UNIVERSITÉ CÔTE D'AZUR

Inria
INVENTEURS DU MONDE NUMÉRIQUE

Browser fingerprinting

Nataliia Bielova

@nataliabielova

February 12th, 2019

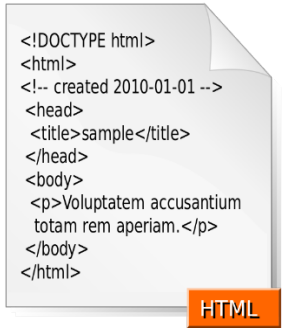
Security and ethical aspects of data

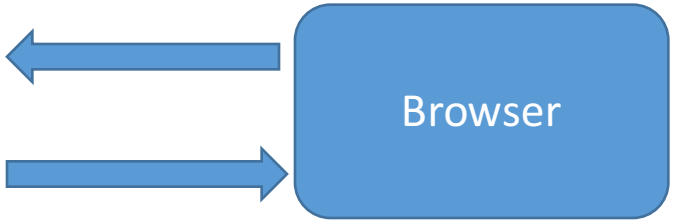
Université Cote d'Azur

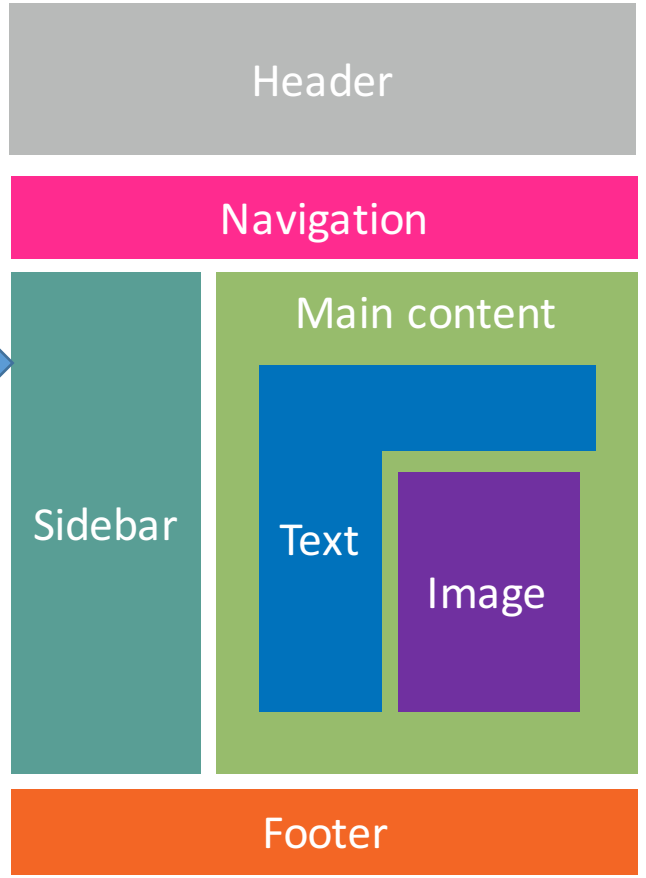
Today’s class

- A brief history of Web browsers
- What is browser fingerprinting?
- From basic to advanced fingerprinting

I. Internet and web browsers







Slides courtesy of Pierre Laperdrix (Stony Brook University)

I. Internet in 1995

HTTP User agent

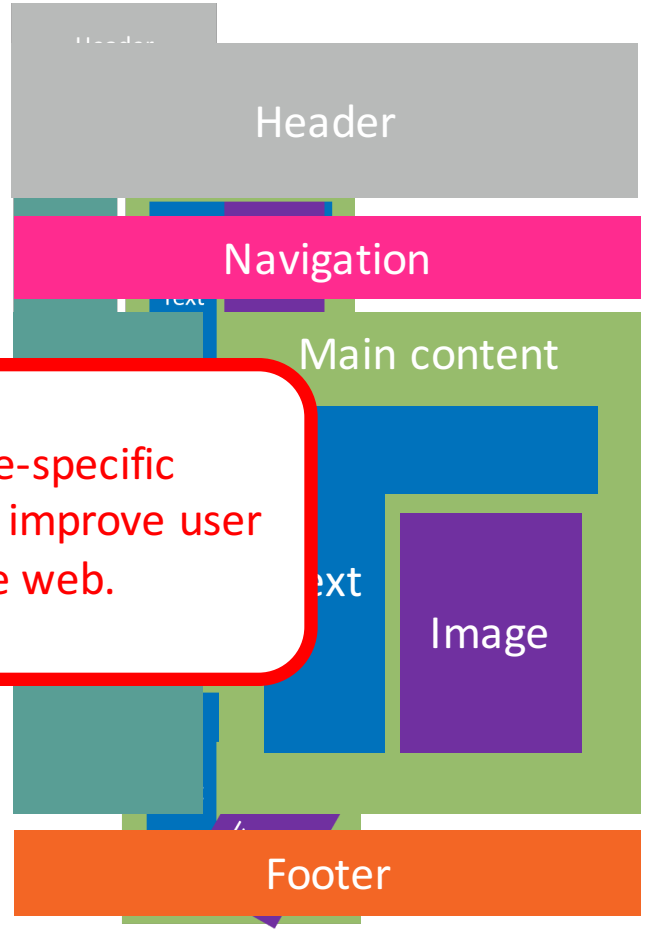
NCSA_Mosaic/2.0 (Windows 3.1)

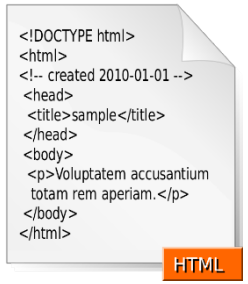
Mozilla/1.22 (compatible; MSIE 2.0; Windows 95)

I am

I am

Browsers send device-specific information to servers to improve user experience on the web.

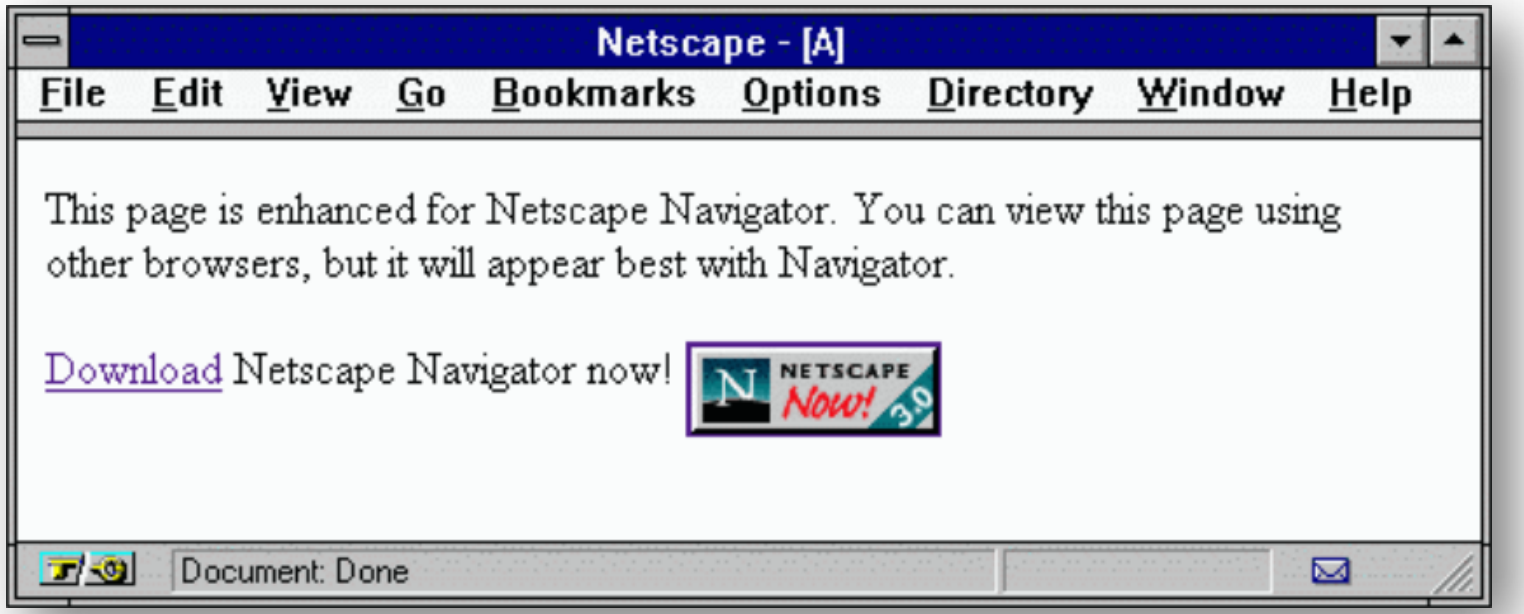




Slides courtesy of Pierre Laperdrix (Stony Brook University)

I. Internet in 1995

- Every website announces with **what browser** it is recommended to visit the website



Slides courtesy of Pierre Laperdrix (Stony Brook University)

I. Internet in 2017

1995	2017
Browser: Netscape Language: Fr	Browser: Chrome v53 OS: Linux Screen: 1920x1080 Language: Fr Timezone: GMT+1 Graphic card: GTX 1080Ti …

A bigger and richer web

- Audio
- Video
- 3D rendering
- Real-time communications
- Web payments
- Virtual reality

…

What happens when we start collecting all the information available in a web browser?

