Tracking on the Web

CS 161: Computer Security
Prof. Raluca Ada Popa

February 19, 2016
Announcements

• Wednesday, Feb 24
  – 8-9:30pm (in 155 Dwinelle)
  – One cheat sheet, double sided, hand written
  – Covers material up to Monday

• Review session on Wed in lecture
  • Sample example questions, review material before

• Homework 2 (due Feb 22)
What does a site learn about you when you visit them?

Discuss with your neighbor
The sites you visit learn:

• The **URLs** you’re interested in
  – Google/Bing also learns *what you’re searching for*
• Your **IP address**
  – Thus, your service provider & geo-location
  – Can often link you to other activity including at other sites
• Your browser’s capabilities, which OS you run, which language you prefer
• Which URL you looked at that took you there
  – Via the HTTP “**Referer**” header

They also learn cookies!
They also learn cookies

Why is that harmful?
Let's remove all of our cookies.
Cool, no web site is tracking us …
We do a search on "private browsing"
Private Browsing - Browse the web without saving information about ...
support.mozilla.org/.../private-browsing-browse-web-without-saving-inf... ▼
When using a shared computer, Private Browsing is great for viewing websites without
saving stuff like cookies, temp files and a history of the pages you visit.

Firefox 20 Launches With Improved Private Browsing, New ...

Privacy mode - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/Privacy_mode ▼
Internet Explorer 8 in InPrivate mode. Google Chrome in Incognito mode. Privacy mode
or "private browsing", sometimes informally referred to as "porn mode", ...

Firefox 20 improves private browsing, fixes three critical flaws | ZDNet
www.zdnet.com/firefox-20-improves-private-browsing-fixes-three-critic... ▼
Apr 3, 2013 – Mozilla has released the latest version of its Firefox web browser which
features new enhancement to private browsing and fixes a number of ...

Private Browsing - Web Browsers - About.com
browsers.about.com › ... › Web Browsers › Web Browser Glossary › FAQs ▼
The methods for activating private browsing mode differ across browsers, operating
systems, and device types. These step-by-step tutorials teach you how to ...
Google has stored a couple of cookies on our system.
Goodness knows what info they decided to put in the cookie.
But it lasts for months ...
Private browsing

You can turn on a mode called private browsing on your browser

What is this?
Does it protect you against tracking?
We click on the top result.
Note that this mode is privacy from your family, not from web sites!
Private browsing

“Private Browsing allows you to browse the Internet without saving any information about which sites and pages you’ve visited.”

- deletes history of URL visits, passwords, cookies too
- Private Browsing maintains cookies for as long as the private browsing window is open. Once you quit the browser, it gets deleted
- So still tracked for a good while!
Ironically, we’ve gained a bunch of cookies in the process.
This one sticks around for two years.

Expires: April 17, 2018
How did YouTube enter the picture??
YouTube is remembering the version of Flash I’m running…

Expires: April 17, 2018
We navigate to *The New York Times* …
U.S. Announces More Sanctions Against Russia Over Ukraine

By PETER BAKER and MARK LANDLER

The United States ordered travel bans and asset freezes for seven Russian officials, including two said to be in President Vladimir V. Putin's inner circle, and froze assets for 17 firms.

284 Comments
- Mayor of Eastern Ukraine City Is Shot
- Putin Rival Takes Message to East Ukraine

Egypt Sentences More Than 680 to Death

The Muslim Brotherhood's spiritual leader and hundreds of others were sentenced on charges of inciting or committing violence. Supporters, above, reacted to the verdict Monday.

130 Comments

Chernobyl: Capping a Nuclear Catastrophe

The president of Ukraine has hinted that the site of the Chernobyl nuclear disaster could one day be a cultural center, but a site visit suggests that a long road lies ahead for the people and the city.

2,264 Comments

Made with IBM

The Opinion Pages

EDITORIAL
Political Executions in Egypt

It is clear from the sentencing of 680 people to death in a mass trial that the country's judges have become a government tool.

- Editorial: Smartphones and the 4th Amendment
- Krugman: High Plains Moochers

THE STONE
What Does Buddhism Require?

The reality of rebirth may not be necessary. But believing in it probably is.

