Mitigating Browser Fingerprint Tracking: Multi-level Reconfiguration and Diversification

Pierre Laperdrix, Walter Rudametkin, Benoit Baudry
Seams – May, 19th 2015
Table of contents

1) What is fingerprint-based tracking?
2) Presentation of Blink
3) Experimental validation
4) Conclusion and perspectives
Device fingerprinting

- Successor of cookies
- Side-effect of software diversity
- Collection of information on the device
  - Browser
  - Operating system
  - Hardware
### Example of a fingerprint

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>User agent</td>
<td>Mozilla/5.0 (X11; Linux i686; rv:25.0) Gecko/20100101 Firefox/25.0</td>
</tr>
<tr>
<td>HTTP accept</td>
<td>text/html, application/xhtml+xml, application/xml;q=0.9, <em>/</em>;q=0.8 gzip, deflate en-US,en;q=0.5</td>
</tr>
<tr>
<td>Plugins</td>
<td>Plugin 0: QuickTime Plug-in 7.6.6; libtotem-narrowspane-plugin.so; Plugin 1: Shockwave Flash; Shockwave Flash 11.2 r202; 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>HTTP DoNotTrack</td>
<td>1</td>
</tr>
<tr>
<td>Cookies enabled</td>
<td>Yes</td>
</tr>
<tr>
<td>Platform</td>
<td>Linux i686</td>
</tr>
<tr>
<td>OS</td>
<td>Linux 3.14.3-200.fc20.x86 32-bit</td>
</tr>
<tr>
<td>Screen resolution</td>
<td>1920x1080x24</td>
</tr>
<tr>
<td>Timezone</td>
<td>-480</td>
</tr>
<tr>
<td>DOM Session storage</td>
<td>Yes</td>
</tr>
<tr>
<td>DOM Local storage</td>
<td>Yes</td>
</tr>
<tr>
<td>I.E. User data</td>
<td>No</td>
</tr>
</tbody>
</table>
How unique are we?

- 83.6% of unique fingerprints
- 86.4% of unique fingerprints
How unique are we?

We can be tracked!

- 83.6% of unique fingerprints
- 86.4% of unique fingerprints

Panopticlick

Am I Unique?
How widespread is fingerprinting?

Google

- Device information
  
  We may collect device-specific information (such as your hardware model, operating system version, unique device identifiers and mobile network information including phone number). Google may associate your device identifiers or phone number with your Google Account.

Amazon

- Automatic Information
  
  Examples of the information we collect and analyze include the internet protocol (IP) address used to connect your computer to the Internet; login; e-mail address; password; computer and connection information such as browser type, version, and time zone setting, browser plug-in types and versions, operating system, and platform; purchase history, which we sometimes aggregate with similar information.

Advertising companies

- Bluecava
- AddThis
Device fingerprinting

- Silent
- Complement usage of cookies
- Hard to detect and block fingerprinting scripts
- Already adopted by major web actors
- Track users without their knowledge
- Real privacy problem
Table of contents

1) What is fingerprint-based tracking?
2) Presentation of Blink
3) Experimental validation
4) Conclusion and perspectives
Properties of a fingerprint

• Uniqueness: we can precisely identify a device thanks to its unique combination of features
• Stability: a fingerprint does not drastically change over time
• These two properties combined are the source of a real privacy problem.
Blink

- Increase temporal diversity of fingerprints.
- Reconfigure platform at runtime.
- Browsing without Blink

- Browsing with Blink
Browsing platform

- Configuration parameters:
  - screen resolution
  - language
  - timezone

- Software components:
  - plugins
  - fonts
  - OS
  - browser

- Hardware components:
  - CPU
  - microphone
  - network

- Dynamic attributes:
  - canvas

Diversifiable elements

Fingerprint:
- user agent
- fonts (name)
- accept header
- plugins (name and version)
  - ...

exhibits
Blink’s generation process

Host machine
- **OS**
- **Plugins**
  - UP1
  - UP2
- **Browser**
  - **Fonts**
    - UF1
    - UF2
- **User Profile**

Diversity reservoir
- **Browsers**
  - B1
  - B2
  - B3
  - B4
- **Plugins**
  - P1
  - P2
  - P3
  - P4
- **Fonts**
  - F1
  - F2
  - F3
  - F4

