Tarzan:
A Peer-to-Peer Anonymizing Network Layer

Michael J. Freedman, NYU

Robert Morris, MIT

ACM CCS 2002

http://pdos.lcs.mit.edu/tarzan/
The Grail of Anonymization

- Participant can communicate *anonymously* with non-participant

- User can talk to CNN.com

- *Nobody* knows who user is
Our Vision for Anonymization

- Thousands of nodes participate
- Bounce traffic off one another

- Mechanism to organize nodes: peer-to-peer
- All applications can use: IP layer
Alternative 1: Proxy Approach

- Intermediate node to proxy traffic
- Completely trust the proxy

Anonymizer.com
Threat model

• Corrupt proxy(s)
  – Adversary runs proxy(s)
  – Adversary targets proxy(s) and compromises, possibly adaptively

• Network links observed
  – Limited, localized network sniffing
  – Wide-spread (even global) eavesdropping

  e.g., Carnivore, Chinese firewall, ISP search warrants
Failures of Proxy Approach

- Proxy reveals identity
- Traffic analysis is easy
Failures of Proxy Approach

- Proxy reveals identity
- Traffic analysis is easy
- CNN blocks connections from proxy
- Adversary blocks access to proxy (DoS)
Alternative 2: Centralized Mixnet

- MIX encoding creates encrypted tunnel of relays
  - Individual malicious relays cannot reveal identity
- Packet forwarding through tunnel

Onion Routing, Freedom

Small-scale, static network
Failures of Centralized Mixnet

- CNN blocks core routers
Failures of Centralized Mixnet

- CNN blocks core routers
- Adversary targets core routers
Alternative 2: Centralized Mixnet

- CNN blocks core routers
- Adversary targets core routers
- So, add cover traffic between relays
Failures of Centralized Mixnet

- CNN blocks core routers
- Adversary targets core routers
Failures of Centralized Mixnet

- CNN blocks core routers
- Adversary targets core routers
- Still allows network-edge analysis
Failures of Centralized Mixnet

- Internal cover traffic does not protect edges
- External cover traffic prohibitively expensive?
  - $n^2$ communication complexity
Tarzan goals

- No distinction between anon proxies and clients
- Anonymity against corrupt relays
  - MIX-net encoding
  - Robust tunnel selection
  - Prevent adversary spoofing or running many nodes
- Anonymity against global eavesdropping
  - Cover traffic protects all edges
  - Restrict topology to make cover practical
  - Choose neighbors in verifiably-random manner
- Application-independence
  - Low-latency IP-layer redirection
Tarzan: Me Relay, You Relay

- Thousands of nodes participate
  - CNN cannot block everybody
  - Adversary cannot target everybody
Tarzan: Me Relay, You Relay

- Thousands of nodes participate
- Cover traffic protects all nodes
  - Global eavesdropping gains little info
Benefits of Peer-to-Peer Design

• Thousands of nodes participate
• Cover traffic protects all nodes
• All nodes also act as relays
  – No network edge to analyze
  – First hop does not know he’s first
Tarzan goals

• No distinction between anon proxies and clients

• Anonymity against corrupt relays

• Anonymity against global eavesdropping

• Application-independence
1. Contacts known peers to learn neighbor lists
2. Validates each peer by directly pinging
4. Nodes begin passing cover traffic with **mimics**:

- Nodes send at some traffic rate per time period
- Traffic rate independent of actual demand
- All packets are same length and link encrypted
5. To build tunnel:

Iteratively selects peers and builds tunnel from among last-hop’s mimics
But, Adversaries Can Join System

User

PNAT
But, Adversaries Can Join System

- Adversary can join more than once by spoofing addresses outside its control

✓ Contact peers directly to validate IP addr and learn PK
But, Adversaries Can Join System

- Adversary can join more than once by running many nodes on each machine it controls

✓ Randomly select by subnet “domain” (/16 prefix, not IP)
But, Adversaries Can Join System

- Adversary can join more than once by running many nodes on each machine it controls

✓ Randomly select by subnet “domain” (／16 prefix, not IP)
But, Adversaries Can Join System

- Colluding adversary can only select each other as neighbors

✓ Choose mimics in universally-verifiable random manner
Tarzan: Selecting mimics

3. Nodes pair-wise choose (verifiable) mimics
Tarzan goals

- No distinction between anon proxies and clients
  - Peer-to-peer model

- Anonymity against corrupt relays
  - MIX-net encoding
  - Robust tunnel selection
  - Prevent adversary spoofing or running many nodes

- Anonymity against global eavesdropping
  - Cover traffic protects all nodes
  - Restrict topology to make cover practical
  - Choose neighbors in verifiably-random manner

- Application-independence
  - Low-latency IP-layer redirection
5. To build tunnel:

Public-key encrypts tunnel info during setup
Maps flowid $\rightarrow$ session key, next hop IP addr
6. Reroutes packets over this tunnel

Diverts packets to tunnel source router
6. Reroutes packets over this tunnel

NATs to private address 192.168.x.x

Pads packet to fixed length
6. Reroutes packets over this tunnel

Layer encrypts packet to each relay

Encapsulates in UDP, forwards to first hop
6. Reroutes packets over this tunnel

- Strips off encryption
- Forwards to next hop within cover traffic
6. Reroutes packets over this tunnel

NATs again to public alias address
Tarzan: Tunneling Data Traffic

6. Reroutes packets over this tunnel

Reads IP headers and sends accordingly
6. Reroutes packets over this tunnel

Response repeats process in reverse
Integrating Tarzan

Use transparently with existing systems

Can build double-blinded channels
Packet forwarding and tunnel setup

• Tunnel Setup (public key ops)

  ~30 msec / hop latency + network delay

• Packet forwarding (without cover traffic)

<table>
<thead>
<tr>
<th>pkt size</th>
<th>latency</th>
<th>throughput</th>
</tr>
</thead>
<tbody>
<tr>
<td>64 bytes</td>
<td>250 µsec</td>
<td>7 Mbits/s</td>
</tr>
<tr>
<td>1024 bytes</td>
<td>600 µsec</td>
<td>60 MBits/s</td>
</tr>
</tbody>
</table>
Summary

• Application-independence at IP layer
  – Previous systems for email, web, file-sharing, etc.

• No network edge through peer-to-peer design
  – Core routers can be blocked, targeted, or black-box analyzed

• Anonymity against corrupt relays and global eavesdropping
  – Cover traffic within restricted topology
  – MIX-net tunneling through verified mimics

• Scale to thousands
  – Towards a critical mass of users
http://pdos.lcs.mit.edu/tarzan/
Packet forwarding and tunnel setup

<table>
<thead>
<tr>
<th>Pkt size (bytes)</th>
<th>Latency (μ-sec)</th>
<th>Throughput (pkts/s)</th>
<th>Throughput (Mbits/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>64</td>
<td>244</td>
<td>14000</td>
<td>7.2</td>
</tr>
<tr>
<td>512</td>
<td>376</td>
<td>8550</td>
<td>35.0</td>
</tr>
<tr>
<td>1024</td>
<td>601</td>
<td>7325</td>
<td>60.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tunnel length</th>
<th>Setup latency</th>
<th>Variance (1 StD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>30.19</td>
<td>1.38</td>
</tr>
<tr>
<td>2</td>
<td>46.54</td>
<td>0.53</td>
</tr>
<tr>
<td>3</td>
<td>68.37</td>
<td>0.73</td>
</tr>
<tr>
<td>4</td>
<td>91.55</td>
<td>1.20</td>
</tr>
</tbody>
</table>

(msec)