Example of a browser fingerprint

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

Slides courtesy of Pierre Laperdrix (Stony Brook University)

I. Definition of browser fingerprinting

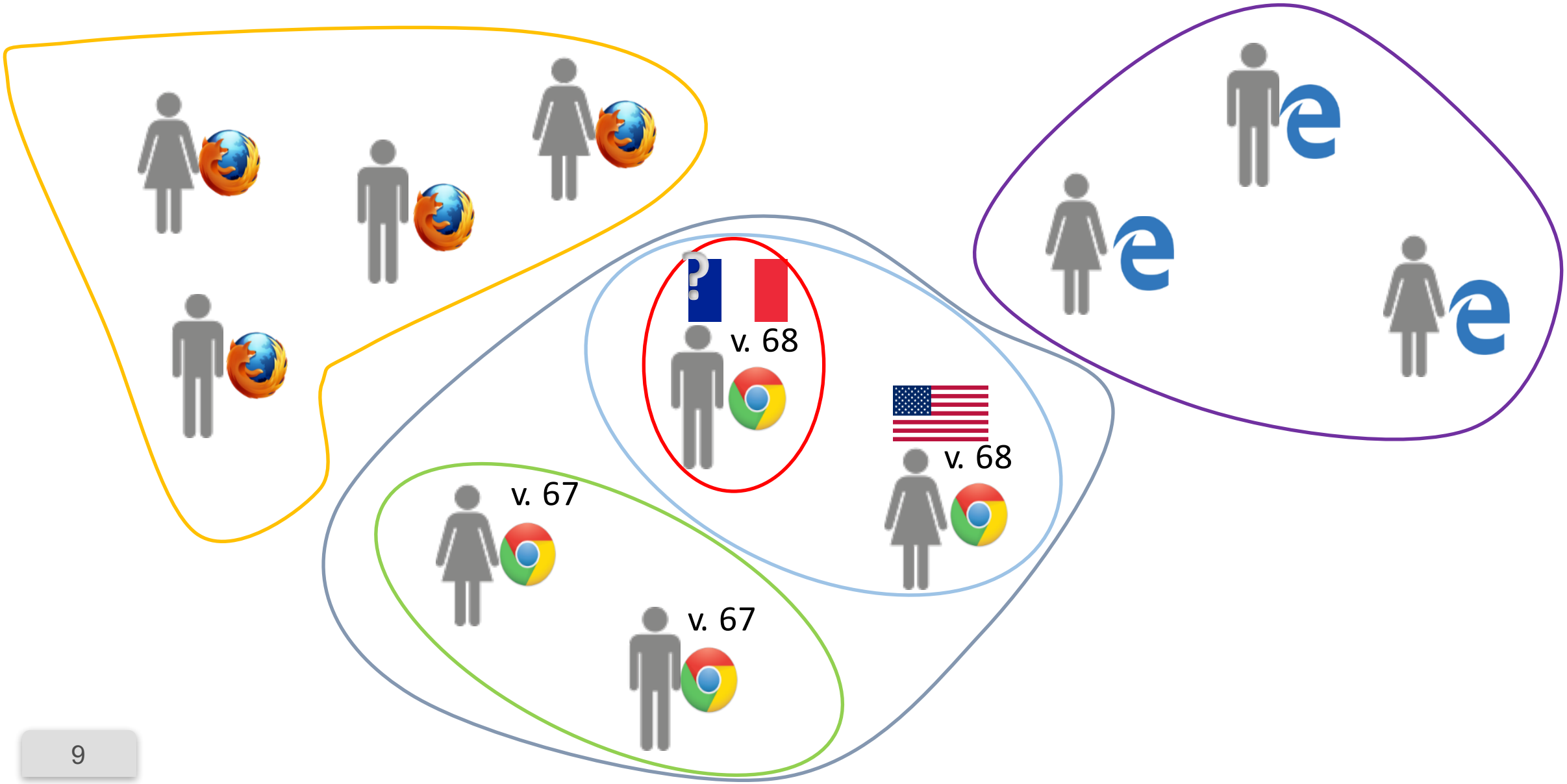
Definitions

- A browser fingerprint is a set of information related to a user's device from the hardware to the operating system to the browser and its configuration.

- Browser fingerprinting refers to the process of collecting information through a web browser to build a fingerprint of a device.

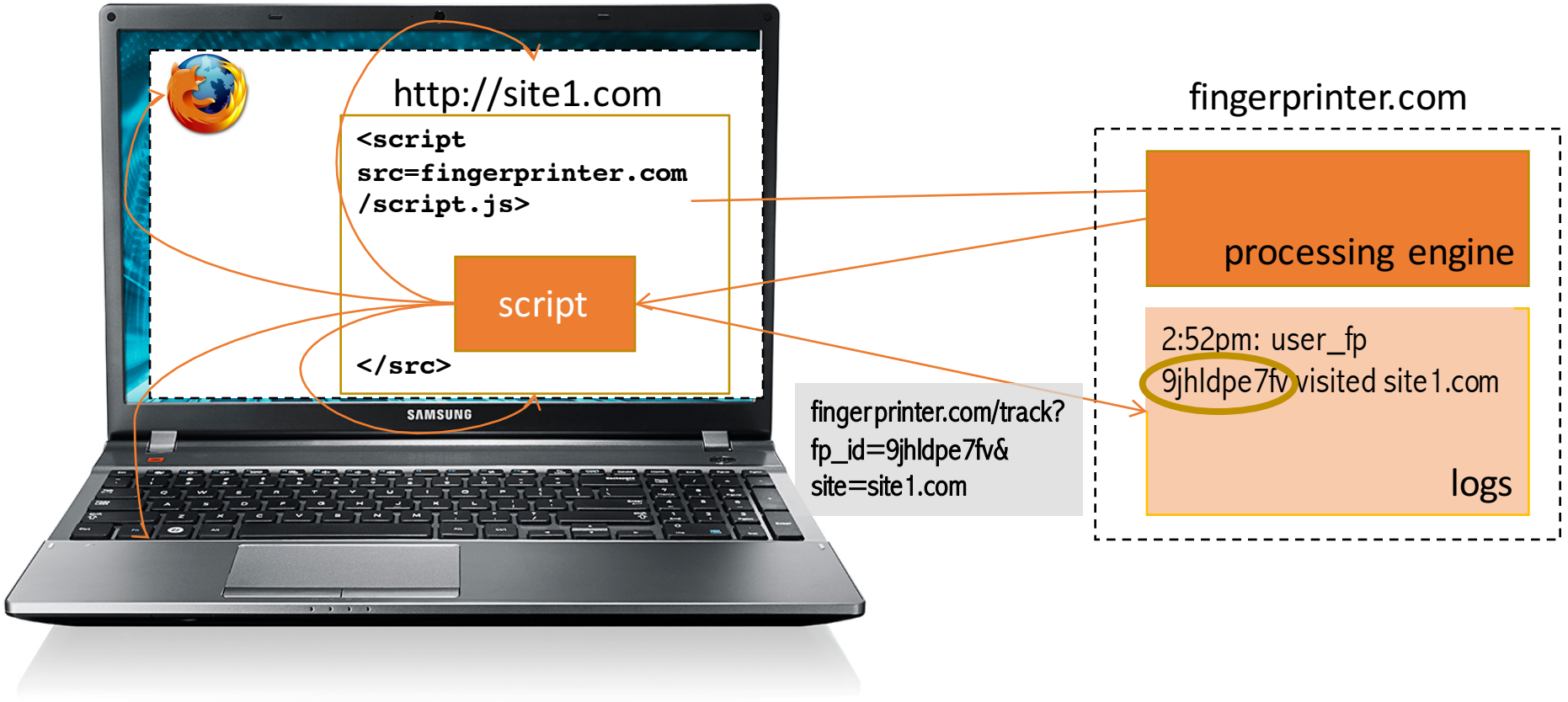
Slides courtesy of Pierre Laperdrix (Stony Brook University)

How can we be identified by a browser fingerprint?