- Gessen: Salon of the Exiled
- Op-Ed: The Wire Next Time
- Op-Docs | ‘Verbatim: What Is a Photocopier?’

Today's Times Insider

Behind the scenes of The New York Times

- Thinking of Wine as Food With Eric Asimov
- Introducing Times Insider

MARKETS »

S&P 500
Dow
Nasdaq

The New York Times

HOME DELIVERY

Register
What a lot of yummy cookies!
Here are the ones from the website itself …

<table>
<thead>
<tr>
<th>Site</th>
<th>Cookie Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>nytimes.com</td>
<td>RMID</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>nyt5_disable</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>_dyid</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>_dyfs</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>_cb_ls</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>nytnow3p</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>kxtag28172.day</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>nyt-m</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>nyt-recommend</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>adxcl</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>adxs</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>tagx-l</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>tagx-s</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>tagx-p</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>WT_FPC</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>_dyaud_page</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>_dysvar_8765260</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>_dyuss_8765260</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>_dyfst</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>_dy_geo</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>_dyaud_nchc</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>_dyaud_sess</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>_dyus_8765260</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>rsi_segs</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>kxtag27935.day</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>kxtag27728.day</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>kxtag21418.day</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>kxtag22998.day</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>kxtag21233.day</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>kxtag28173.day</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>_chartbeat?</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>_chartbeat_uuniq</td>
</tr>
</tbody>
</table>
This one tracks the details of my system & browser

<table>
<thead>
<tr>
<th>Site</th>
<th>Cookie Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>nytimes.com</td>
<td>_dyus_8765260</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>rsi_segs</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>kxtag27935.day</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>kxtag27728.day</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>kxtag15486.day</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>kxtag21418.day</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>kxtag22998.day</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>kxtag21233.day</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>_chartbeat2</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>_chartbeat_uuniq</td>
</tr>
<tr>
<td>nytimes.com</td>
<td>kxtech</td>
</tr>
<tr>
<td>nvtimes.com</td>
<td>kxsegs</td>
</tr>
<tr>
<td>nvtimes.com</td>
<td>krux_segs</td>
</tr>
</tbody>
</table>

Name: kxtech
Content: device%3DComputer%26manufacturer%3DApple%2520Inc.%26os%3DMac%2520OS%2520X%26browser%3DFirefox%25202
Host: www.nytimes.com
Path: /
doubleclick.net - who's that? And how did it get there from visiting www.nytimes.com?
Third-Party Cookies

• How can a web site enable a third party to plant cookies in your browser & later retrieve them?
  – Include on the site’s page (for example):
    • `<img src="http://doubleclick.net/ad.gif" width=1 height=1>`

• Why would a site do that?
  – Site has a business relationship w/ DoubleClick

• Why can this track you?
  – Now DoubleClick sees all of your activity that involves their web sites
  – Because your browser dutifully sends them their cookies for any web page that has that img
  – Identifier in cookie ties together activity as = YOU

* Owned by Google, by the way
Remember this 2-year Mozilla cookie?
Google Analytics

• Any web site can (anonymously) register with Google to instrument their site for analytics
  – Gather information about who visits, what they do when they visit
• To do so, site adds a small Javascript snippet that loads http://www.google-analytics.com/ga.js
  – You can see sites that do this because they introduce a "__utma" cookie
• Code ships off to Google information associated with your visit to the web site
  – Shipped by fetching a GIF w/ values encoded in URL
  – Web site can use it to analyze their ad “campaigns”
  – Not a small amount of info …
Values Reportable via Google Analytics

Affiliation
Billing City
Billing Country
Billing Region
Browser Lang.
Complete URL
Cookie Values
Current Page
Event Tracking
Flash Version
Grand Total
Host Name
Java-enabled
Language Encoding
Order ID
Page Title
Product Code
Product Name
Profile Number
Repeat Campaign Visit
Quantity
Screen Color Depth
Screen Resolution
Shipping Cost
Special Event
Start Campaign Sess.
Tax
Tracking Code Version
Unique GIF ID
Unit Price
User Defined Var
Variations on an Item
Still More Tracking Techniques …

• Any scenario where browsers execute programs that manage persistent state can support tracking by cookies
  – Such as …. Flash?
My browser had Flash cookies from 67 sites!

