CSE 484 / CSE M 584: Computer Security and Privacy

Web Security
[Web Privacy]

Spring 2019

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

• Guest lecture on Friday
  – Emily McReynolds from Microsoft, on law+policy

• Rest of quarter overview
  – Final project checkpoint due Friday (May 17)
  – Lab 2 due next Friday (May 24)
  – Homework 3 out soon, due May 31
  – Final project checkpoint 2 also due May 31
  – Lab 3 on smart home security coming up
  – No section in last week; No class on last day
Last Word on Web App Security...
Storing State in Hidden Forms

• Dansie Shopping Cart (2006)
  – “A premium, comprehensive, Perl shopping cart. Increase your web sales by making it easier for your web store customers to order.”

```html
<FORM METHOD=POST
ACTION="http://www.dansie.net/cgi-bin/scripts/cart.pl">
  Black Leather purse with leather straps
  <INPUT TYPE=HIDDEN NAME=name VALUE="Black leather purse">
  <INPUT TYPE=HIDDEN NAME=price VALUE="20.00">
  <INPUT TYPE=HIDDEN NAME=sh VALUE="1">
  <INPUT TYPE=HIDDEN NAME=img VALUE="purse.jpg">
  <INPUT TYPE=HIDDEN NAME=custom1 VALUE="Black leather purse with leather straps">
  <INPUT TYPE=SUBMIT NAME="add" VALUE="Put in Shopping Cart">
</FORM>

Change this to 2.00
Bargain shopping!

Fix: MAC client-side data, or, more likely, keep on server.
Web Privacy
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

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
Outline

1. Understanding web tracking
2. Measuring web tracking
3. Defenses
Recall: 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.
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&bust=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=ENkBRgGHD4GYEA35MMIL74MKiyDs1A2MQI1Q
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.
History Sniffing

How can a webpage figure out which sites you visited previously?

• Color of links
  – CSS :visited property
  – getComputedStyle()

• Cached Web content timing

• DNS timing
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
Measurement Study

• **Questions:**
  – How prevalent is tracking (of different types)?
  – How much of a user’s browsing history is captured?
  – How effective are defenses?

• **Approach:** Build tool to automatically crawl web, detect and categorize trackers based on our taxonomy.

Longitudinal studies since then: tracking has increased and become more complex.
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 are users affected?

• Question: How much of a real user’s browsing history can top trackers capture?

• Measurement challenges:
  – Privacy concerns.
  – Users may not browse realistically while monitored.

• Insight: AOL search logs (released in 2006) represent real user behaviors.
How are users affected?

• Idea: Use AOL search logs to create 30 hypothetical browsing histories.
  – 300 unique queries per user → top search hits.

• Trackers can capture a large fraction:
  – Doubleclick: Avg 39% (Max 66%)
  – Facebook: Avg 23% (Max 45%)
  – Google: Avg 21% (Max 61%)
How are users affected?

- Trackers can capture a large fraction:
  - Doubleclick: Avg 39% (Max 66%)
  - Facebook: Avg 23% (Max 45%)
  - Google: Avg 21% (Max 61%)

NSA reportedly 'piggybacking' on Google advertising cookies to home in on surveillance targets

See also: ADINT (2017)
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!

  [USENIX Security ’16]
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?

 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?

![Diagram showing a browser window with a website, Bar's Server, and Foo's Server. The diagram indicates that Bar's cookie (1st party) is allowed, while Foo's cookie (3rd party) is blocked.](image-url)
Quirks of 3rd Party Cookie Blocking

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

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.

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”