Browser fingerprinting used for tracking

Browser and operating system properties are used to track repeated visits to a site.



Comparison of the emoji on different devices and OSs

(a) Windows 7

(b) Windows 10

(c) Linux

(d) iOS

(e) Firefox OS

(f) Android 4.3 and before

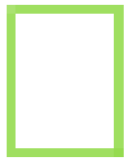
(g) Android 4.4

(h) Android 5.0

(i) Android on an LG device

(j) Android on a Samsung device

(k) Android on an HTC device



(l) Emoji not supported

https://hal.inria.fr/hal-01285470/document

What makes fingerprinting a threat to online privacy?

Two studies have investigated the diversity of browser fingerprints.

Panopticlick
How Unique – and Trackable – Is Your Browser?

EFF

Am I Unique?

470,161 fingerprints
94.2% were unique

118,934 fingerprints
89.4% were unique

Tracking is possible

Slides courtesy of Pierre Laperdrix (Stony Brook University)

Fingerprinting

- Panopticlick [Eckersley, PET'2010]

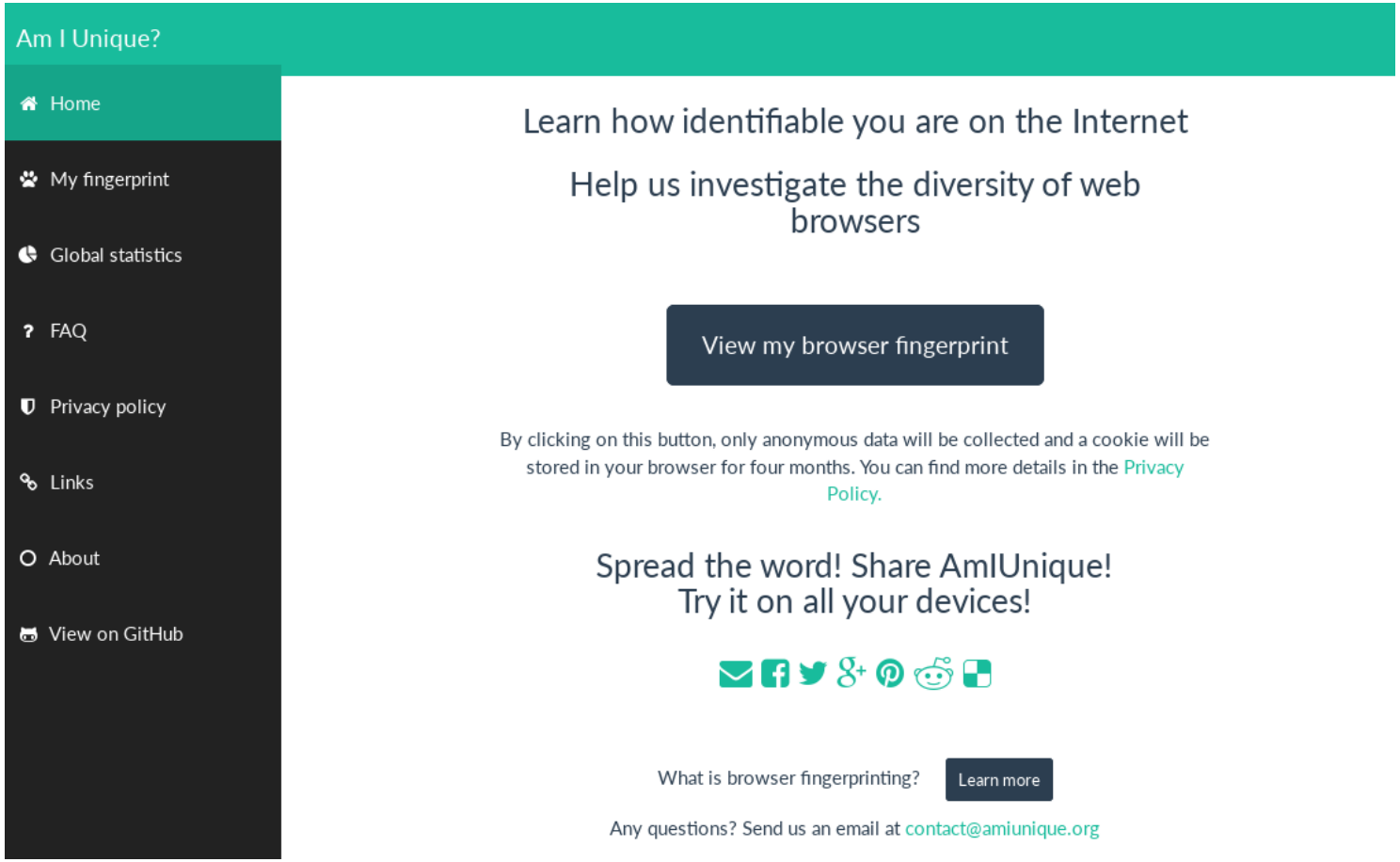
> Your browser fingerprint **appears to be unique** among the 2,419,678 tested so far.
>
> Currently, we estimate that your browser has a fingerprint that conveys **at least 21.21 bits of identifying information.**

- Information needed to **uniquely identify a browser**
  - n – number of connected devices: 5 000 000 000
  - $log_2 n$ – number of bits for a unique id: 33 bits

- **Idea: distinguish user's browsers** by accessing browser features and using their probability distributions

I. See your own fingerprint

https://amiunique.org (Am I Unique)



- Website launched in November 2014
- Collected 660,000+ fingerprints so far
- Browser extension available to see the evolution of your own browser fingerprint

Slides courtesy of Pierre Laperdrix (Stony Brook University)

I. Example of values collected on AmIUnique

Slides courtesy of Pierre Laperdrix (Stony Brook University)

How to compare datasets: Anonymity sets

- User-agent on Desktop vs Mobile devices

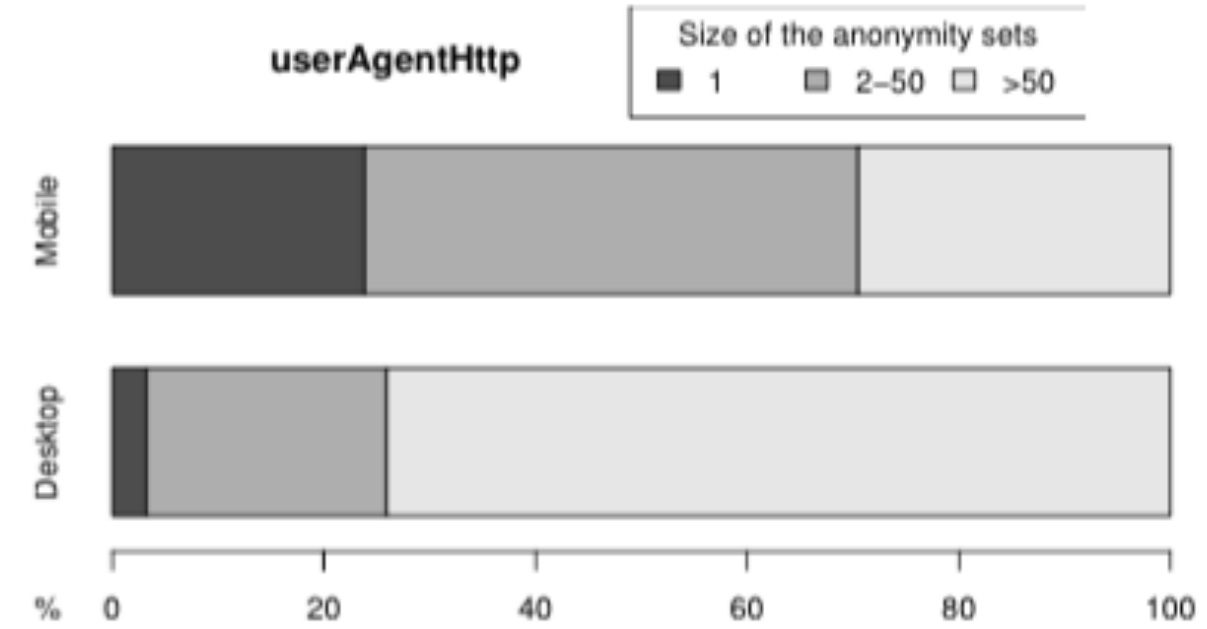


Fig. 4. Comparison of anonymity set sizes on the user-agent between desktop and mobile devices

