What information do we leak through our online activities?

Claudia Diaz
K.U.Leuven ESAT/SCD-COSIC
May 30 2011
“offline world” vs “online world”

face-to-face conversation
letters in the post
Knowing who your friends are
Paying with cash

Papers in physical archives

Following your movements
Learning your shopping profile

- Easy/cheap to collect, store, search, and process
- Easy to copy/disseminate, but hard to destroy
- Easy to aggregate, make profiles and inferences
- Information never forgotten
When you browse to a website

“Privacy Diffusion on the Web: A Longitudinal Perspective” B. Krishnamurthy and C. Wills
Study of 1200 popular Web sites
NHS.uk allowing Google, Facebook, and others to track you

The NHS is allowing Google, Facebook, and others to track your http://www.nhs.uk/ browsing habits, regardless of the fact that people use the page to seek medical advice. It was recently pointed out to me that the NHS Choices website’s social features include the Facebook Like button (see e.g. the page on Testicular Cancer). Due to the fact that the standard method of Facebook Like button deployment is intrusive to say the least, I thought I would look into identifying which third party companies have been given permission to track users on NHS Choices, and my results are rather disconcerting.

In short there are four third-party, advertising/ tracking companies which are informed every time a user visits one of the “conditions pages” on the NHS Choices website. These listed below, all get to make a call from the user’s browser, in turn allowing the four companies to access their cookies, tracking the users (explained in a previous blog post of mine, and in Bala’s research). This means, that if one has ever logged into a Google account, or a Facebook account and then visits one of the pages on the NHS site, the company will then know that their user X was just looking at a page about condition Y on the NHS website.
Hidden inside Ashley Hayes-Beaty's computer, a tiny file helps gather personal details about her, all to be put up for sale for a tenth of a penny.

The file consists of a single code—4c812db292229295e5416a323e79bd37—that secretly identifies her as a 28-year-old female in Nashville, Tenn.

When a user like you logs onto the internet...

The code knows that her favorite movies include "The Princess Bride," "50 First Dates" and "10 Things I Hate About You." It knows she enjoys the "Sex and the City" series. It knows she browses entertainment news and likes to take quizzes.

"Well, I like to think I have some mystery left to me, but apparently not!" Ms. Hayes-Beaty said when told what that snippet of code reveals about her. "The profile is eerily correct."

Ms. Hayes-Beaty is being monitored by Lotame Solutions Inc., a New York company that uses sophisticated software called a "beacon" to capture what people are typing on a website—their comments on movies, say, or their interest in expecting and pregnancy. Lotame packages that data into profiles about individuals, without determining a person's name, and sells the profiles to companies seeking customers. Ms. Hayes-Beaty's tastes can be sold wholesale (a batch of movie lovers is $1 per thousand) or customized (26-year-old Southern fans of "50 First Dates").
By EMILY STEEL and JULIA ANGWIN
(Please see Corrections & Amplifications item below)

You may not know a company called [x+1] Inc., but it may well know a lot about you.

From a single click on a web site, [x+1] correctly identified Carrie Isaac as a young Colorado Springs parent who lives on about $50,000 a year, shops at Wal-Mart and rents kids’ videos. The company deduced that Paul Bouliffard, a Nashville architect, is childless, likes to travel and buys used cars. And [x+1] determined that Thomas Eurney, a Colorado building contractor, is a skier with a college degree and looks like he has good credit.

The company didn’t get every detail correct. But its ability to make snap assessments of individuals is accurate enough that Capital One Financial Corp. uses [x+1]’s calculations to instantly decide which credit cards to show first-time visitors to its website.

In short: Websites are gaining the ability to decide whether or not you’d be a good customer, before you tell them a single thing about yourself.
Life insurers are testing an intensely personal new use for the vast dossiers of data being amassed about Americans: predicting people's longevity.

Insurers have long used blood and urine tests to assess people's health—a costly process. Today, however, data-gathering companies have such extensive files on most U.S. consumers—online shopping details, catalog purchases, magazine subscriptions, leisure activities and information from social-networking sites—that some insurers are exploring whether data can reveal nearly as much about a person as a lab analysis of their bodily fluids.
Panopticlick

How Unique — and Trackable — Is Your Browser?

Is your browser configuration rare or unique? If so, web sites may be able to track you, even if you limit or disable cookies.

Panopticlick tests your browser to see how unique it is based on the information it will share with sites it visits. Click below and you will be given a uniqueness score, letting you see how easily identifiable you might be as you surf the web.

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.
Your Apps Are Watching You

A WSJ Investigation finds that iPhone and Android apps are breaching the privacy of smartphone users

By SCOTT THURM and YUKARI IWATANI KANE

What we found on one app

The iPhone version of music app Pandora sent information to eight trackers. It sent location data to seven of these, a unique phone ID to three and demographic data to two.

Few devices know more personal details about people than the smartphones in their pockets: phone numbers, current location, often the owner’s real name—even a unique ID number that can never be changed or turned off.

WSJ’s Julia Angwin explains to Simon Constable how smartphone apps collect and broadcast data about your habits. Many don’t have privacy policies and there isn’t much you can do about it.

These phones don’t keep secrets. They are sharing this personal data widely and regularly, a Wall Street Journal investigation has found.

An examination of 101 popular smartphone “apps”—games and other software applications for iPhone and Android phones—showed that 56 transmitted the phone’s unique device ID to other companies without users’ awareness or consent. Forty-seven apps transmitted the phone’s location in some way. Five sent age, gender and other personal details to outsiders.
Gaydar Algorithm Outs Facebook Users

By Susannah F. Locke  Posted 09.21.2009 at 12:27 pm  9 Comments

What are your friends saying about you? Online social networks like this Facebook one might reveal more about you than you think [Jurvetson (CC licensed)]

A pair of MIT students claim that they have created an algorithm that outs gay members of Facebook by analyzing the sexual orientations of their networks of friends.
Beware online “filter bubbles”: Eli Pariser on TED.com

As web companies strive to tailor their services (including news and search results) to our personal tastes, there’s a dangerous unintended consequence: We get trapped in a “filter bubble” and don’t get exposed to information that could challenge or broaden our worldview. Eli Pariser argues powerfully that this will ultimately prove to be bad for us and bad for democracy. (Recorded at TED2011, March 2011, in Long Beach, CA. Duration: 9:05)
Conclusions

- Bottom line: our actions and interactions are increasingly mediated by technology
  - We leave digital traces everywhere
- We have no control over
  - who can access our information
  - how it is combined and which inferences are derived from it
  - how information is interpreted, how decisions are made, and how these decisions may affect us: price discrimination, access to services, access to information, etc.
- Not just about our name or address, but our interests, thoughts, lifestyle, ...
Thank you!

http://homes.esat.kuleuven.be/~cdiaz/

claudia.diaz@esat.kuleuven.be