CSE 484 / CSE M 584: Computer Security and Privacy

Web Security
[Web Privacy]

Autumn 2020

Franziska (Franzi) Roesner
franzi@cs.washington.edu

Thanks to 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 12/4

• Guest lecture on Friday
  – Emily McReynolds, law & policy

• No class the day before Thanksgiving
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

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
Outline

1. Understanding web tracking
2. Measuring web tracking
3. Defenses
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).
Anonymous Tracking

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

cookie: id=789
user 789: theonion.com, cnn.com, adult-site.com, ...

referer header: theonion.com

CSE 484 / CSE M 584 - Autumn 2020
Basic Tracking Mechanisms

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

Hypertext Transfer Protocol

GET /pixel/p-3aud4J6uA4Z6Y.gif?labels=InvisibleBox&busty=2710 HTTP/1.1
Host: pixel.quantserve.com
Connection: keep-alive
Accept: image/webp, */*; q=0.8
User-Agent: Mozilla/5.0 (Macintosh; Intel Mac OS X 10_9_2) AppleWebKit/537.36
Referer: http://www.theonion.com
Accept-Encoding: gzip, deflate, sdch
Accept-Language: en-US, en; q=0.8
Cookie: mc=52a65386-f1de1-00ade-0b26e; d=ENkBRgGH44YEA35MMIL74MKiyDs1A2MQ1I1Q
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)
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!)
Your browser fingerprint appears to be unique among the 3,435,834 tested so far.

Only anonymous data will be collected by this site.

A paper reporting the statistical results of this experiment is now available: How Unique Is Your Browser?, Proceedings of the Privacy Enhancing Technologies Symposium (PETS 2010), Springer Lecture Notes in Computer Science.

Learn about Panopticlick and web tracking., The Panopticlick Privacy Policy., Learn about the Electronic Frontier Foundation.
Other Trackers?

“Personal” Trackers
Personal Tracking

- Tracking is **not anonymous** (linked to accounts).
- Users **directly visit tracker’s site** → evades some defenses.
Outline

1. Understanding web tracking
2. Measuring web tracking
3. 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)

Top 20 Cross-Site Trackers on Top 500 Domains

- DoubleClick (Google)
- Facebook
- QuantServe
- Twitter
- Adnxt
- Revsci
- AdThis
- Invitemedia
- Serving-sys
- YouTube

_tracker prevalence (from left to right): 189, 154, 149, 109, 105, 93, 81, 60, 45, 44, 40, 34, 33, 32, 32, 30, 29, 27, 26, 25
How has this changed over time?

- The web has existed for a while now...
  - What about tracking before 2011? (our first study)
  - What about tracking before 2009? (first academic study)

- 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
1996-2016: More & More Tracking

- More trackers of more types, more per site
1996-2016: More & More Tracking

- More trackers of more types, more per site, more coverage
Outline

1. Understanding web tracking
2. Measuring web tracking
3. Defenses
Defenses to Reduce Tracking

• Do Not Track proposal?

Do Not Track is not a technical defense: trackers must honor the request.

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

Diagram:
- www.bar.com (1st party)
- www.foo.com (3rd party)
- Bar’s Server
- Foo’s Server
Quirks of 3rd Party Cookie Blocking

So if a third-party cookie is somehow set, it can be used.

How to get a cookie set?
One way: be a first party.

In some browsers, this option means third-party cookies cannot be set, but they CAN be sent.

etc.
Defenses to Reduce Tracking

- Do Not Track header?
- Private browsing mode?
- Third-party cookie blocking?
- Browser add-ons?

Often rely on blacklists, which may be incomplete.

“uses algorithmic methods to decide what is and isn't tracking”