I. Example of values collected on AmIUnique

Some user-agents

- Mozilla/5.0 (Windows NT 6.1; WOW64; rv:34.0) Gecko/20100101 Firefox/34.0
- Mozilla/5.0 (iPhone; CPU iPhone OS 8_1_2 like Mac OS X) AppleWebKit/600.1.4 (KHTML, like Gecko) Version/8.0 Mobile/12B440 Safari/600.1.4
- Mozilla/5.0 (Android; Mobile; rv:27.0) Gecko/27.0 Firefox/27.0
- Mozilla/5.0 (Macintosh; Intel Mac OS X 10_10_2) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/39.0.2171.95 Safari/537.36
- Mozilla/5.0 (X11; Ubuntu; Linux i686; rv:34.0) Gecko/20100101 Firefox/34.0

I. Example of values collected on AmIUnique

Other custom user-agents

- godzilla/5.0 (X122; BSD; rv:500.0) Gecko/20100101
- pouet
- “54. When a warlike prince attacks a powerful state, his generalship shows itself in preventing the concentration of the enemy's forces. He overawes his opponents, and their allies are prevented from joining against him.”
- Deepnet Explorer 1.5.3; Smart 2x2; Avant Browser; .NET CLR 2.0.50727; InfoPath.1)
- NSA
- Game Boy Advance
- eat it

Slides courtesy of Pierre Laperdrix (Stony Brook University)

Anonymity sets for mobile devices

- User-agent on Android vs iOS devices

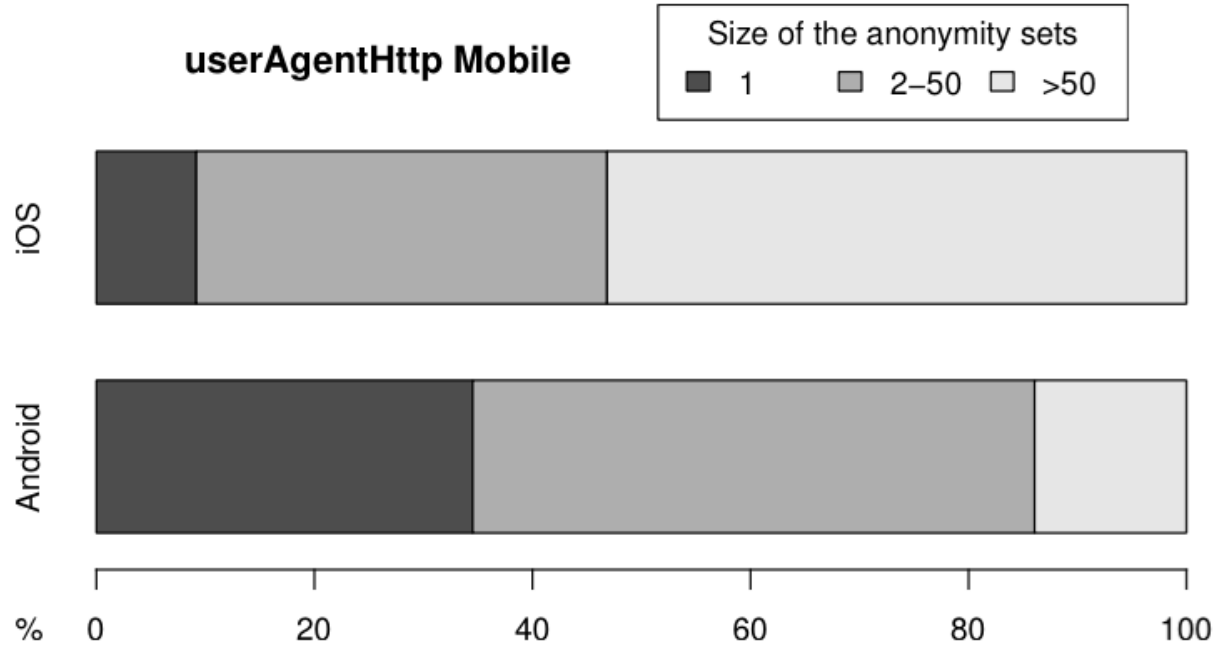


Fig. 5. Comparison of anonymity set sizes on the user-agent between Android and iOS devices

What if I disable JavaScript?

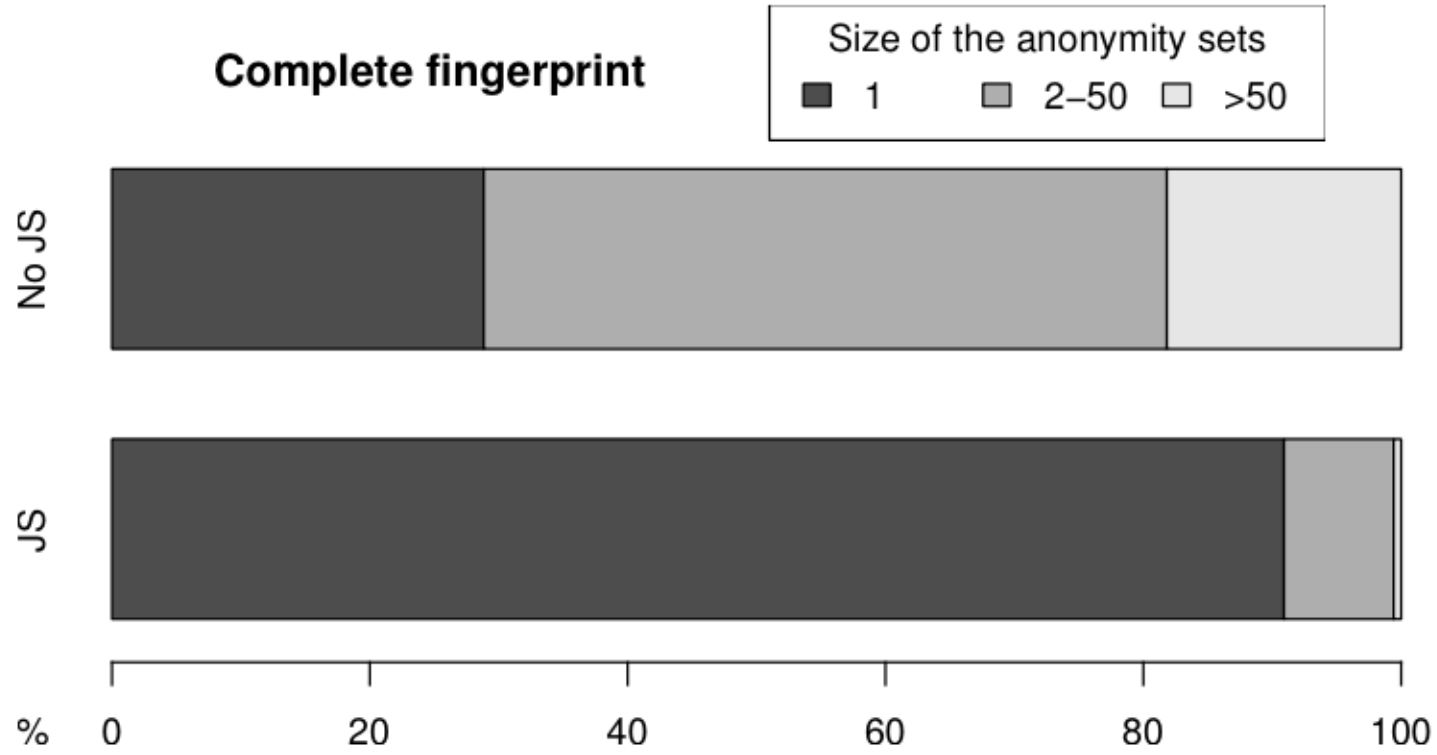


Fig. 9. Comparison of anonymity set sizes on the complete fingerprint between devices with and without JavaScript

I. Summary

- Servers can easily collect information about a device to form what is called a **browser fingerprint**.
- There is so much diversity that users can be **tracked** online if their fingerprint is **unique**.
- Test your device on

  https://amiunique.org and https://extensions.inrialpes.fr

Slides courtesy of Pierre Laperdrix (Stony Brook University)

Very hard to opt-out

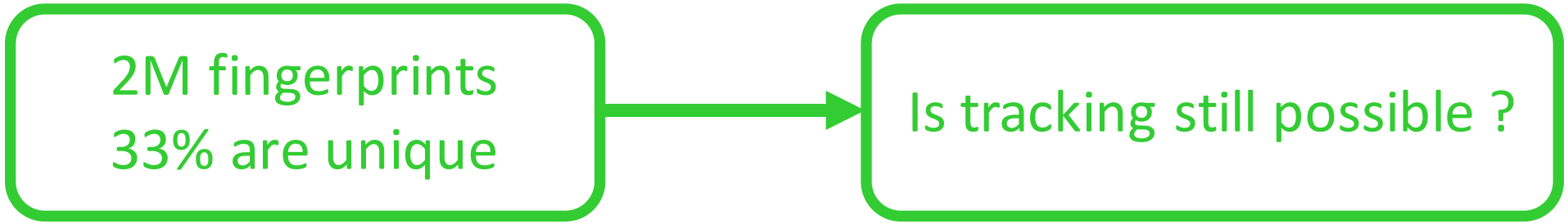
- Even if
  - you delete all the cookies
  - you clean all the storages (HTML5, Flash)
  - you use browser private mode

…your fingerprint remains the same!

