10. Privacy and Trust on the Web

Blase Ur and David Cash
February 3rd, 2020
CMSC 23200 / 33250
Online Tracking

- Advertisers want to show you advertisements targeted to your interests and demographics.
Online Tracking

• First party = the site you are visiting (whose address is in the URL bar)
• Third party = other sites contacted as a result of your visit to that site
• First-party tracking (e.g., for search)
  – Consider DuckDuckGo and alternatives
Online Tracking

• JavaScript / images from advertising networks loaded as part of your page
  – In iframes
  – Or sometimes not
  – Why does this matter?
  – Does this also apply to email? (Yes)
Ubiquity of Online Tracking

<table>
<thead>
<tr>
<th>TRACKERS</th>
<th>Block All</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Advertising</strong></td>
<td></td>
</tr>
<tr>
<td>10 TRACKERS</td>
<td>3 Blocked</td>
</tr>
<tr>
<td>Advertising.com</td>
<td>✓</td>
</tr>
<tr>
<td>DoubleClick</td>
<td></td>
</tr>
<tr>
<td>Google Adsense</td>
<td></td>
</tr>
<tr>
<td>Korrelate</td>
<td></td>
</tr>
<tr>
<td>Metav</td>
<td>✓</td>
</tr>
<tr>
<td>NetRatings-Site-Center</td>
<td>✓</td>
</tr>
<tr>
<td>Polar-Mobile</td>
<td>✓</td>
</tr>
<tr>
<td>ScoreCard Research Beacon</td>
<td></td>
</tr>
<tr>
<td>Tacoda</td>
<td></td>
</tr>
</tbody>
</table>

22 Trackers found on www.mynews.com

4 Blocked

- Trust Site
- Restrict Site
- Pause Ghostery
- Map these trackers
Ad Bidding Marketplaces

DATA LEAKAGE IN ONLINE ADVERTISING

This is the current process of real-time bidding that is used in online behavioural advertising.

Legend
- Channel of data leakage
- Money
- Personally identifiable information

Step 1. User requests webpage

Step 2. Ad server selects an SSP

Step 3. SSP selects an exchange

Step 4. Exchange sends bid requests to hundreds of partners

Step 5. Exchange lets some DMPs/DSPs to refresh cookie sync

Step 6. Exchange serves winning bid

Step 7. DSP serves agency creative

Step 8. Assets load from CDN

Step 9. Agency ad server loads verification vendor

Winning DSP

PageFair
Do not track

- Proposed W3C standard
- User checks a box
- Browser sends “do not track” header to website
- Website stops “tracking”
- W3C working group trying to define what that means
Tools to stop tracking, effective?

- Browser privacy settings
  - Cookie blocking
  - P3P
  - Tracking Protection Lists
  - Do Not Track
- Browser add-ons
- Opt-out cookies
- Digital Advertising Alliance (DAA) AdChoices icon and associated opt-out pages
Existing Privacy Tools
Existing Privacy Tools

Privacy Badger detected 45 potential trackers on this page. These sliders let you control how Privacy Badger handles each one. You shouldn't need to adjust them unless something is broken.

- weather.api.cnn.io
- rtax.criteo.com
- ad.doubleclick.net
- googleads.g.doubleclick.net
- securepubads.g.doubleclick.net
- connect.facebook.net

Disable Privacy Badger for This Site
Did Privacy Badger break this site? Let us know!
Donate to EFF

Ghostery found 15 trackers on www.cnn.com, with 14 blocked.

- Advertising: 10 trackers, 10 blocked
  - Amazon Associates
  - ChartBeat
  - Criteo
  - DoubleClick
  - Google-Publisher-Tags
  - Krux-Digital
  - NetRatings-SiteCensus
  - Outbrain
  - Rubicon
  - ShareThrough

- Other: 2 trackers, 2 blocked
  - Site Analytics

Trust Site
Restrict Site
Pause Ghostery
Map These Trackers
Connection Graphs
Browser fingerprinting

- Use features of the browser that are relatively unique to your machine
  - Fonts
  - GPU model anti-aliasing (Canvas fingerprinting)
  - User-agent string
  - *(Often not)* IP address *(Why not?)*
Device Fingerprinting

• Use unique(-ish) combination of device features as an identifier

• [https://panopticlick.eff.org/](https://panopticlick.eff.org/)