Forensic analysis of your browser fingerprint

Walter Rudametkin
Inria Spirals & University of Lille
http://rudametw.github.io
Browser cookies

client

cookie
id = b5d0

server

request-some-content

send-content

cookie
id = b5d0
Le hasard, une nécessité contre les pirates numériques

LE MONDE SCIENCE ET TECHNO | 09.06.2014 à 15h41 • Mis à jour le 10.06.2014 à 16h49

Par David Larousserie

Sans que nous le remarquions, l'arsenal technologique des ordinateurs, smartphones et cartes à puce en tout genre ont un point commun avec les.
Le hasard, une nécessité contre les pirates numériques

Par David Larousserie

Sans que nous le remarquions, l'arsenal technologique des ordinateurs, smartphones et cartes à puce en tout genre ont un point commun avec les pirates numériques.
74 trackers (most invisible)

Accuen Media, Acuity Ads, Adap.tv, Adify, Adroit Digital Solutions, AdScale, ADTECH, Advertising.com, Aggregate Knowledge, AppNexus, AT Internet, Atlas, BidSwitch, Casale Media, Cedexis Radar, Chango, ChartBeat, Connexity, Criteo, Datalogix, DataXu, Digilant, Dotomi, DoubleClick, DoubleClick Bid Manager, DoubleClick Spotlight, EQ Advertising, Eulerian, Experian Marketing Services, eyeReturn Marketing, Ezakus, Facebook Connect, Facebook Exchange (FBX), Facebook Social Plugins, Google Adsense, Google Analytics, Improve Digital, Integral Ad Science, Jumptap, Kameleoon, Ligatus, Lijit, Magnetic, Media Innovation Group, Media Optimizer (Adobe), Media6Degrees, MediaMath, Netmining, Neustar AdAdvisor, OpenX, Optimix Media Delivery, Outbrain, OwnerIQ, PubMatic, PulsePoint, Quantcast, Right Media, Rocker Fuel, Rubicon, ScoreCard Research Beacon, SiteScout, Sizmek, SMART AdServer, SpotXchange, TradeDesk, TubeMogul, Turn, Twitter Button, Veruta, Videology, Video Step, Visual Revenu, Yandex.Metrics, Yieldr
client

- cookie id = b5d0
- cookie id = 13pm
- cookie id = w7nq

server

- request-some-content
- send-content

triangle nodes:
- advertiser
- advertiser
- social-media
- social-media

script nodes:
- JS

network diagram showing interactions between client and server, with cookies and script elements.
“Cookie awareness”

• Legal action
  • CNIL, GDPR, ePrivacy

• Extensions
  • 95 add-ons for Firefox for cookies
  • Tor, Ad-Blockers, Anti-trackers

• Browser parameters
  • accept cookies, third-party cookies, etc.

• Private navigation
  • removes cookies after each session

• Browsers
  • Brave
More techniques

• But also *more aggressive* cookies & fingerprinting

  • Evercookies (multiple cookies), flash cookies (LSO), Permacookie (e.g., Verizon), Google Cookie (used by NSA).

  • Advanced Device Identification using Browser Fingerprinting
## 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; Linux x86_64; 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-narrowospace-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 x86_64</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>
A special attribute: Canvas fingerprint

Cwm fjordbank glyphs vext quiz, 😊

Cwm fjordbank glyphs vext quiz, 😊
Browser fingerprinting

• Began in 2010 and is growing
  [PETS’10\textsuperscript{1}, CCS’13\textsuperscript{2}, CCS’14\textsuperscript{3}, CCS’16\textsuperscript{4}]

• Some defenses exist
  • NoScript, randomization/blocking extensions, Brave, …

• Discovering new attributes all the time
  • Searching for unique and stable fingerprints

1. Eckersley, P. How unique is your web browser?
Learn how identifiable you are on the Internet

Help us investigate the diversity of web browsers

View my browser fingerprint

By clicking on this button, only anonymous data will be collected and a cookie will be stored in your browser for four months. You can find more details in the Privacy Policy.

Spread the word! Share AmlUnique!
Try it on all your devices!

