Technologies et Protection de votre vie privée sur Internet

Nataliia Bielova

joint work with Claude Castelluccia, Imane Fouad, Gabor Gulyas, Arnaud Legout, Tamara Rezk, Natasa Sarafijanovic-Djukic, Doliere Francis Some

nataliia.bielova@inria.fr
@nataliabielova

13 novembre 2017
Who am I?

• Nataliia Bielova, PhD 2011

• **Research scientist (CR)** at **INDES** team
  French National Institute in computer science and automation

• Interests:
  – Privacy protection
  – Web Tracking technologies
  – Discrimination on the Web
  – ePrivacy Regulation (EU privacy law)
Joint work with my colleagues

Pixel tracking measurement

Browser extensions and Web logins experiment

Server-side protection for Web developers
Back in 1993...

“On the Internet, nobody knows you’re a dog.”
Today...

It’s the Internet! Of course they know you’re a dog. They also know your favorite brand of pet food and the name of the cute poodle at the park that you have a crush on!
Lit mezzanine Sacha

DESCRIPTION DU LIT MEZZANINE ENFANT SACHA :

- Un lit mezzanine
- Structure en bois
- Taille du couchage : 90*200
- Compartiment inclus : le sommier, un bureau, un escalier, un 2ème lit, une bibliothèque.
- Les lits superposés conviennent aux enfants de moins de 6 ans. Hauteur de nos matelas : 16 cm.

AVIS DE NOS CLIENTS

8.4/10

8 étoiles

MÉTÉO PARIS - 75000 - PRÉVISION À 15 JOURS

METEO > France > Île-de-France > Paris > Météo Paris - 75000

PARTAGEZ À PARIS

66 € 64 € 41 € 64 € 35 € 55 € 50 € 35 € 38 € 67 € 50 € 59 € 60 €
Lit mezzanine Sacha

Lit mezzanine, voici un meuble qui regroupera tout ce dont votre enfant aura besoin. Le lit mezzanine enfant Sacha vous propose en plus de son couchage accessible par un escalier, un espace bureau, mais aussi un deuxième lit encastré dans la tête d'une bibliothèque qui lui permettra de ranger les livres et déposer les objets précieux. Ce lit mezzanine aux formes arrondies vous permettra un gain de place.

Matelas :
Sans matelas

Disponibilité : 8 À 12 SEMAINES

1 390 €
3 x 463.33 €
SANS FRAIS + d'infos
dont 450 € d'éco-participation

138,00 € sur 10 mois

Commandez par téléphone :
05 63 61 57 48

Rubiconproject.com
googletagservices.com
pixel.adsafeprotected.com
effectivemeasure.net
doubleclick.net
googlesyndication.com
yahoo.com
sync.mathtag.com
mobilienitro.com
Web Tracking

doubleclick.net

likes

mobiliernitro.com
Web Tracking

doubleclick.net

likes

paris.lachainmeteo.com
How many trackers per website?
How many trackers per website?

18 companies will know that I visited this website!
How many trackers per website?

Pourquoi sommes-nous suivis sur les sites d'actualités?

“L'activité principale du plaignant consiste à diffuser des annonces à ses visiteurs. Le contenu journalistique est juste un moyen d'amener les lecteurs à voir les publicités.”

Axel Springer’s lawyer, 2015
Collecting data
• websites visited
• buttons pressed
• ads viewed

Displaying ads
• take data from trackers
• find relevant ads
• display ads to the users
Blocking ads does block tracking

doubleclick.net

likes

paris.lachainmeteo.com
Web Tracking Technologies

TRACKING VIA COOKIES
What’s the original use of cookies?

- Keep the session through different windows/tabs
- Shopping basket
Within-Site Tracking

First-party cookies are used to track repeat visits to the same site.
Within-Site Tracking

First-party cookies are used to track repeat visits to the same site.

2:52pm: user visited my website
3:00pm: user visited my website

mobiliernitro.com
First-party cookies have more benefits

• Website owners can evaluate
  – website statistics
  – popularity of certain pages
  – popularity of links
  – selected and copied phrases
First-party cookie setting
Cross-site Tracking

Third-party cookies are used to track users across sites and collect their browsing history.

doubleclick.net
likes

9:30am: user visited paris.lachainmeteo.com

paris.lachainmeteo.com
Cross-site Tracking

Third-party cookies are used to track users across sites and collect their browsing history.

- Doubleclick.net
- likes
- 9:30am: user visited paris.lachainmeteo.com
- 9:31am: user visited amazon.fr

Amazon.fr
Third-party cookies blocking
Facebook lawsuit

This works because of third-party cookies!

A Belgian court on Monday gave Facebook 48 hours to stop tracking internet users who do not have accounts with the social network or risk fines of up to €250,000 a day.

