Online Privacy and ISPs: New Factual Study

Peter Swire
Huang Professor of Law and Ethics
Georgia Institute of Technology

January 25, 2016
Introduction

- Released yesterday:
  - “Online Privacy and ISPs: ISP Access to Consumer Data is Limited and Often Less than Access by Others”
  - Co-authors Justin Hemmings & Alana Kirkland
  - iisp.gatech.edu/policy
- ISP access to user data is not comprehensive
- ISP access to user data is not unique
- Factual report, no policy recommendations
- Any regulatory action should be based on an accurate factual record
ISPs See Less Than You Think

Non-ISPs have access to substantial user online information while encryption and VPNs block ISP access to traffic.

And

Individuals use multiple devices and individuals use more than one ISP.

ISPs do not have comprehensive or unique visibility into users' online activity.
Limits on ISP Visibility: Encryption

• Pervasive Encryption:
  – HTTPS blocks ISP visibility into content and full URLs
  – HTTPS allows ISPs to see only host name, www.example.com

• Appendix 1
  – Top 10 web sites either encrypt by default or upon user log-in
  – 42 of top 50 web sites either encrypt by default or upon user log-in

• Appendix 2: CAIDA Internet backbone data
  – April 2014 share of HTTPS: 13%
  – February 2016 share of HTTPS: 49%
APPLICATION BITS/S - 1 WEEK

FEBRUARY 03, 2016 - FEBRUARY 10, 2016 UTC

<table>
<thead>
<tr>
<th>Application</th>
<th>Min</th>
<th>Avg</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>HTTPS</td>
<td>1.04G</td>
<td>1.53G</td>
<td>2.12G</td>
</tr>
<tr>
<td>HTTP</td>
<td>480.93M</td>
<td>1.14G</td>
<td>2.31G</td>
</tr>
<tr>
<td>UNKNOWN_UDP</td>
<td>89.05M</td>
<td>232.12M</td>
<td>454.18M</td>
</tr>
<tr>
<td>UNKNOWN_TCP</td>
<td>80.40M</td>
<td>180.76M</td>
<td>516.36M</td>
</tr>
</tbody>
</table>
Contrasting Views

• Letter last week:
  ISPs have “comprehensive” visibility into user Internet activity

• Our research:
  • No
  • SSL encryption blocks content
  • SSL encryption blocks detailed links
  • Encryption is becoming predominant

Diagram 4 - screen shot for week in Feb 2016 statistics.
THE EFFECTS OF HTTPS:

STEP 1 - DNS LOOKUP

I want to visit comfyshoes.com

Now I know how to locate comfyshoes

Can you tell me where comfyshoes is located?

comfyshoes is located at 208.67.222.222

comfyshoes.com

location: 208.67.222.222

ISP

DNS

STEP 2 - CONTENT RETRIEVAL

OVER HTTP

Can you send me to 208.67.222.222?

User can browse comfyshoes

Welcome to ComfyShoes!

ISP retrieves content from 208.67.222.22

ISP delivers content to user

OVER HTTPS

Over HTTPS, ISPs can no longer see full URLs or content

Can you send me to 208.67.222.222?

User can browse comfyshoes

Q29tZnJTaG9tYWQ=

ISP retrieves content from 208.67.222.22

ISP delivers content to user
WITH VIRTUAL PRIVATE NETWORKS

1. Establishing a connection to a VPN

- **USER**: I want to establish a connection with my VPN
- **ISP**: The user wants to establish a secure connection with you.
- **VPN** (proxy server): OK, establish a secure connection with the user

Connection with VPN established! ENCRYPTED TUNNEL
WITH VIRTUAL PRIVATE NETWORKS

2. DNS Lookup

I want to go to HTTPS://www.comfyshoes.com

ISP
Forwards encrypted messages between user and VPN

VPN (proxy server)
Decrypts message and performs DNS lookup for IP address of comfyshoes.com

comfyshoes.com is located at 916.47.63.137

In this scenario, the VPN has its own DNS lookup System

Connection with VPN established!

WHAT THEY CAN SEE
WWW
HOST NAME
FULL URLs
CONTENT

DNS

encrypted tunnel encrypted tunnel encrypted tunnel encrypted tunnel
WITH VIRTUAL PRIVATE NETWORKS

3. Browsing a Website after the IP address has been retrieved

- User can browse comfyshoes
- ISP forwards encrypted messages between user and VPN
- VPN (proxy server) retrieves content from 208.67.222.22
- VPN delivers content to the user

Legend:
- WHAT THEY CAN SEE
  - FULL URLS
  - CONTENT
- HOST NAME

Welcome to: COMFYSHOES.COM!
ISP's See Less Than You Think

Non-ISP's have access to substantial user online information while

Encryption and VPNs block ISP access to traffic

And

Individuals use multiple devices

And

Individuals use more than one ISP

ISP's do not have comprehensive or unique visibility into users' online activity
“Unique” Visibility of ISPs?

• “Contexts” and cross-context tracking:
  – Social networks
  – Search
  – Webmail/messaging
  – OS, especially mobile
  – Mobile apps
  – Interest-based advertising
  – Browsers
  – Internet video
  – E-commerce

• Combining user data about any one device
• Then, cross-device tracking
Conclusion

• Report seeks to describe online ecosystem, personal information, the role of ISPs and other major players
• Policy should be based on accurate facts about today’s and tomorrow’s markets
• We welcome comments and corrections
Credits

• Images for attribution to Marie Le Pichon, under a Creative Commons Attribution 4.0 license.
• Research support from: Georgia Tech Scheller College of Business; Georgia Tech Institute for Information Security and Privacy; Broadband for America.