CSE 484: Computer Security and Privacy

Web Security
[Web Privacy]

Winter 2021

David Kohlbrenner
dkohlbre@cs.washington.edu

Thanks to Franzi Roesner, Dan Boneh, Dieter Gollmann, Dan Halperin, Yoshi Kohno, Ada Lerner, John Manferdelli, John Mitchell, Vitaly Shmatikov, Bennet Yee, and many others for sample slides and materials
...
Admin

• Lab 2:
  • Check access ASAP
  • Read FAQs 😊

• Homework 3: out soon, due 03/5

• Guest lecture on Wednesday
  • Alex Gantman – Qualcomm – Embedded systems security and careers
A topic in flux

• Tracking via cookies

• Tracking via other methods

• Fingerprinting
Ads That Follow You

Advertisers (and others) track your browsing behaviors for the purposes of targeted ads, website analytics, and personalized content.
Third-Party Web Tracking

Browsing profile for user 123:

- cnn.com
- theonion.com
- adult-site.com
- political-site.com

These ads allow criteo.com to link your visits between sites, even if you never click on the ads.
Concerns About Privacy

The New York Times

‘Do Not Track’ Privacy Bill Appears in Congress
By TANZINA VEGA

And the privacy legislation just keeps on coming.

On Friday, two bills were introduced in Washington in support of a Do Not Track mechanism that would give users control over how much of their data was collected by advertisers and other online companies.

By JENNIFER VALENTINO-DEVRIES, JEREMY SINGER-VINE and ASHKAN SOLTANI
December 24, 2012
First and Third Parties

• **First-party cookie**: belongs to top-level domain.
• **Third-party cookie**: belongs to domain of embedded content (such as image, iframe).

![Diagram of First and Third Parties](image-url)
Anonymous Tracking

Trackers included in other sites use third-party cookies containing unique identifiers to create browsing profiles.
Basic Tracking Mechanisms

• Tracking requires:
  (1) re-identifying a user.
  (2) communicating id + visited site back to tracker.
Tracking Technologies

- HTTP Cookies
- HTTP Auth
- HTTP Etags
- Content cache
- IE userData
- HTML5 protocol and content handlers
- HTML5 storage
- Flash cookies
- Silverlight storage
- TLS session ID & resume
- Browsing history
- window.name
- HTTP STS
- DNS cache
- “Zombie” cookies that respawn
  (http://samy.pl/evercookie)
Other Trackers?

“Personal” Trackers
Personal Tracking

- Tracking is **not anonymous** (linked to accounts).
- Users directly visit tracker’s site $\rightarrow$ evades some defenses.
How prevalent is tracking? (2011)

524 unique trackers on Alexa top 500 websites (homepages + 4 links)

457 domains (91%) embed at least one tracker. (97% of those include at least one cross-site tracker.)

50% of domains embed between 4 and 5 trackers.

One domain includes 43 trackers.
Who/what are the top trackers? (2011)
How has this changed over time?

• The web has existed for a while now...
  - What about tracking before 2011?
  - What about tracking before 2009?

• Solution: time travel!
The Wayback Machine to the Rescue

Time travel for web tracking: http://trackingexcavator.cs.washington.edu
1996-2016: More & More Tracking
- More trackers of more types, more per site, more coverage

Rise And Fall of Historical Champion Trackers

- come.to
- go.com
- v3.com
- doubleclick.net
- allyes.com
- 2o7.net
- google-analytics.com
- google.com
- quantserve.com
- scorecardresearch.com
- gstatic.com
Defenses to Reduce Tracking

• Do Not Track?

Send a ‘Do Not Track’ request with your browsing traffic

Do Not Track is not a technical defense: trackers must honor the request.
Defenses to Reduce Tracking

• Do Not Track proposal?
• Private browsing mode?

Private browsing mode protects against local, not network, attackers.

You’ve gone incognito

Now you can browse privately, and other people who use this device won’t see your activity. However, downloads and bookmarks will be saved. Learn more

Chrome won’t save the following information:
• Your browsing history
• Cookies and site data
• Information entered in forms

Your activity might still be visible to:
• Websites you visit
• Your employer or school
• Your internet service provider
Defenses to Reduce Tracking

- Do Not Track proposal?
- Private browsing mode?
- Third-party cookie blocking?
Its real!

• Safari and FF (mostly) now block 3rd party cookies
  • https://webkit.org/blog/10218/full-third-party-cookie-blocking-and-more/

• Chrome...

  “By undermining the business model of many ad-supported websites, blunt approaches to cookies encourage the use of opaque techniques such as fingerprinting (an invasive workaround to replace cookies), which can actually reduce user privacy and control. We believe that we as a community can, and must, do better.”
Fingerprinting is out there

• Better than a ‘voluntary’ cookie: involuntary, unchangeable id!
  • “Fingerprint”

• Idea: Measure ‘behavior’ of browser
  • Smash into unique ID
Fingerprinting Web Browsers

• User agent
• HTTP ACCEPT headers
• Browser plug-ins
• MIME support
• Clock skew

• Installed fonts
• Cookies enabled?
• Browser add-ons
• Screen resolution
• HTML5 canvas
  (differences in graphics SW/HW!)

2/19/2021
HTML5 Canvas Fingerprinting - Text

Original Image:
How quickly daft jumping zebras vex.

Linux:
How quickly daft jumping zebras vex.
How quickly daft jumping zebras vex.

OSX:
How quickly daft jumping zebras vex.

Windows (XP, Vista, 7):
How quickly daft jumping zebras vex.
How quickly daft jumping zebras vex.

Windows 8:

Figure 7: Difference maps for a group on text_arial

Mowery and Shacham, 2012
HTML5 Canvas Fingerprinting - Image

Figure 10: Original render and difference maps for Group 24

Mowery and Shacham, 2012
And its out there!

Figure 4: Different images printed to canvas by fingerprinting scripts. Note that the phrase “Cwm fjordbank glyphs vext quiz” in the top image is a perfect pangram, that is, it contains all the letters of the English alphabet only once to maximize diversity of the outcomes with the shortest possible string.
Fingerprinting Web Browsers

- User agent
- HTTP ACCEPT headers
- Browser plug-ins
- MIME support
- Clock skew
- Installed fonts
- Cookies enabled?
- Browser add-ons
- Screen resolution
- HTML5 canvas
  (differences in graphics SW/HW!)
Here are your Cover Your Tracks results. They include an overview of how visible you are to trackers, with an index (and glossary) of all the metrics we measure below.

**Our tests indicate that you have strong protection against Web tracking.**

**IS YOUR BROWSER:**

<table>
<thead>
<tr>
<th>Blocking tracking ads?</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blocking invisible trackers?</td>
<td>Yes</td>
</tr>
<tr>
<td>Protecting you from fingerprinting?</td>
<td>Your browser has a nearly-unique fingerprint</td>
</tr>
</tbody>
</table>

Still wondering how fingerprinting works?

[LEARN MORE]

---

One in 145,235 browsers have the same fingerprint
Fingerprinting as a security measure

• Blocking bots (e.g. reCAPTCHA)

• Validating users over-time
How should we view tracking and fingerprinting efforts?
“Privacy preserving” personalized ads

- [https://github.com/WICG/turtledove](https://github.com/WICG/turtledove)
  - The browser, not the advertiser, holds the information about what the advertiser thinks a person is interested in.
  - Advertisers can serve ads based on an interest, but cannot combine that interest with other information about the person — in particular, with who they are or what page they are visiting.
  - Web sites the person visits, and the ad networks those sites use, cannot learn about their visitors' ad interests.