ExperimenTor: A Testbed for Safe and Realistic Tor Experimentation

Kevin Bauer¹  Micah Sherr²
Damon McCoy³  Dirk Grunwald⁴

¹University of Waterloo  ²Georgetown University
³UCSD  ⁴University of Colorado

http://crysp.uwaterloo.ca/software/exiptor
4th Workshop on Cyber Security Experimentation and Test (CSET ’11)
What is Tor and why is it important?

Tor is a low-latency overlay network and a software package that allows you to use TCP-based applications anonymously.

Tor has an estimated 350,000 daily users world-wide and its network consists of over 2,500 volunteer-operated Tor routers.

Ordinary Citizens
- Protect web browsing habits
- Research sensitive or taboo topics
- Circumvent censorship

Activists & Whistleblowers
- Expose human rights violations
- Promote democracy
- Protest election results

Corporations
- Research the competition
- Safeguard trade secrets

Law Enforcement
- Online surveillance
- Sting operations
Tor uses layered encryption to hide your online behaviors.

The Tor Network

Tor provides *anonymity for TCP applications* by tunneling traffic through a *virtual circuit* of three Tor routers using layered encryption. Communicating parties are *unlinkable* as long as the entry and exit routers do not collude.
Tor is still an evolving research network

• Past and current research aims to improve Tor’s:
  – Security and anonymity  [CCS ‘07, NDSS ’08, USENIX Security ‘10]
  – Quality of service  [USENIX Security ‘09, CCS ‘10, PETS ’11]

• **Problem:** There is no standard methodology for conducting Tor research in a *realistic* and *safe* manner; prior methods include:

---

**Realism**

- Abstract modeling
- Network simulators
- Small-scale network emulation
- Distributed research networks
- Live Tor network

**Safety**
ExperimenTor: A whole-network Tor emulation testbed

Goal: Propose a standard experimental methodology

- Replicates all components of the Tor network *in isolation*
- Reproduces plausible network conditions through *scalable network emulation*
- Fuels experiments with *empirically derived models*

ExperimenTor

Allows investigators to study global, whole-network effects
Talk outline

• Motivating case studies from prior Tor research
• Challenges of building a Tor network testbed
• Design and implementation of ExperimenTor
• Early experiences and lessons learned
• Conclusions and future work
Case study: Whole-network PlanetLab experiments

Uniform router selection: Probability of attack’s success is \((c/n)^2\), \(c\) malicious routers in a network of \(n\) total routers

Tor routers are selected in proportion to their perceived bandwidth capacities for load balancing, but malicious routers can lie
Case study: Whole-network PlanetLab experiments (2)

**Experiment:** Evaluated the attack on two small Planetlab deployments with 40 and 60 honest Tor routers

**Details:** Sample the bandwidth distribution of the real Tor network

<table>
<thead>
<tr>
<th>Tier</th>
<th>Tor Networks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Real Tor</td>
</tr>
<tr>
<td>996 KB/s</td>
<td>38</td>
</tr>
<tr>
<td>621 KB/s</td>
<td>43</td>
</tr>
<tr>
<td>362 KB/s</td>
<td>55</td>
</tr>
<tr>
<td>111 KB/s</td>
<td>140</td>
</tr>
<tr>
<td>29 KB/s</td>
<td>123</td>
</tr>
<tr>
<td>20 KB/s</td>
<td>21</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>103.9 MB/s</strong></td>
</tr>
</tbody>
</table>

**Limitations:**
1. Reduced scale
2. Need to run many measurements to find suitable PlanetLab nodes
3. Repeatability?

[Bauer et al., WPES ‘07]
Case study: Small-scale experiments with the live Tor network

*Tunable Tor* was proposed to help users manage their risk of the previous attack \[\text{[Snader and Borisov, NDSS '08]}\]

**Uniform selection**

**Skewed to high bandwidth nodes**

**User-tunable router selection**

**High anonymity**

**High performance**

**Experiment:** Deployed one “Tunable Tor” client on the live Tor network

**Details:** Measured download times at different “selection levels”

**Limitations:** What happens when *many* Tor clients use Tunable Tor? Global effects?
A case for whole-network Tor emulation

**Goal:** Capture all salient dynamics of the live Tor network and reproduce in isolation → **Realistic** and **safe** experiments

** Desired features:**

1. Allow investigators to deploy small-scale, large-scale, or global changes to any part of Tor’s design
2. Should eliminate any risk to the live Tor network
3. Experimental results should be meaningful to the live Tor network

**Our argument:** All can be realized with whole-network Tor emulation
Design challenges

• Modeling the live Tor network is difficult
  – Tor routers: Bandwidths, guard statuses, exit policies?
  – Tor clients: How many? Applications? Behaviors?

