CSE508 Network Security

5/2/2016 Privacy

Michalis Polychronakis
Stony Brook University
Privacy

“The right of an entity (normally a person), acting in its own behalf, to determine the degree to which it will interact with its environment, including the degree to which the entity is willing to share information about itself with others.” [RFC2828]
Real-world Privacy

Large-scale data collection examples
  Credit cards, Metrocards, Loyalty cards
  Street/public space cameras
  E-ZPass
  Named tickets
  …

Part of our everyday activities and personal information is (voluntarily or compulsorily) recorded
  Information from different sources can be correlated
  Did you buy your Metrocard with your credit card?

The same happens in the online world…
Third parties have access to...
Our email (Gmail, Yahoo, …)
Our files (Dropbox, Google Drive, …)
Our finances (e-banking, credit reporting, Mint, …)
Our communication (Skype, Facebook, …)
Our traffic (WiFi hotspots, ISPs, …)
Our location (3/4G, GPS, WiFi, …)
Our activities (browsing history, daily routine, …)
Our preferences (“Likes,” Amazon, Netflix, …)
Our health (Fitbit, iWatch, 23andMe, …)
…
Millions of Anthem Customers Targeted in Cyberattack

By REED ABELSON and MATTHEW GOLSTEIN  FEB. 5, 2015

Outside the Anthem facility in Indianapolis. Anthem said it detected a data breach on Jan. 29, and that it was working with the Federal Bureau of Investigation. Aaron P. Bernstein/Getty Images

Anthem, one of the nation’s largest health insurers, said late
WASHINGTON — The Obama administration on Thursday revealed that 21.5 million people were swept up in a colossal breach of government computer systems that was far more damaging than originally believed.

Katherine Archuleta, director of the Office of Personnel Management, right, at hearing before the House Oversight and Government Reform Committee last month. Mark Wilson/Getty Images
World's Biggest Data Breaches
Selected losses greater than 30,000 records
(updated 2nd October 2015)
Apple Says It Will Add New iCloud Security Measures After Celebrity Hack

By BRIAN X. CHEN  SEPTEMBER 4, 2014 11:32 PM  21 Comments
Armed With Facebook 'Likes' Alone, Researchers Can Tell Your Race, Gender, and Sexual Orientation

REBECCA J. ROSEN | MAR 12, 2013, 2:59 PM ET

But the deeper aspects of your personality remain hard to detect.
Facebook knows you’re gay before you do

3/20/13 4:29pm by Jon Green  39 Comments

Am I the only one creeped out that Facebook is now guessing, sometimes correctly, if its users are gay?

In the world of Big Data, our private lives are increasingly becoming intermingled with the shadowy, yet public, world of cyberspace.

Whenever we go online we are providing data that can be used to market to us; from Google searches to Facebook likes to eBay purchases, we are inputting data into a series of mathematical models which make incredibly educated guesses about the kinds of people we are.

Facebook creepily offers help to a gay guy thinking of “coming out”

Enter Matt. As BuzzFeed notes, Matt was your typical Facebook user who suddenly found an ad in his news feed for help in coming out. The weird thing was that Matt “did” need help coming out, and understandably he was more than a bit curious as to how Facebook knew.

At first, Matt wondered if Facebook had accessed his text messages, as he had confided in a close friend the previous
Facebook users unwittingly revealing intimate secrets, study finds

Personal information including sexuality and drug use can be correctly inferred from public 'like' updates, according to study
How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did

Every time you go shopping, you share intimate details about your consumption patterns with retailers. And many of those retailers are studying those details to figure out what you like, what you need, and which coupons are most likely to make you happy. Target, for example, has figured out how to data-mine its way into your womb, to figure out whether you have a baby on the way long before you need to start buying diapers.