Courtesy of Nick Nikiforakis

I. Summary

- How effective is fingerprinting at large scale?



Hiding in the Crowd: an Analysis of the Effectiveness of Browser Fingerprinting at Large Scale

Alejandro Gómez-Boix, Pierre Laperdrix, Benoit Baudry

The Web Conference (**WWW 2018**)

Slides courtesy of Pierre Laperdrix (Stony Brook University)

2M users in France (WWW 2018)

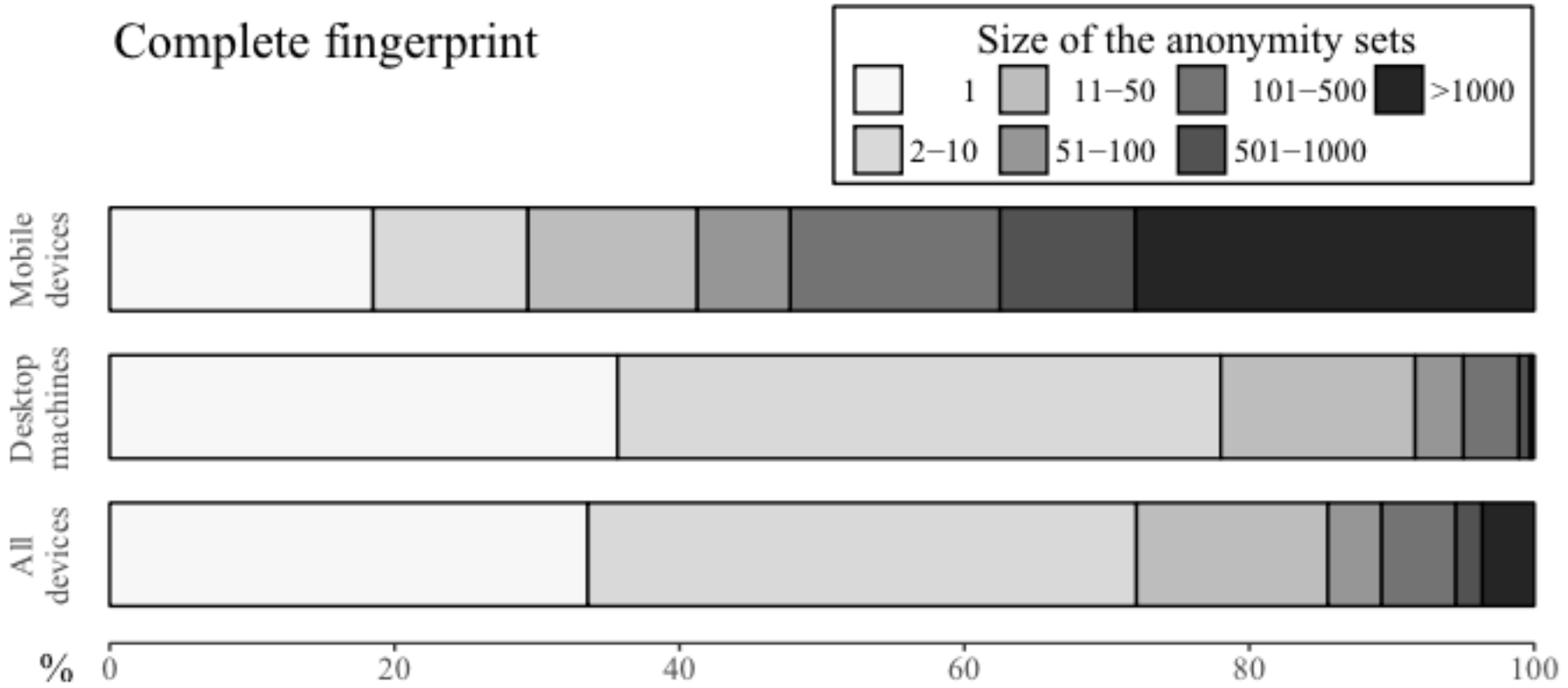
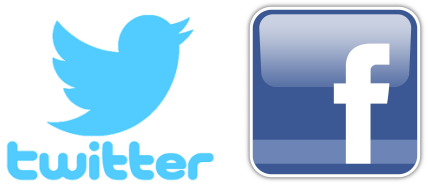


Figure 3: Comparison of anonymity set sizes between mobile devices and desktop/laptop machines.

Slides courtesy of Pierre Laperdrix (Stony Brook University)

New Fingerprinting Methods

- **Privacy Paradox**
  - Users' fingerprints can be enriched by their browser extensions
  - Moreover, we found an attack allows to detect 58 web services where the user is logged in!



G.G. Gulyás, D. F. Some, **N. Bielova** and C. Castelluccia. To Extend or not to Extend: on the Uniqueness of Browser Extensions and Web Logins. *WPES@ACM CCS 2018.*

I. Plugins VS Browser extensions

- **Plugins** were created to display content not supported by the browser
  - Flash Java Silverlight



- All installed plugins are accessible via the `navigator.plugins` JavaScript object

- **Extensions** extend or modify default behavior of a browser
  - AdBlockPlus, LastPass, Ghostery, Pinterest

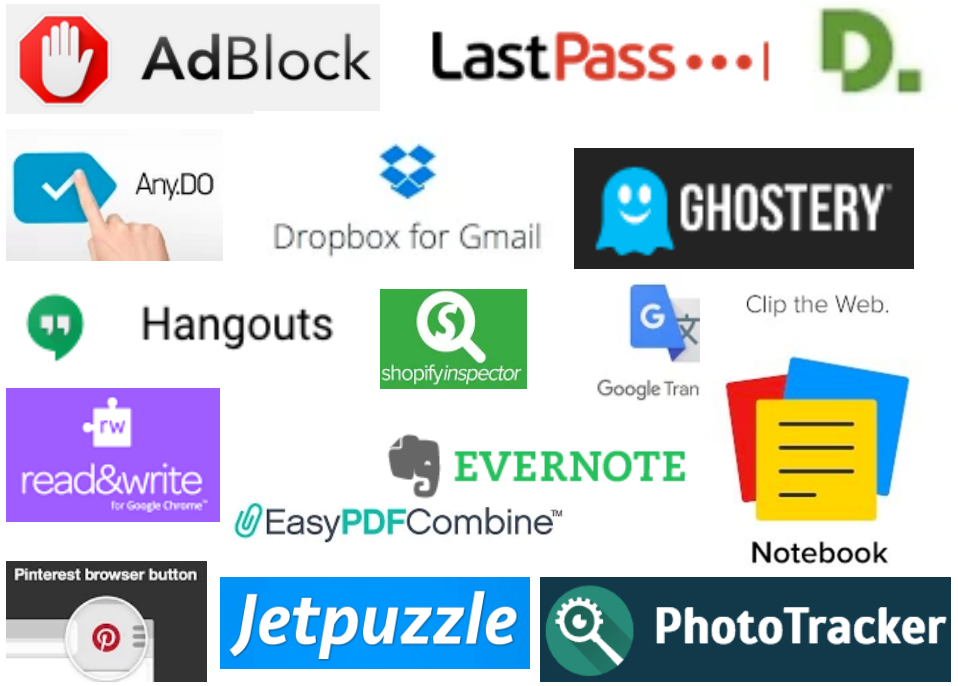

- There is no API that webpages can use to detect all installed extensions

Slides courtesy of Pierre Laperdrix (Stony Brook University)

How unique is your browser?