• Large-scale network emulation
  – Emulab and DETER have limited and shared resources

• Need to run unmodified Tor and application code
  – Avoids re-implementation errors; promotes realism
Meeting the design challenges

Modeling Tor routers:
- Publicly-available router metadata from Tor’s directories
- Historical router data aggregated by the Tor Metrics Portal

Replicate live Tor’s router state, or scale things up or down
Meeting the design challenges (2)

Modeling Tor clients: Leverage existing empirical data on Tor clients and their behaviors [McCoy et al., PETS ’08]

Number of Exit TCP Connections by Protocol

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Number of Connections</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>HTTP</td>
<td>12,160,437</td>
<td>92.45%</td>
</tr>
<tr>
<td>SSL</td>
<td>534,666 (4.06%)</td>
<td></td>
</tr>
<tr>
<td>BitTorrent</td>
<td>438,395 (3.33%)</td>
<td></td>
</tr>
<tr>
<td>Instant Messaging</td>
<td>10,506 (0.08%)</td>
<td></td>
</tr>
<tr>
<td>E-Mail</td>
<td>7,611 (0.06%)</td>
<td></td>
</tr>
<tr>
<td>FTP</td>
<td>1,338 (0.01%)</td>
<td></td>
</tr>
<tr>
<td>Telnet</td>
<td>1,045 (0.01%)</td>
<td></td>
</tr>
</tbody>
</table>

Also leverage existing empirical studies of HTTP traffic to emulate realistic workloads [e.g., Hernández-Campos et al., MASCOTS ’03; Google web metrics 2010]
Meeting the design challenges (3)

**Modeling Tor clients:** Leverage existing empirical data on Tor clients and their behaviors  [McCoy et al., PETS ’08]

Aggregate Exit Traffic Volume by Protocol (GB)

- HTTP: 411 GB (57.90%)
- SSL: 11 GB (1.55%)
- BitTorrent: 285 GB (40.20%)
- Instant Messaging: 735 MB (0.10%)
- E-Mail: 291 MB (0.04%)
- FTP: 792 MB (0.11%)
- Telnet: 110 MB (0.02%)

Model the distribution of client traffic by connection and volume
Meeting the design challenges (4)

**Modeling Tor clients:** Can also leverage publicly-available data about Tor clients from the Tor Metrics Project

![Graph showing the number of users directly connecting from all countries between May 2011 and July 2011.](https://metrics.torproject.org/)

Replicate live Tor’s client state, or scale things up or down
Meeting the design challenges (5)

Large-scale network emulation with ModelNet [Vahdat et al., OSDI ‘02]
- Emulates a specified network topology
- Runs native code without modification
- Commodity hardware and OSes; can be deployed at local institution

• High-level system architecture
  - “Emulator” machine: Emulates a network topology in a kernel module
  - “Virtual node” machine: Runs applications within the virtual topology

Applications run on “virtual node” machines
Putting it all together

Network topology emulation on a ModelNet core

Per-link bandwidth, latency, queues, drop rate

Prototype: FreeBSD 6.3
Linux 2.6.32

Accompanying toolkit:
- Topology generation
- Configure Tor clients, routers, & directories
- Run experiments & perform analyses

Tor and applications run on edge nodes in virtual topology

Testbed and toolkit are publicly available
http://crysp.uwaterloo.ca/software/exptor
Early experiences

• ExperimenTor prototypes are deployed at four research institutions (single emulator)
• Used to support two ongoing research projects:
  – Evaluate the effects of link-based router selection
  – Re-design Tor’s congestion control and flow control
• Both projects require global design changes to Tor
Limitations and future work

- **Scalability**
  - Scaling experiments to Tor’s estimated 350K users is likely not possible; necessary to “down sample”

- **Improve client and traffic models**
  - Data on Tor usage are limited
  - Is it possible to emulate diverse versions and configurations of Tor users?
Summary and conclusion

• Experimenter is a whole-network emulation-based testbed and toolkit for safe and realistic Tor experiments

• Enables large-scale Tor experiments that:
  – Use real Tor router bandwidths to inform topology
  – Emulate Tor clients and their traffic
  – Enable experiments with global changes to Tor’s design
  – Can be deployed cheaply on commodity systems

For more information:
http://crysp.uwaterloo.ca/software/exptor