Charles Duhigg outlines in the New York Times how Target tries to hook parents-to-be at that crucial moment before they turn into rampant — and
Web Browsing Tracking

Webpages are often mashups of content loaded from different sources

- Ads, images, videos, widgets, …
- IMG URLs, IFRAMEs, JavaScript, web fonts, Flash/applets, …
- Hosted on third-party servers: CDNs, cloud providers, ad networks, …

A third party involved in many different websites can track user visits across all those websites

- 2+ third parties may collude to expand their collective “view”

Need to learn two key pieces of information

*What webpage was visited*

*Who visited it*
Microsoft Announces Continuum, Turning Windows 10 Phones Into Desktops

Microsoft just demonstrated one of the intriguing possibilities from its single platform/multiple-device ecosystem: turning a Windows 10 phone into a desktop computer.

The company demonstrated a Windows 10 phone connected to a monitor, keyboard, and mouse, effectively turning it into a desktop computer. The demonstration was presented at the Build developers conference in San Francisco.

The Continuum feature allows users to switch between phone and desktop modes without having to restart the device. In desktop mode, the phone runs the full Windows 10 operating system, providing access to all the apps and services available on the desktop.

Microsoft has been working on Continuum for several months, and the feature is expected to be available in the Windows 10 update slated for later this year. The company has also been working on other features, such as the Continuum for Tablets mode, which allows users to use a tablet computer as a desktop device.

The demonstration at Build showed how Continuum works in practice, with users being able to type on a keyboard, use a mouse, and run desktop apps on the phone. Microsoft has also been working on improving the battery life of Windows 10 phones, which is a key consideration for users who want to use them as desktop devices.

Overall, Continuum represents a significant step forward for Microsoft, as it allows the company to leverage its existing desktop ecosystem and attract more users to its platform.

---

**Disconnect**

Browse the web normally. As you do, the graph in this popup and the counter in the toolbar will update. Each circle in the graph represents a site that’s been or would’ve been sent some of your personal info.

Circles with a halo are sites you’ve visited. Circles without a halo are sites you haven’t.

Red circles are known tracking sites. Gray circles aren’t but may still track you.

Mouse over a circle to view that site’s tracking footprint. Click a red circle to block or unblock that site.

- Unblock tracking sites
- Hide sidebar
What webpage was visited?

HTTP Referer [sic] header

The URL of the webpage from which a link was followed
Useful for statistics/analytics, bad for privacy
Can be turned off through browser options/extensions
HTML5 rel="noreferrer" anchor attribute to indicate to the user agent not to send a referrer when following the link

Page-specific, session-specific, user-specific URLs

Unique URL per page (even for the same resource) ➔ track what page was visited
Unique URL per session/user ➔ distinguish between visits from different users
Tracking URLs are also commonly used in promotional emails

**Embedded image loading**
This is an active email address!
Detect the time a user viewed a message
The request reveals much more: user agent, device, location, ...

**Embedded links**
Learn which email addresses resulted in visits (click-through rate)

Default behavior of email clients varies
Gmail used to block images by default, now uses image proxy servers
Tracking through unique images still possible: senders can track the first time a message is opened (user’s IP is not exposed though)
Who visited the page?

Browsing to a web page reveals a wealth of information

Source IP address

Not very accurate (e.g., NAT, DHCP, on-the-go users) but still useful

Third-party cookies: precise user tracking

Easy to block (configurable in most browsers, defaults vary)

“Evercookies:” exploit alternative browser state mechanisms

Flash/Silverlight/other plugin-specific storage, ETags, HTML5 session/local/global storage, caches, …

Browser/device fingerprinting: recognize unique system characteristics

Browser user agent, capabilities, plugins, system fonts, screen resolution, time zone, and numerous other properties
What do web tracking techniques really track?

Distinguish between different visitors

Track anonymous individuals

Actually: track the pages visited by a particular browser running on a particular device

Better: distinguish between different persons

Track named individuals

The transition is easy…

Personally identifiable information (PII) is often voluntarily provided to websites:

Social networks, cloud services, web sites requiring user registration, …

Cookies/sessions are associated with PII

Contamination: trackers may collude with services

Previously “anonymous” cookies/fingerprints can be associated with named individuals
VERIZON’S ‘PERMA-COOKIE’ IS A PRIVACY-KILLING MACHINE
AT&T charges $29 more for gigabit fiber that doesn’t watch your Web browsing

AT&T goes head to head against Google in KC on fiber and targeted ads.

by Jon Brodkin - Feb 16, 2015 12:38pm EST

AT&T's gigabit fiber-to-the-home service has just arrived in Kansas City, and the price is the same as Google Fiber—if you let AT&T track your Web browsing history.
Users register on trackers!

