsuperscript{1}, Mohsen Imani\textsuperscript{1}, Jiang Ming\textsuperscript{1}, and Matthew Wright\textsuperscript{2}

\textsuperscript{1}University of Texas at Arlington, \textsuperscript{2}Rochester Institute of Technology
Anonymity with Tor

- Millions of users
- 7,000 + Tor relays
Anonymity with Tor

- Millions of users
- 7,000 + Tor relays
Anonymity with Tor

Tor Network

Guard

Middle

Exit

WWW
Latency in Tor
Tor Congestion
Path Selection Algorithms

- Bandwidth Weighted Selection
- Snader and Borisov Selection
- Congestion Aware Routing
Bandwidth Weighted Path Selection
Tor Congestion
Tor Congestion

BW
Congestion Aware Routing
PredicTor

Tor Network

Trusted Authority

Consensus

Clients

k-NN
Random Forests
Feature Extraction

Guard

Middle

Exit

Training Features

AS

BW

日本

AS

BW

美国

AS

BW

德国

Training Label

True

False
Evaluating Accuracy

Shadow Simulation

- 1000 clients, 400 relays, 70 servers
- 320 KiB
- Training set: 120,000 streams
- Test set: 25,000 streams

Live Tor

- Server hosting 20 instances of Tor
- 80 KiB from a US server
- Training set: 50,000 streams
- Test set: 20,000 streams
Evaluating Accuracy

<table>
<thead>
<tr>
<th>Model</th>
<th>Shadow</th>
<th>Live Tor</th>
</tr>
</thead>
<tbody>
<tr>
<td>k-NN</td>
<td>70%</td>
<td>64%</td>
</tr>
<tr>
<td>Random Forests</td>
<td>76%</td>
<td>70%</td>
</tr>
</tbody>
</table>
PredicTor Evaluation

Implemented PredicTor in the Tor source code

- Tested on Shadow and Live Tor
- Compared with
  - BW (Vanilla)
  - Congestion Aware Routing (CAR)
  - Snader and Borisov (SB) - 9
  - SB-15
Shadow Experiment

PredicTor Improved Performance

- 23% compared to Vanilla
- 13% compared to CAR
- Speed up over 500ms in the med.
- Over 1.5s in the 90th.
- SB-9 and SB-15 performed the slowest.
Live Tor Experiment

PredicTor Improved Performance

- 11% compared to Vanilla
- 6% compared to CAR
- Over 1.0s in the 90th.
- SB-9 and SB-15 performed the slowest.
Live Tor Experiment

Circuit Bandwidth

- SB-9
  - 22% BW compared to Vanilla
- SB-15
  - 97% BW compared to Vanilla
- Indicates
  - relays experience persistent congestion.
  - performance gains in PredicTor are not solely attributed to BW.
Key Findings

Shadow Simulation

- Relay Utilization
  - SB-9, SB-15 utilized 50%
  - Vanilla utilized 65%
  - PredicTor utilized 85%
Key Findings

Live Tor Experiment

- Circuit Length
  - 680 km shorter compared to vanilla.
PredicTor Performance Gains

- Avoiding nodes with persistent congestion.
- Better relay utilization.
- Builds circuits of shorter geographic distance.
Security Evaluation

- Entropy based metrics
- All-or-nothing compromise
- AnoA Framework
Client AS Inference (CLASI)

\[ P' \leftarrow PS \]

\[ L' \leftarrow \text{Predict}(P') \]

\[ \text{Compare}(L, L') \]
Feature Extraction

Guard

Middle

Exit

AS
BW
Japan

AS
BW
USA

AS
BW
Germany

Client
AS

Destination

Training Features

Training Label
Client AS Inference (CLASI)

\[ Pr[L = L'] = \frac{1}{S_L} + \varepsilon_s \]
PredicTor Security Evaluation CLASI

![Graph showing network security data]
## PredicTor Security Evaluation

### Uniformity Degree

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Uniformity Degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vanilla</td>
<td>0.84</td>
</tr>
<tr>
<td>CAR</td>
<td>0.83</td>
</tr>
<tr>
<td>PredicTor</td>
<td>0.79</td>
</tr>
</tbody>
</table>
Conclusion

PredicTor performance gains

- Avoiding congestion
- Load distribution
- Shorter circuits

PredicTor security evaluation

- PredicTor had Similar sender AS leakage compared to Vanilla
- Lower AS leakage compared to CAR

Conclude: PredicTor had the best security / performance trade-off compared to both Vanilla and CAR.
Questions?