**OS**
- VM1
- VM2
- VM3

Browsing platform
Blink’s generation process
Blink’s generation process

Host machine
- OS
- Plugins
  - UP1
  - UP2
- Browser
- Fonts
  - UF1
  - UF2
- User Profile

Diversity reservoir
- Browsers
  - B1
  - B2
  - B3
  - B4
- Plugins
  - P1
  - P2
  - P3
  - P4
- Fonts
  - F1
  - F2
  - F3
  - F4

VM2
- B1

Browsing platform
Blink’s generation process
Blink’s generation process

Host machine
- OS
- Browsers
- Plugins
- Fonts
- User Profile

Diversity reservoir
- Browsers: B1, B2, B3, B4
- Plugins: P1, P2, P3
- Fonts: F1, F2, F3, F4

VMs
- VM1
- VM2
- VM3

Browsing platform
Blink’s generation process
Blink’s generation process

Host machine

- OS
- Plugins
- Browser
- Fonts

Diversity reservoir

- Browsers
  - B1
  - B2
  - B3
  - B4
- Plugins
  - P1
  - P2
  - P3
- Fonts
  - F1
  - F2
  - F3
  - F4

VM1

OS

VM2

B1

UP1

P1

UF1

F2

User Profile

VM3

P4

UF2

UP2

B2

F3

B3

F4

B4
Blink’s generation process
Blink’s generation process

Host machine
- OS
- Browser
- Plugins
  - UP1
  - UP2
- Fonts
  - UF1
  - UF2
- User Profile

Diversity reservoir
- Browsers
  - B1
  - B2
  - B3
  - B4
- Plugins
  - P1
  - P2
  - P3
  - P4
- Fonts
  - F1
  - F2
  - F3
  - F4
- OS
  - VM1
  - VM2
  - VM3

Browsing platform
Table of contents

1) What is fingerprint-based tracking?
2) Presentation of Blink
3) Experimental validation
4) Conclusion and perspectives
Objectives

1. Does dynamic reconfiguration break fingerprint stability?

2. What is the impact of the user’s plugins and fonts on global diversity?

Requirements

• A fingerprinting script

• A metric to quantify the difference between two synthesized platforms
Experimental protocol

Initial situation n°1

User browsing platform

1 plugin  20 fonts

Initial situation n°25

User browsing platform

25 plugins  500 fonts

Synthesized b. p.

OS  browser

plugins  fonts

1 plugin  20 fonts

25 plugins  500 fonts
Experimental protocol

Initial situation n°1
User browsing platform
1 plugin 20 fonts

100 synthesized b.p.

1 5 97
2 6 98
3 7 99
4 8 100

Initial situation n°25
User browsing platform
25 plugins 500 fonts

100 synthesized b.p.

1 5 97
2 6 98
3 7 99
4 8 100

2500 synthesized browsing platforms
Research questions

1. How diverse is the set of fingerprints exhibited by the synthesized platforms?
Research questions

1. How diverse is the set of fingerprints exhibited by the synthesized platforms?

More and more similar
Less and less similar

Boxplot

Dissimilarity

Blink on 25 original platforms
Research questions

1. How diverse is the set of fingerprints exhibited by the synthesized platforms?

Coffee-break mode
→ Faster mode
→ Modify fonts and plugins
→ OS and browser are not changed
Research questions

2. How diverse are the platforms in the eyes of actual fingerprinters?

http://bluecava.com/opt-out/
Research questions

2. How diverse are the platforms in the eyes of actual fingerprinters?

http://bluecava.com/opt-out/

More than 99% of unique IDs!
Table of contents

1) What is fingerprint-based tracking?
2) Presentation of Blink
3) Experimental validation
4) Conclusion and perspectives
Conclusion and perspectives

• Blink: break fingerprint stability thanks to automatic and complete synthesis of a new browsing platform for each browsing session

• AmIUnique.org: Collection of fingerprints to refine Blink’s randomization algorithms
  https://amiunique.org

• Blink on Docker: Fast and lightweight prototype already available
  https://github.com/plaperdr/blink-docker
Thank you for listening!

Any questions?