Sure, this is where you’d think to look to analyze what Flash cookies are stored on your machine.

Some Flash cookies “respawn” regular browser cookies that you previously deleted!
Internet Explorer bug lets hacker control your PC

By Jose Pagliery @Jose_Pagliery April 28, 2014: 2:27 PM ET

Facebook “Like” button (an IFRAME hosted on facebook.com)

A new bug in Internet Explorer allows hackers to commandeer your computer.

NEW YORK (CNNMoney)
If you’re using Internet Explorer and click on the wrong link, a hacker could hijack your computer.
What does Facebook learn?

• Many pages include a Facebook “Like” button.
• What are the implications, for user tracking?

• Facebook can track you on every site that you visit that embeds such a button
What information does Facebook get when I visit a site with the Like button?

If you’re logged into Facebook and visit a website with the Like button, your browser sends us information about your visit. Since the Like button is a little piece of Facebook embedded on another website, the browser is sending info about the request to load Facebook content on that page.

We record some of this info to help show you a personalized experience on that site and to improve our products. For example, when you go to a website with a Like button, we need to know who you are in order to show you what your Facebook friends have liked on that site. The data we receive includes your user ID, the website you’re visiting, the date and time and other browser-related info.
Cookies form the core of how Internet advertising works today

- Without them, arguably you’d have to pay for content up front a lot more
  - (and payment would mean you’d lose anonymity anyway)
- A “better ad experience” is not necessarily bad
  - Ads that reflect your interests; not seeing repeated ads

But: ease of gathering so much data so easily \(\rightarrow\) concern of losing control how it’s used

- Privacy concerns
- Large amounts of private data in one place
More Employers Screening Candidates via Social Networking Sites

Five tips for creating a positive online image
Rosemary Haefner, Vice President of Human Resources at CareerBuilder

Gone are the days when all job seekers had to worry about were their résumés and cover letters. Today, those documents remain a staple of the job-search process, but they are joined by a growing phenomenon: social networking.

Forty-five percent of employers reported in a June 2009 CareerBuilder survey that they use social networking sites to screen potential employees, compared to only 22 percent of employers last year. Eleven percent of employers plan to start using social networking sites for the screening process. More than 2,600 hiring managers participated in the survey.
Why employers disregard candidates after screening online

Thirty-five percent of employers reported they have found content on social networking sites that caused them not to hire the candidate, including:

- Candidate posted provocative or inappropriate photographs or information -- 53 percent
- Candidate posted content about them drinking or using drugs -- 44 percent
- Candidate bad-mouthed their previous employer, co-workers or clients -- 35 percent
- Candidate showed poor communication skills -- 29 percent
- Candidate made discriminatory comments -- 26 percent
- Candidate lied about qualifications -- 24 percent
- Candidate shared confidential information from previous employer -- 20 percent
Tracking – So What?

• Cookies etc. form the core of how Internet advertising works today
  – Without them, arguably you’d have to pay for content up front a lot more
    • (and payment would mean you’d lose anonymity anyway)
  – A “better ad experience” is not necessarily bad
    • Ads that reflect your interests; not seeing repeated ads

• But: ease of gathering so much data so easily → concern of losing control how it’s used
  – Content shared with friends doesn’t just stay with friends …
  – You really don’t have a good sense of just what you’re giving away …
My baby girl.... http://t.co/5qLfLV6

2 minutes ago via Twitter for Android

BritBangert
Brittany Bangert
I Can Stalk U

Raising awareness about inadvertent information sharing

Who have we stalked recently?