https://extensions.inrialpes.fr

- Browser extension detection
- ~13 000 extensions



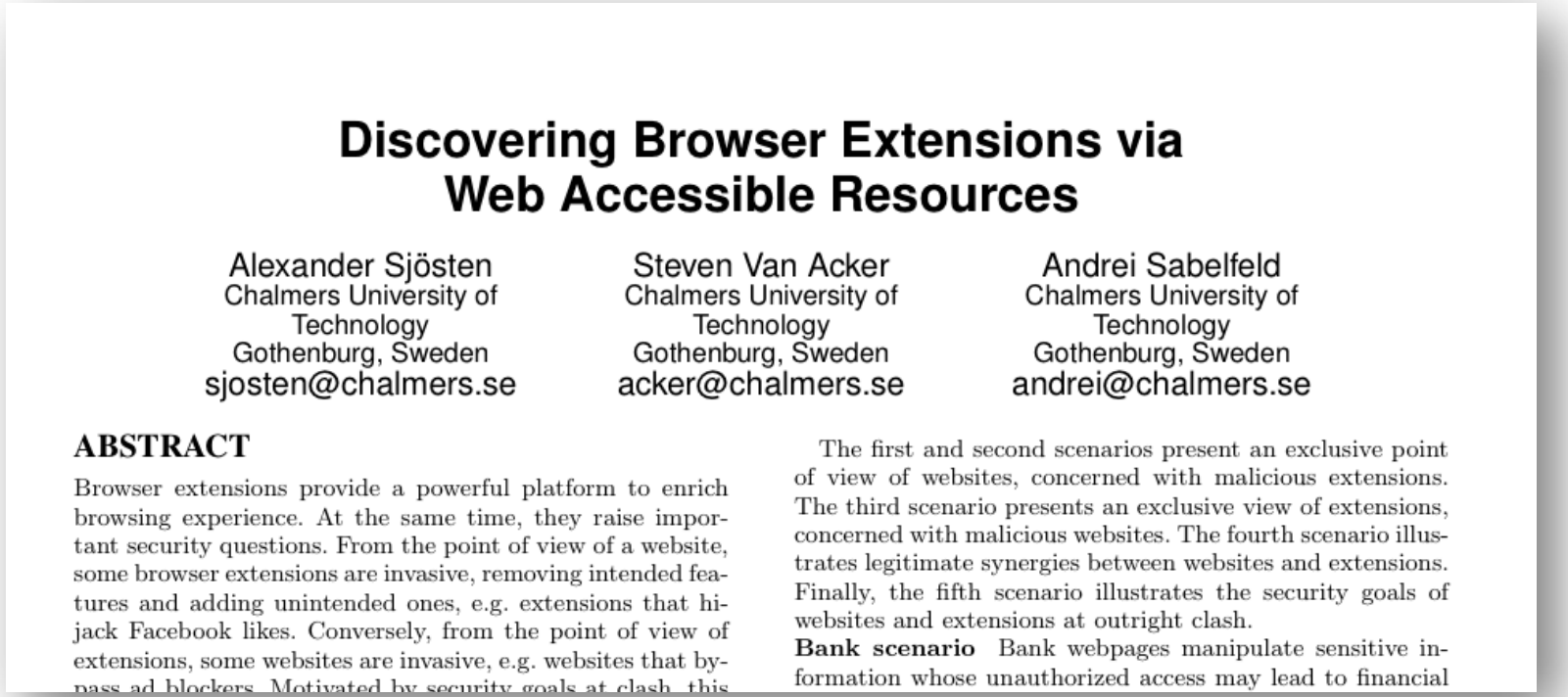
- Websites a user is logged in
- 58 websites

Browser extension detection

- via **Web Accessible Resources**

chrome-extension://gpdjojdkbbmdfjfahjcgigfpmkopogic/img/icon_48.png

unique extension ID



Discovering Browser Extensions via Web Accessible Resources

Alexander Sjösten
Chalmers University of Technology
Gothenburg, Sweden
sjosten@chalmers.se

Steven Van Acker
Chalmers University of Technology
Gothenburg, Sweden
acker@chalmers.se

Andrei Sabelfeld
Chalmers University of Technology
Gothenburg, Sweden
andrei@chalmers.se

ABSTRACT

Browser extensions provide a powerful platform to enrich browsing experience. At the same time, they raise important security questions. From the point of view of a website, some browser extensions are invasive, removing intended features and adding unintended ones, e.g. extensions that hijack Facebook likes. Conversely, from the point of view of extensions, some websites are invasive, e.g. websites that bypass ad blockers. Motivated by security goals at clash, this

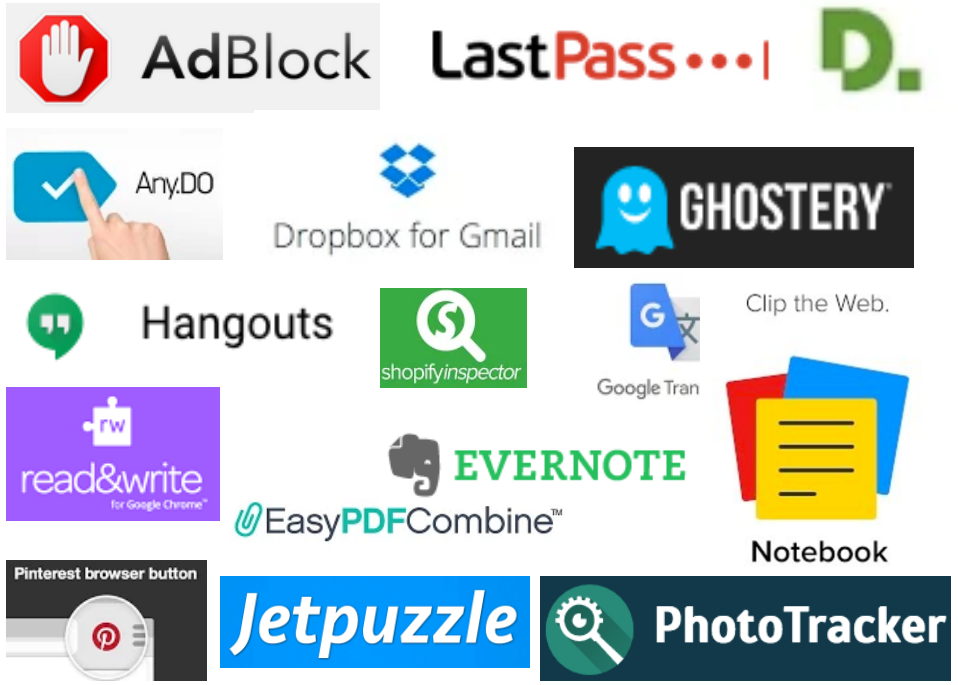
The first and second scenarios present an exclusive point of view of websites, concerned with malicious extensions. The third scenario presents an exclusive view of extensions, concerned with malicious websites. The fourth scenario illustrates legitimate synergies between websites and extensions. Finally, the fifth scenario illustrates the security goals of websites and extensions at outright clash.

Bank scenario Bank webpages manipulate sensitive information whose unauthorized access may lead to financial

How unique is your browser?

https://extensions.inrialpes.fr

- Browser extension detection
- ~13 000 extensions

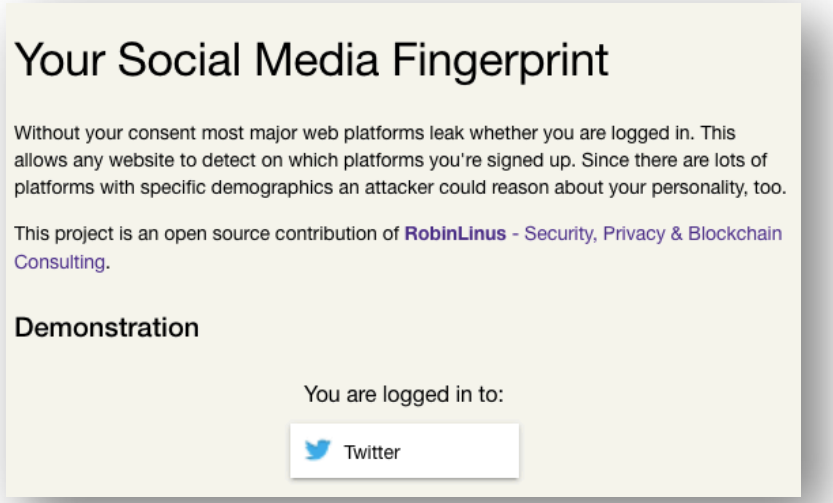


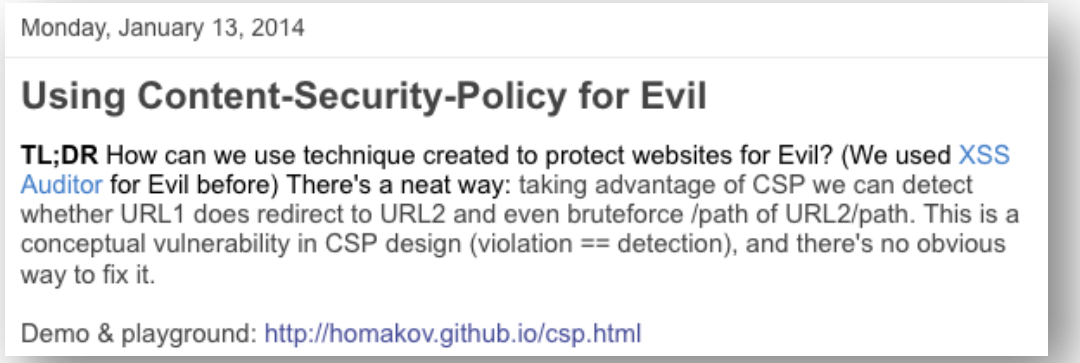
- Websites a user is logged in
- 58 websites