Social plugins are prevalent

1.65+ billion Facebook users
33% of the top 10K websites have Like Buttons
Twitter, Google+, LinkedIn, Pinterest, AddThis, …
OS/app integration

A growing part of our browsing history can be tracked

Not as merely anonymous visitors, but as named persons
Just visiting the page is enough (no interaction needed)
Cross-device tracking
Existing Solutions

Log out
  Some cookies persist

Block third-party cookies
  Not always effective

Block social widgets completely

Incognito mode

All existing solutions disable content personalization
  Privacy vs. functionality dilemma
Single Sign-on/Social Login

Pros

Convenience – fewer passwords to remember
Rich experience through social features
Outsource user registration and management

Cons

Same credentials for multiple sites
User tracking
Access to user’s profile
surfingneighbors.com is requesting permission to do the following:

Access my basic information
Includes name, profile picture, gender, networks, user ID, list of friends, and any other information I've made public.

Send me email
surfingneighbors.com may email me directly at diego.ridaz@yahoo.com. Change

Post to Facebook as me
surfingneighbors.com may post status messages, notes, photos, and videos on my behalf.

Access posts in my News Feed

Access my data any time
surfingneighbors.com may access my data for application.

Access my profile information
Birthday and Facebook Status

Logged in as Diego Ridaz • Log Out

Take it or leave it

Allow  Don't Allow
Location Tracking

IP addresses reveal approximate location information

MaxMind statistics: 99.8% accurate on a country level, 90% accurate on a state level in the US, and 81% accurate for cities in the US within a 50 kilometer radius

Mobile devices allow for precise location tracking

Cell tower triangulation/trilateration
GPS
WiFi access points in known locations

Per-app permissions

Android vs. iOS:
installation vs. usage time
Attention, Shoppers: Store Is Tracking Your Cell

By STEPHANIE CLIFFORD and QUENTIN HARDY  JULY 14, 2013

Like dozens of other brick-and-mortar retailers, Nordstrom wanted to learn more about its customers — how many came through the doors, how many were repeat visitors — the kind of information that e-commerce sites like Amazon have in spades. So last fall the company started testing new technology that allowed it to track customers’ movements by following the Wi-Fi signals from their smartphones.

But when Nordstrom posted a sign telling customers it was tracking them, shoppers were unnerved.

“We did hear some complaints,” said Tara Darrow, a spokeswoman for the store. Nordstrom ended the experiment in May, she said, in part because of the comments.

Nordstrom’s experiment is part of a movement by retailers to gather data about in-store shoppers’ behavior and moods, using video surveillance and signals from their cellphones and apps to learn.
Online Behavioral Tracking

An increasing part of our daily activities are recorded
  What we are interested in  (Searches, Likes, …)
  What we read  (News, magazines, blogs, …)
  What we buy  (Amazon, Freshdirect, …)
  What we watch  (Netflix, Hulu, …)
  What we eat  (Seamless, GrubHub, …)
  Where we eat  (OpenTable, Foursquare, …)
  Where we go  (online travel/hotel/event booking)
  What we own/owe  (e-banking, credit services, Mint, …)

Mobile apps make behavioral tracking easier and more accurate

Behavioral profiles have desirable and not so desirable uses
  Recommendations, content personalization, …
  Targeted advertising, price discrimination (e.g., insurance premiums based on past behavior, higher prices for high-end device users), …
Health and Activity

Health records
  How securely are they handled and stored?

Devices track our activities and health
  Activity tracking devices
  Health monitoring devices
  Mobile phones

Many upload all data to the “cloud”…
  Who can access them?