ICanStalkU was able to stalk RangeLifeEnt at 51 Great Jones St New York NY
1 minute ago · Map Location · View Tweet · View Picture · Reply to RangeLifeEnt

ICanStalkU was able to stalk Inicklasson at http://maps.google.com/?q=57.1344444444,12.7141666667
2 minutes ago · Map Location · View Tweet · View Picture · Reply to Inicklasson

ICanStalkU was able to stalk Welerson13 at http://maps.google.com/?q=-15.7380555556,-47.8986111111
2 minutes ago · Map Location · View Tweet · View Picture · Reply to Welerson13

ICanStalkU was able to stalk BritBangert at 920 Hawley St Taylorville IL
1 minute ago · Map Location · View Tweet · View Picture · Reply to BritBangert

ICanStalkU was able to stalk jiggy_Owl at http://maps.google.com/?q=13.7830055879,100.518500685
4 minutes ago · Map Location · View Tweet · View Picture · Reply to jiggy_Owl

ICanStalkU was able to stalk gcolony at http://maps.google.com/?q=37.7851666667,-122.404166667
4 minutes ago · Map Location · View Tweet · View Picture · Reply to gcolony

Links

- Mayhemic Labs
- PaulDotCom
- SANS ISC
- Electronic Frontier Foundation
- Center for Democracy & Technology

How did you find me?

Did you know that a lot of smart phones encode the location of where pictures are taken? Anyone who has a copy can access this information.

read more

Help me fix this!

Disabling Geo-Tagging on your phone is easy. Check your phone's settings for more information.
How To Gain Better Privacy?

discuss with your neighbor
How To Gain Better Privacy?

• Force of law
  – Example #1: web site privacy policies
    • US sites that violate them commit false advertising
    • But: policy might be “Yep, we sell everything about you, Ha Ha!”
7. Collection of Viewing Information. You acknowledge that you are aware of and consent to the collection of your viewing information during your use of the Software and/or Content. Viewing information may include, without limitation, the time spent viewing specific pages, the order in which pages are viewed, the time of day pages are accessed, IP address and user ID. This viewing information may be linked to personally identifiable information, such as name or address and shared with third parties.
How To Gain Better Privacy?

• Force of law
  – Example #1: web site privacy policies
    • US sites that violate them commit false advertising
    • But: policy might be “Yep, we sell everything about you, Ha Ha!”
  – Example #2: SB 1386
    • Requires an agency, person or business that conducts business in California and owns or licenses computerized 'personal information' to disclose any breach of security (to any resident whose unencrypted data is believed to have been disclosed)
    • Quite effective at getting sites to pay attention to securing personal information
May 8, 2009 1:53 PM PDT

UC Berkeley computers hacked, 160,000 at risk

by Michelle Meyers

This post was updated at 2:16 p.m. PDT with comment from an outside database security software vendor.

Hackers broke into the University of California at Berkeley's health services center computer and potentially stole the personal information of more than 160,000 students, alumni, and others, the university announced Friday.

At particular risk of identity theft are some 97,000 individuals whose Social Security numbers were accessed in the breach, but it's still unclear whether hackers were able to match up those SSNs with individual names, Shelton Waqgner, UCB's chief technology officer, said in a press conference Friday afternoon.
How To Gain Better Privacy?

• Technology
  – Various browser additions
  – Special browser extensions
  – Tor and anonymizers – later in course
Browser: “Tracking protection”

Private browsing includes tracking protection

Blocks third-party trackers based on Disconnect.me

- **basic**: blocks commonly known analytics trackers, social sharing trackers, and advertising trackers, but allows some known content trackers to reduce website breakage.

- **strict**: blocks all known trackers, including analytics, trackers, social sharing trackers, and advertising trackers as well as content trackers. The strict list will break some videos, photo slideshows, and some social networks.
Browsers: Do not track flag

You can turn on this flag in your browser

What does it do?
- Tells web servers you want to opt-out of tracking
- It does this by transmitting a Do Not Track HTTP header every time your data is requested from a web server

It does not enforce that there is no tracking, it is up to the web servers
Some ad companies do provide more generic ads than
Browser extension: Ghostery

User installs browser extension:

1. Recognizes third-party tracking scripts on a web page based on an actively curated database of such scripts

2. Blocks HTTP requests to these sites
   • as a result, Facebook buttons don’t even show

3. Users can create “Whitelists” of allowed sites
   • e.g., allow FB button but note that you allow tracking by FB too
But you have to be careful…

• Users can opt-in to sending anonymously data back to Evidon, the parent company, to improve its tracking database
• Evidon sells this data to ad companies..
• But strategy is transparent, users opt into this
Conclusions

• Third-party apps can track us even if we don’t visit their website
• Tracking is very common on the web and can collect a lot of data about you
• Some solutions exist, but have caveats