The effect of DNS on Tor’s anonymity

Benjamin Greschbach  
KTH Royal Institute of Technology

Tobias Pulls  
Karlstad University

Laura M. Roberts  
Princeton University

Philipp Winter  
Princeton University

Nick Feamster  
Princeton University
Internet Sehat

Sebagai bentuk dukungan terhadap program "Internet Sehat" milik pemerintah Indonesia, maka kami menutup akses ke situs yang Anda kunjungi.

We are blocking this abusive site as stated by the Indonesia regulation in order to provide Internet Sehat

Apabila Anda merasa situs yang dikunjungi tidak bertentangan dengan program pemerintah silakan hubungi bagian pengaduan konten depkominfo:

If you think this site is not containing abusive content, please contact depkominfo content department at:

aduankonten @ mail.kominfo.go.id

Untuk informasi lebih jelas, silahkan kunjungi website:

Further information, please go to:

http://trustpositif.kominfo.go.id
This site has been blocked as per the instructions of Competent Authority
How does DNS work over Tor?

Tor client  Guard  Middle  Exit  example.com
How does DNS work over Tor?

Where's example.com?

Tor client  Guard  Middle  Exit  example.com
How does DNS work over Tor?

Tor client → Guard → Middle → Exit → DNS resolver

example.com
How does DNS work over Tor?

Example.com

DNS resolver

Where's example.com?

Tor client
Guard
Middle
Exit

example.com
How does DNS work over Tor?

- Tor client
- Guard
- Middle
- Exit

DNS resolver

Where's example.com?
How does DNS work over Tor?

Tor client → Guard → Middle → Exit → example.com

Where's example.com?
How does DNS work over Tor?

- Tor client
- Guard
- Middle
- Exit
- DNS resolver
- example.com

Where's example.com?
How exposed are DNS queries?
How exposed are DNS queries?
How exposed are DNS queries?

- Simulate resolution process for Alexa top 1,000
  - Run traceroutes for DNS delegation path
  - Run traceroutes to web server IP address
  - Map IP addresses to autonomous system numbers
- For half of all websites, 57% or more ASes were only traversed for DNS
- New class of adversaries
What resolvers do exit relays use?
What resolvers do exit relays use?

- Machines under our control
- Guard relay
- Tor network
- Exit relays
- Third-party DNS resolver

DD8BD7307017407FCC36F8D04A688F74A0774C02.2017-02-17-08.tor.nymity.ch
A10C4F666D27364036B562823E5830BC448E046A.2017-02-17-08.tor.nymity.ch
...
What resolvers do exit relays use?
What resolvers do exit relays use?
What resolvers do exit relays use?

<table>
<thead>
<tr>
<th>Resolver</th>
<th>Min (%)</th>
<th>Max (%)</th>
<th>Median (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Google</td>
<td>23.57</td>
<td>42.33</td>
<td>32.84</td>
</tr>
<tr>
<td>Local</td>
<td>7.71</td>
<td>15.95</td>
<td>11.56</td>
</tr>
<tr>
<td>OVH</td>
<td>1.96</td>
<td>14.13</td>
<td>6.57</td>
</tr>
<tr>
<td>OpenDNS</td>
<td>0.05</td>
<td>5.62</td>
<td>0.76</td>
</tr>
</tbody>
</table>

Percentage of observed DNS queries
Can we improve website fingerprinting attacks?
Can we improve website fingerprinting attacks?
Can we improve website fingerprinting attacks?

- We extended Wang et al.’s Wa-kNN classifier (USENIX Security’14)

- **High precision** attack
  - Training phase identical to Wa-kNN
  - Testing phase throws out sites that weren’t observed in DNS traffic when calculating nearest neighbors

- **Close-the-world** attack
  - Accepts Wa-kNN’s website classification only if that website was observed in DNS traffic

- **Great results for unpopular websites**
  - Small anonymity set to hide in
Our attacks at Internet-scale

- Place Tor clients in top five usage countries
- Simulate clients’ online behavior
  - Cf. Johnson et al. CCS’13
- Simulate Tor clients’ path selection
  - TorPS (github.com/torps/torps)
- Run traceroutes client → guard and exit → destination
  - Use RIPE Atlas!
- Check for overlapping autonomous systems
  - Simple set intersection
RIPE Atlas for traceroutes
RIPE Atlas for traceroutes
Fraction of compromised streams

(a) The fraction of compromised streams of simulated Tor clients.
(b) The time until simulated Tor clients got first compromised.
How do we fix this mess?

- Reach out to exit relay operators
  - Don’t use Google
  - Use QNAME minimisation

- Add confidentiality to DNS
  - T-DNS (Zhu et al. Oakland’15)
  - Push for more onion services
  - Improve website fingerprinting defenses
Thanks

- Paper, data, code, and replication:
  - https://nymity.ch/tor-dns/

- Contact
  - pwinter@cs.princeton.edu
  - @_phw