Detection of websites a user logged in

- Redirection URL hijacking @robin_linus

- Abusing Content Security Policy (CSP) – no JavaScript needed @homakov



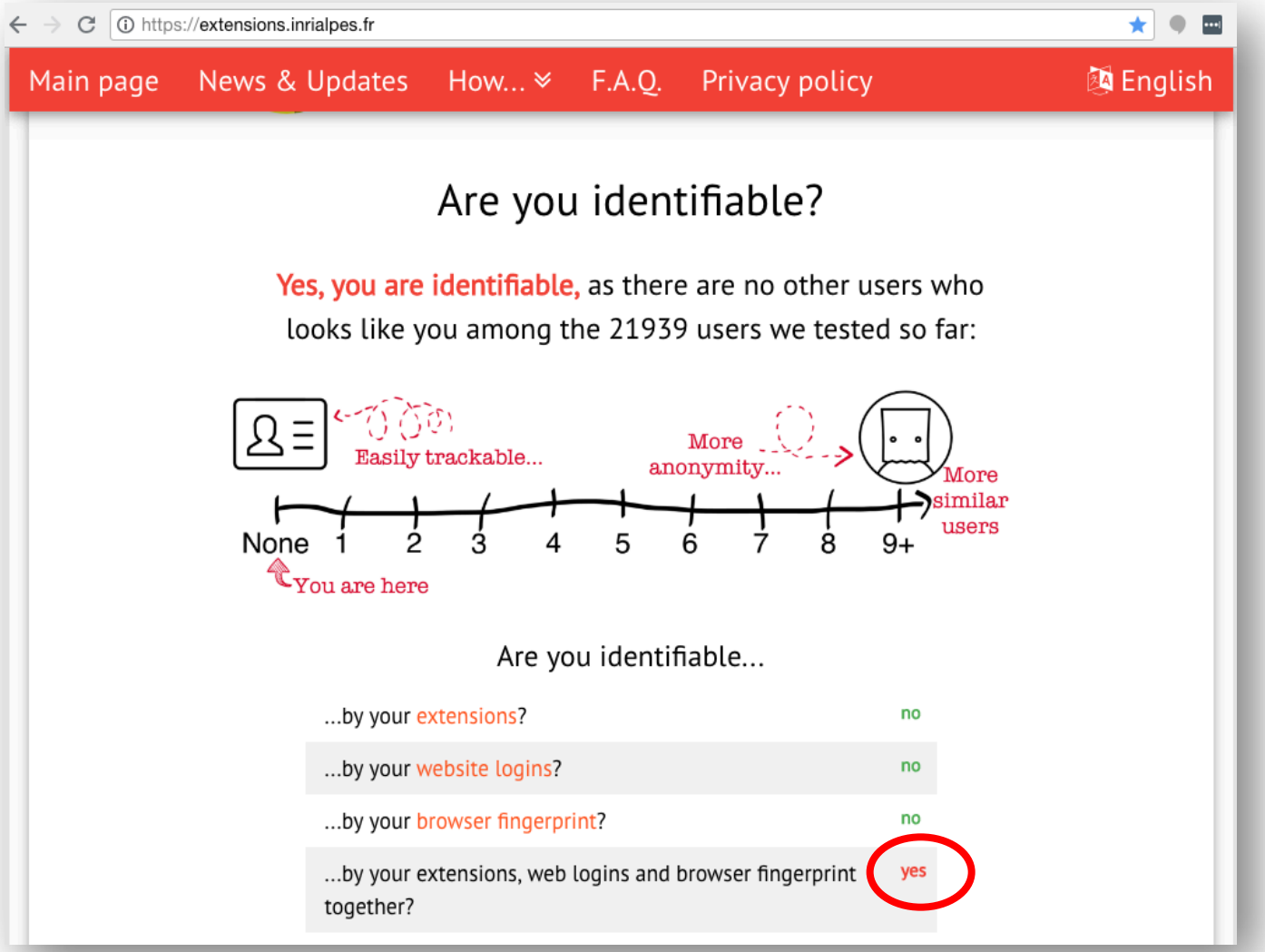


How unique is your browser?

https://extensions.inrialpes.fr

How unique is your browser?

https://extensions.inrialpes.fr



User dataset w.r.t previous studies

Table 2: Previous studies on measuring uniqueness based on browser extensions and our estimation of uniqueness.

Study	Fingerprints collected in a study	Extensions targeted in a study	Unique finger-prints in a study	Unique fin-gerprints in our dataset
Timing leaks [54]	204	2,000	56.86%	55.64%
XHOUND [58]	854	1,656	14.10%	49.60%
Ours	7,643	13k	39.29%	39.29%

Uniqueness grows as the dataset grows!

How to get a meaningful dataset?

How to define when we have enough users?

