Who Is Answering My Queries?
Understanding and Characterizing Hidden Interception of the DNS Resolution Path

Baojun Liu, Chaoyi Lu, Haixin Duan,
Ying Liu, Zhou Li, Shuang Hao and Min Yang
DNS Resolution

- **ISP DNS Resolver**
  - Might have security problems [Dagon, NDSS’08] [Weaver, SATIN’11] [Weaver, FOCI’11] [Kuhrer, IMC’15] [Chung, IMC’16] …

![Diagram showing DNS resolution process](image-url)
DNS Resolution

• Public DNS Resolver
  – Performance (e.g., load balancing)
  – Security (e.g., DNSSEC support)
  – DNS extension (e.g., EDNS Client Subnet)
DNS Interception

• Who is answering my queries?

Spoof the IP address and **intercept** queries.
Potential Interceptors

Internet Service Provider (ISP)

Censorship / firewall

Anti-virus software / malware
(E.g., Avast anti-virus)

Enterprise proxy
(E.g., Cisco Umbrella intelligent proxy)
Q1: How to **globally measure** the hidden DNS interception?

Q2: What are the **characteristics** of the hidden DNS interception?
Motivation

Threat Model

Methodology

Analysis
Threat Model

1. Client -> On-path Device
2. On-path Device -> Public DNS
3. Public DNS -> Root NS
4. Public DNS -> TLD NS
5. Public DNS -> SLD NS
6. Middlebox -> On-path Device

original path
intercepted path
Threat Model

- Taxonomy (request only)
  - [1] Normal resolution

![Diagram of DNS resolution process]

- Client requests to 8.8.8.8
  - On-path Device
  - Public DNS 8.8.8.8
  - Alternative resolver 1.2.3.4
  - Authoritative nameserver
Threat Model

- Taxonomy (request only)
  - [2] Request redirection

Client → Request to 8.8.8.8 → On-path Device → Public DNS 8.8.8.8 → Authoritative nameserver

Request redirection from 1.2.3.4 → Alternative resolver 1.2.3.4
Threat Model

• Taxonomy (request only)
Threat Model

- **Taxonomy (request only)**
  - [4] Direct responding
How to Detect?

- At a glance

**Send DNS requests.**

Client → Request to 8.8.8.8 → On-path Device → Public DNS 8.8.8.8 → Alternative resolver 1.2.3.4 → Authoritative nameserver

Check where they are from.

From 8.8.8.8 → From 1.2.3.4
How to Detect?

1. Open the refrigerator
2. Put in the elephant
3. Close the door

1. Collect vantage points
2. Send DNS requests
3. Collect requests on NS

*Pic source: cdc.tencent.com
Collect vantage points

Diversify DNS requests

Identify egress IP
Vantage Points

• Requirements
  – Ethical
  – Large-scale and geo-diverse
  – Directly send DNS packets to specified IP
Measurement frameworks

- **Advertisement Networks**
  - Flash applet [Huang, W2SP’11] [Chen, CCS’16]
  - JavaScript [Burnett, Sigcomm’15]
- **HTTP Proxy Networks**
  - Luminati [Chung, IMC’16] [Tyson, WWW’17], [Chung, Security’17]
- **Internet Scanners**
  - Open DNS resolver [Kuhrer, IMC’15] [Pearce, Security’17]
  - Scanners [Zakir, Security’13] [Pearce, SP’17]

*Cannot be used in this study.*
Vantage Points

- **Phase I: Global Analysis**
  - ProxyRack: SOCKS5 residential proxy networks
  - Limitation: **TCP** traffic only
Vantage Points

• Phase I: Global Analysis
  – ProxyRack: SOCKS5 residential proxy networks
  – Limitation: **TCP** traffic only

• Phase II: China-wide Analysis
  – A network debugger module of security software
  – Similar to *Netalyzr* [Kreibich, IMC’ 10]
  – Capability: **TCP and UDP; Socket level**
# Vantage Points

- **Ethics considerations**

<table>
<thead>
<tr>
<th>Global (ProxyRack)</th>
<th>Pay for access</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Abide by ToS</td>
</tr>
<tr>
<td></td>
<td>Only query our domain</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>China-wide (network debugging tool)</th>
<th>One-time consent</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Restrict traffic amount</td>
</tr>
<tr>
<td></td>
<td>Only query our domain</td>
</tr>
</tbody>
</table>
Collect vantage points

Diversify DNS requests

Identify egress IP
DNS Requests

• **Requirements**
  – **Diverse**: triggering interception behaviors
  – **Controlled**: allowing fine-grained analysis

<table>
<thead>
<tr>
<th>Public DNS</th>
<th>Google, OpenDNS, Dynamic DNS, EDU DNS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protocol</td>
<td>TCP, UDP</td>
</tr>
<tr>
<td>QTYPE</td>
<td>A, AAAA, CNAME, MX, NS</td>
</tr>
<tr>
<td>QNAME (TLD)</td>
<td>com, net, org, club</td>
</tr>
<tr>
<td>QNAME</td>
<td>UUID.[Google].OurDomain. [TLD]</td>
</tr>
</tbody>
</table>
Collect vantage points

Diversify DNS requests

Identify egress IP
Egress IP

• Ownership of resolver IP
  – Is a request from public DNS?
Egress IP

• Ownership of resolver IP
  – Is a request from public DNS?