Facebook said it would appeal against the order, which followed a case lodged by Belgium’s privacy watchdog in June saying the US company indiscriminately tracks internet users when they visit pages on the site or click “like” or “share”, even if they are not members.
Want to include “like” buttons but not to track users?

http://panzi.github.io/SocialSharePrivacy/
Why should I care not to track my users?

• Current law:
  – ePrivacy directive 2009 (known as “cookie law”)

29
thanks to ePrivacy directive 2009
Why should I care not to track my users?

- EU privacy regulations updated by May 2018
  - GDPR (General Data Protection Regulation)
  - ePrivacy Regulation (known as “cookie law”)

- Companies will become **liable** for inclusion of third parties that track their users

- **Fines**: 20M euro or 4% global annual turnover (whichever is bigger) **per violation**!
Alternative tracking

- What if you could track users without the need of cookies?

- **Web-based device fingerprinting**
  - Some properties of your browser can be used to track you
  - These attributes, when combined, create a quite unique fingerprint of your system

Courtesy of Nick Nikiforakis
Browser fingerprinting

• A project of the Electronic Frontier Foundation
• Analyzed 470,161 fingerprints collected in early 2010
• Each fingerprint is composed of 7 attributes (HTTP headers + JS)

94.2% of unique fingerprints

Tracking is possible

Courtesy of Pierre Laperdrix
Example of a browser fingerprint

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>User agent</td>
<td>Mozilla/5.0 (X11; Fedora; Linux x86_64; rv:55.0) Gecko/20100101 Firefox/55.0</td>
</tr>
<tr>
<td>HTTP headers</td>
<td>text/html, application/xhtml+xml, application/xhtml+xml; q=0.9, <em>/</em>; q=0.8 gzip, deflate, br en-US,en;q=0.5</td>
</tr>
<tr>
<td>Plugins</td>
<td>Plugin 0: QuickTime Plug-in 7.6.6; libtotem-narrowospace-plugin.so; Plugin 1: Shockwave Flash; Shockwave Flash 26.0 r0; libflashplayer.so.</td>
</tr>
<tr>
<td>Fonts</td>
<td>Century Schoolbook, Source Sans Pro Light, DejaVu Sans Mono, Bitstream Vera Serif, URW Palladio L, Bitstream Vera Sans Mono, Bitstream Vera Sans, ...</td>
</tr>
<tr>
<td>Platform</td>
<td>Linux x86_64</td>
</tr>
<tr>
<td>Screen resolution</td>
<td>1920x1080x24</td>
</tr>
<tr>
<td>Timezone</td>
<td>-480 (UTC+8)</td>
</tr>
<tr>
<td>OS</td>
<td>Linux 3.14.3-200.fc20.x86 32-bit</td>
</tr>
<tr>
<td>WebGL vendor</td>
<td>NVIDIA Corporation</td>
</tr>
<tr>
<td>WebGL renderer</td>
<td>GeForce GTX 650 Ti/PCIe/SSE2</td>
</tr>
<tr>
<td>Canvas</td>
<td>Cwm fjordbank glyphs vext quiz, 😊</td>
</tr>
</tbody>
</table>

Courtesy of Pierre Laperdrix
Very hard to opt-out

• Even if
  – you delete all the cookies
  – you clean all the storages (HTML5, Flash)
  – you use browser private mode

...your fingerprint remains the same!

Courtesy of Nick Nikiforakis
How unique is your browser? https://extensions.inrialpes.fr

- Browser extension detection
- ~13 000 extensions
- Websites a user is logged in
- 58 websites

Based on slides of Gábor György Gulyás
How unique is your browser?

https://extensions.inrialpes.fr

Browser Extension and Login-Leak Experiment

When you browse the web, small beacons (trackers) are spying on your online activities. Even though such trackers are invisible, they collect information about you such as which pages you visit, which buttons clicked, and what text you typed. This information is often used to show you targeted advertisements and may require you to pay a higher price during online shopping depending on the collected information.

Did you know websites can track you by your browser extensions and web logins?

Recent studies show that you can be tracked based on your web browser properties. In this experiment, we demonstrate that you can also be tracked by

- your browser extensions (such as AdBlock, Pinterest, or Ghostery), and
- the websites you have logged in (such as Facebook, Gmail, or Twitter).

You can learn more here about how these detection techniques work.

In the experiment, we will collect your browser fingerprint, together with the browser extensions installed and a list of websites you have logged in. We only collect anonymous data during the experiment (see our Privacy Policy), we will securely store the data on an Inria server, use it only for research purpose and not share it with anyone outside of Inria. You can also read the frequently asked questions here.

20 000 users have already tested!
How unique is your browser?

https://extensions.inrialpes.fr

Are you identifiable?

Yes, you are identifiable, as there are no other users who looks like you among the 20089 users we tested so far:

Are you identifiable...

...by your extensions? yes

...by your website logins? no

...by your browser fingerprint? no

...by your extensions, web logins and browser fingerprint together? yes