What is browser fingerprinting? Learn more

Any questions? Send us an email at contact@amiunique.org
http://amiunique.org  (am I unique?)

• Inform users

• Study advanced fingerprinting techniques

• Highly visible project
  • Website, browser extensions, research papers, news articles, interviews, and lots of vulgarization

• More than 1 million fingerprints collected
Finding #1: Mobile fingerprints are also unique

• Mobile browser fingerprinting is feasible
  • High uniqueness (depends on make/model)
  • High stability
  • Different attributes (e.g., user agent, emojis)

• As browsers close privacy holes, they also add new APIs
  • Increased attack surface
  • Little thought to privacy

Beauty and the Beast: Diverting modern web browsers to build unique browser fingerprints [IEEE S&P’16]
## Emoji fingerprinting

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(a)</td>
<td>Windows 7</td>
<td>(b)</td>
<td>Windows 10</td>
</tr>
<tr>
<td>(c)</td>
<td>Linux</td>
<td>(d)</td>
<td>iOS</td>
</tr>
<tr>
<td>(e)</td>
<td>Firefox OS</td>
<td>(f)</td>
<td>Android 4.3 and before</td>
</tr>
<tr>
<td>(g)</td>
<td>Android 4.4</td>
<td>(h)</td>
<td>Android 5.0</td>
</tr>
<tr>
<td>(i)</td>
<td>Android on an LG device</td>
<td>(j)</td>
<td>Android on a Samsung device</td>
</tr>
<tr>
<td>(k)</td>
<td>Android on an HTC device</td>
<td>(l)</td>
<td>Emoji not supported</td>
</tr>
</tbody>
</table>

[IEEE S&P’16]
Long-term fingerprint tracking

FPStalker: Machine learning for tracking fingerprint evolutions

[IEEE S&P’18]
Example of Decision Tree [IEEE S&P’18]

Output: Link  Don’t Link
Finding #2: Fingerprints can be tracked

• ~26% of browsers are highly trackable
  • More than 100 days
  • If you try to hide you’re worse off

• ~20% of browsers are “untrackable”
  • Fingerprints too similar (not enough attributes)
  • Unpredictable evolutions

[IEEE S&P’18]
What about fingerprinting defenses?

FP-Scanner: Detecting incoherencies in countermeasures

[USENIX Security’18]
Finding #3: Current defenses are not very effective [USENIX Security’18]

• Current solutions are (very) bad
  • Can be counterproductive

• Privacy is difficult
  • Many side-effects to watch for
  • Spoofing must be coherent throughout multiple channels
  • Browsers are very complex

• The better solutions…
  • Integrated in the browser
  • Block or randomize attributes
  • Reducing APIs/Features
Can we design better tools?

• What do developers need?
• What kind of feedback to provide?
• Can we automate this?
Finding #4: We can design better countermeasures but it’s hard

- Development and testing phases
  - Uniqueness and stability algorithms
- Automated reports for developers to inform on
  - Uniqueness (entropy)
  - Trackability (stability)
- Automate configuration exploration

FP-Tester: Designing fingerprint-proof browser extensions [IWPE ‘18]
Summary

• Fingerprinting is a threat to privacy
  • Difficult to solve (arguably harder than cookies)
  • Use is growing
  • Can be used to track users and to complement cookies
• Currently no good defenses
  • Many are counterproductive
• But... it is possible to build better countermeasures
  • By providing automated tools
Perspectives

• Combine fingerprinting with other techniques
• Explore good uses for fingerprinting
  • Second-tier security, bot detection, fraud detection
    • Verify, block, honey pots, poisoning
• Design and build a fingerprint resilient browser
  • In-browser randomization, attribute blocking, whitelisting
• ADT FingerKit [proposal under review]
  • Spirals (Inria), Indes (Inria), Stony Brook University
  • New AmlUnique project + Extensions.inria.fr
  • Scientific workbench, tools to end-users
QUESTIONS ?
Towards AmlUnique2.0 with the ADT Fingerkit
Blink: Proactive diversification [SEAMS’14]
Automatic randomization of the platform: Leery mode
Automatic randomization of the platform: Coffee-break mode