• Solution
  – PTR & SOA records of reverse lookups

$ dig -x 74.125.41.1

;; AUTHORITY SECTION:
125.74.in-addr.arpa.60 IN SOA ns1.google.com.
dns-admin.google.com. 207217296 900 900 1800 60
Collected Dataset

- DNS requests from vantage points
  - A wide range of requests collected

<table>
<thead>
<tr>
<th>Phase</th>
<th># Request</th>
<th># IP</th>
<th># Country</th>
<th># AS</th>
</tr>
</thead>
<tbody>
<tr>
<td>ProxyRack</td>
<td>1.6 M</td>
<td>36K</td>
<td>173</td>
<td>2,691</td>
</tr>
<tr>
<td>Debugging tool</td>
<td>4.6 M</td>
<td>112K</td>
<td>87</td>
<td>356</td>
</tr>
</tbody>
</table>
Q1: Interception Characteristics
Q2: DNS Lookup Performance
Q3: Response Manipulation
Q4: Security Threats
Q5: Interception Motivations
Q6: Solutions
Interception Characteristics

- Magnitude (% of total requests)
  - Normal resolution
  - Request redirection
  - Request replication

<table>
<thead>
<tr>
<th>Service</th>
<th>Direct Responding</th>
<th>Request Redirection</th>
<th>Request Replication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Google</td>
<td>72.1%</td>
<td>22.3%</td>
<td>5.6%</td>
</tr>
<tr>
<td>OpenDNS</td>
<td>87.4%</td>
<td>7.8%</td>
<td>4.8%</td>
</tr>
<tr>
<td>Dyn DNS</td>
<td>83.9%</td>
<td>9.7%</td>
<td>6.3%</td>
</tr>
<tr>
<td>EDU DNS</td>
<td>90.2%</td>
<td>6.3%</td>
<td>3.5%</td>
</tr>
</tbody>
</table>

Direct responding is rare.

Request redirection > Request replication
Interception Characteristics

• Magnitude (% of total requests)
  – Normal resolution  Request redirection  Request replication

Requests to popular public DNS services are more likely to be intercepted.
Interception Characteristics

- **ASes (% of total requests)**
  - Sorted by # of total requests

<table>
<thead>
<tr>
<th>AS</th>
<th>Organization</th>
<th>Redirection</th>
<th>Replication</th>
<th>Alternative Resolver</th>
</tr>
</thead>
<tbody>
<tr>
<td>AS4134</td>
<td>China Telecom</td>
<td>5.19%</td>
<td>0.2%</td>
<td>116.9.94.* (AS4134)</td>
</tr>
<tr>
<td>AS4837</td>
<td>China Unicom</td>
<td>4.59%</td>
<td>0.51%</td>
<td>202.99.96.* (AS4837)</td>
</tr>
<tr>
<td>AS9808</td>
<td>China Mobile</td>
<td><strong>32.49%</strong></td>
<td><strong>8.85%</strong></td>
<td>112.25.12.* (AS9808)</td>
</tr>
<tr>
<td>AS56040</td>
<td>China Mobile</td>
<td><strong>45.09%</strong></td>
<td>0.04%</td>
<td>120.196.165.* (AS56040)</td>
</tr>
</tbody>
</table>

Interception strategies can be complex, and vary among ASes.
DNS Lookup Performance

- RTT of requests
  - Which requests complete faster?

Request replication vs. Normal resolution: Better.

Request redirection vs. Request to local resolver: Very similar.
DNS Lookup Performance

• Arrival time of replicated requests
  – Which requests reach NS faster?

In AS4812, **ALL** replicated requests arrive **slower than** their original counterparts.
Response Manipulation

- DNS record values
  - Which responses are tampered?

<table>
<thead>
<tr>
<th>Classification</th>
<th>#</th>
<th>Response Example</th>
<th>Client AS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gateway</td>
<td>54</td>
<td>192.168.32.1</td>
<td>AS4134, CN, China Telecom</td>
</tr>
<tr>
<td>Monetization</td>
<td>10</td>
<td>39.130.151.30</td>
<td>AS9808, CN, GD Mobile</td>
</tr>
<tr>
<td>Misconfiguration</td>
<td>26</td>
<td>::218.207.212.91</td>
<td>AS9808, CN, GD Mobile</td>
</tr>
<tr>
<td>Others</td>
<td>54</td>
<td>fe80::1</td>
<td>AS4837, CN, China Unicom</td>
</tr>
</tbody>
</table>
Response Manipulation

• Example: traffic monetization

China Mobile Group of Yunnan: advertisements of an APP.
Security Threats

• Ethics & privacy
  – Users may **not be aware** of the interception behavior

• Alternative resolvers’ security
  – An analysis on **205 open alternative resolvers**

  ![DNSSEC](image1.png)
  ![BIND](image2.png)

  **Only 43% resolvers support DNSSEC**

  **ALL BIND versions should be deprecated before 2009**
Interception Motivations

• Vendors
  – Routers
  – Software platforms

• Motivations
  – Improving DNS security ?
  – Improving DNS lookup performance ?
  – Reducing traffic financial settlement ✔
Solutions

• Encrypted DNS
  – Resolver authentication (RFC8310)
  – DNS-over-TLS (RFC7858)
  – DNS-over-DTLS (RFC8094, experimental)
  – DNS-over-HTTPS

• Online checking tool
  – Which resolver are you really using?
  – http://whatismydnsresolver.com/
Conclusions

• Understanding
  – A measurement platform to systematically study DNS interception

• Findings
  – DNS interception exists in 259 ASes we inspected globally
  – Up to 28% requests from China to Google are intercepted
  – Brings security concerns

• Motivations
  – Reducing traffic financial settlement

• Mitigation
  – Online checking tool
Who Is Answering My Queries?
Understanding and Characterizing Hidden Interception of the DNS Resolution Path

Baojun Liu, Chaoyi Lu, Haixin Duan,
Ying Liu, Zhou Li, Shuang Hao and Min Yang

lbj15@mails.tsinghua.edu.cn