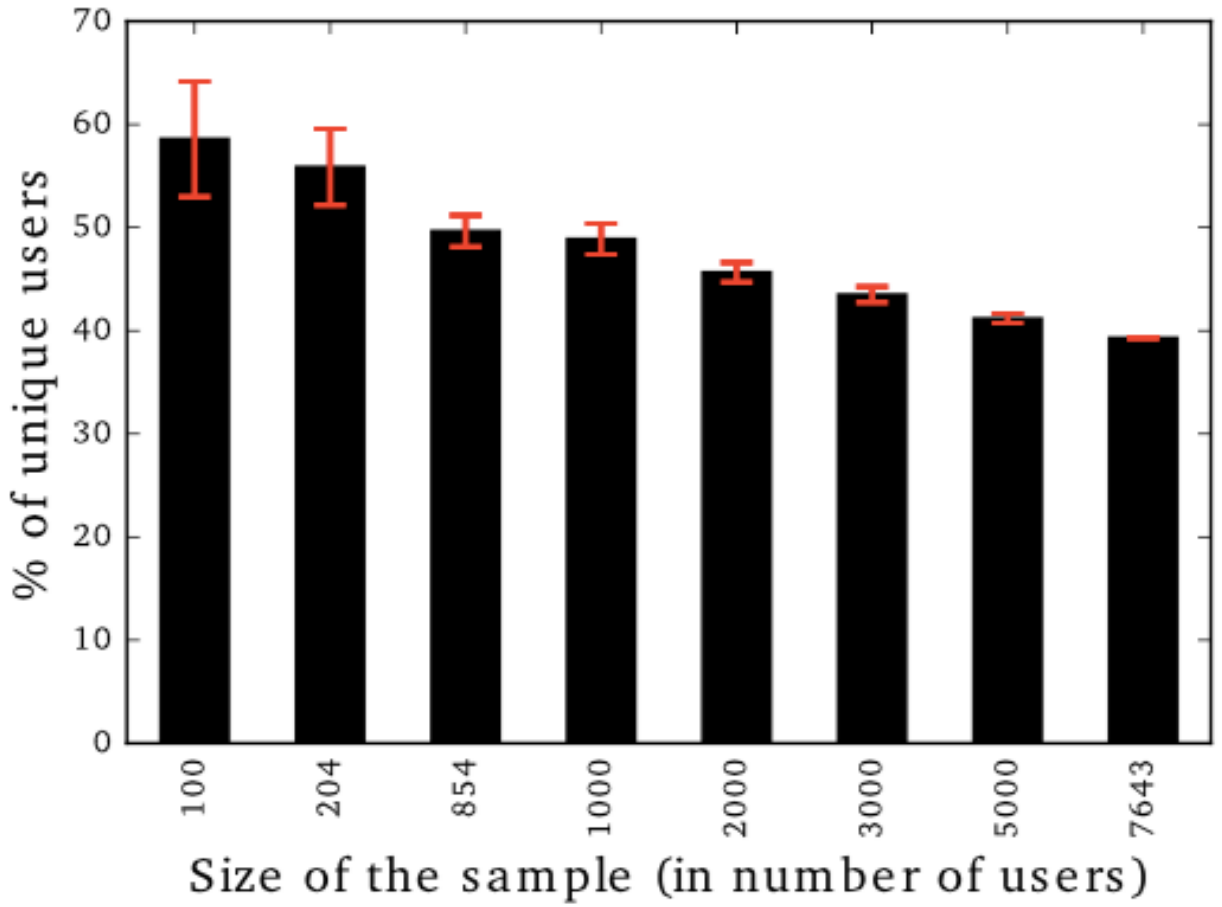
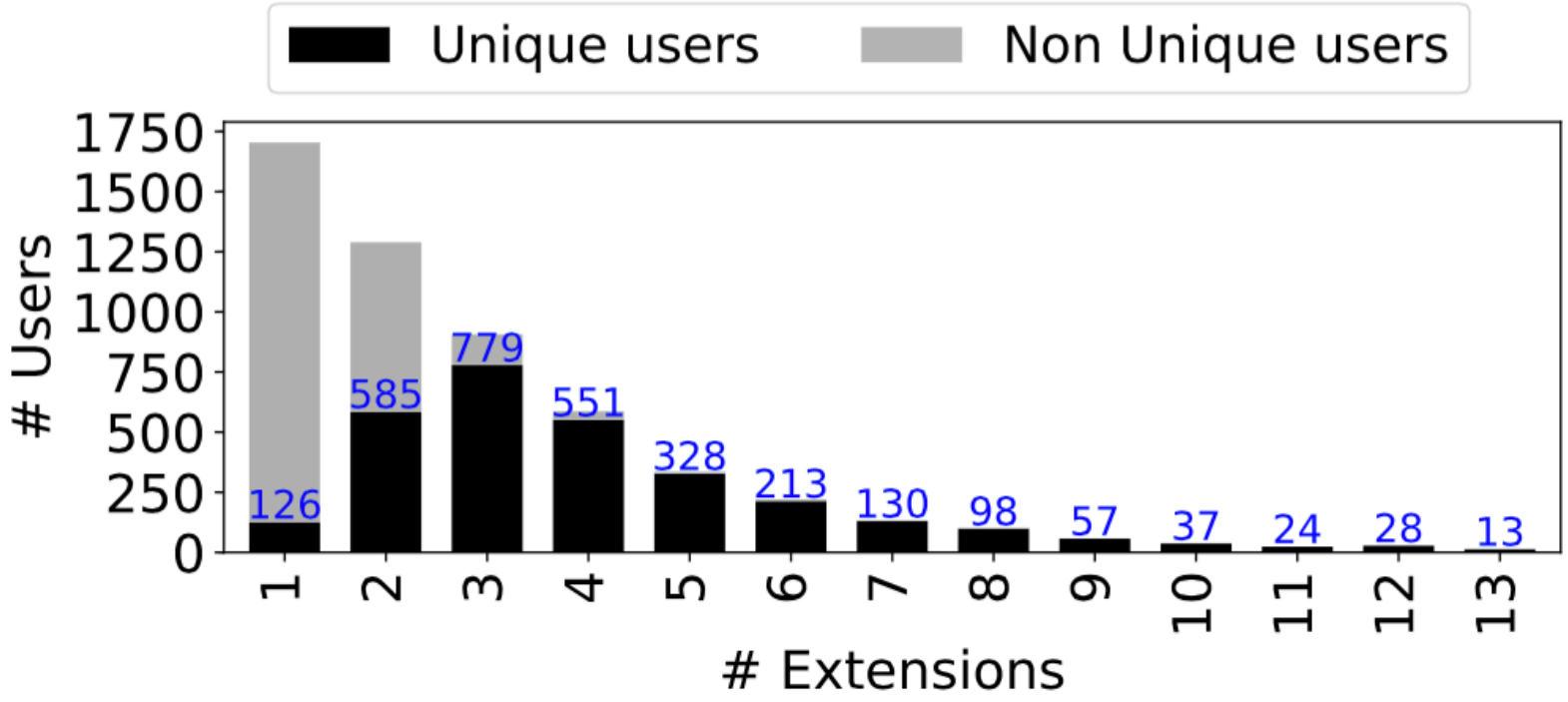


Figure 13: Uniqueness of Chrome users based on their extensions only vs. number of users - 204 is the number of users used in [54] and 854 the number of users considered in [58]

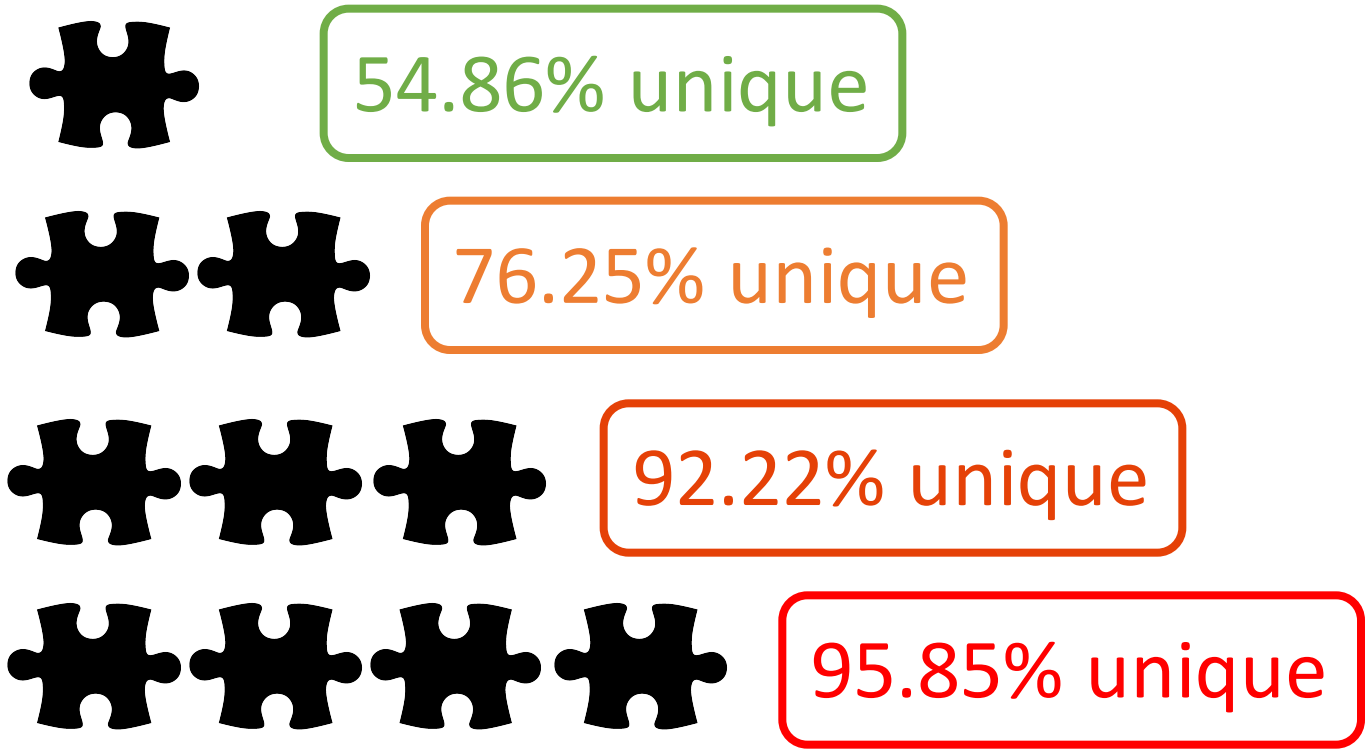
How many extensions our users have?

7,643 users of Google Chrome browser



G.G. Gulyás, D. F. Some, **N. Bielova** and C. Castelluccia. To Extend or not to Extend: on the Uniqueness of Browser Extensions and Web Logins. *WPES@ACM CCS 2018.*

Am I really unique if I use a few extensions?



G.G. Gulyás, D. F. Some, **N. Bielova** and C. Castelluccia. To Extend or not to Extend: on the Uniqueness of Browser Extensions and Web Logins. *WPES@ACM CCS 2018.*

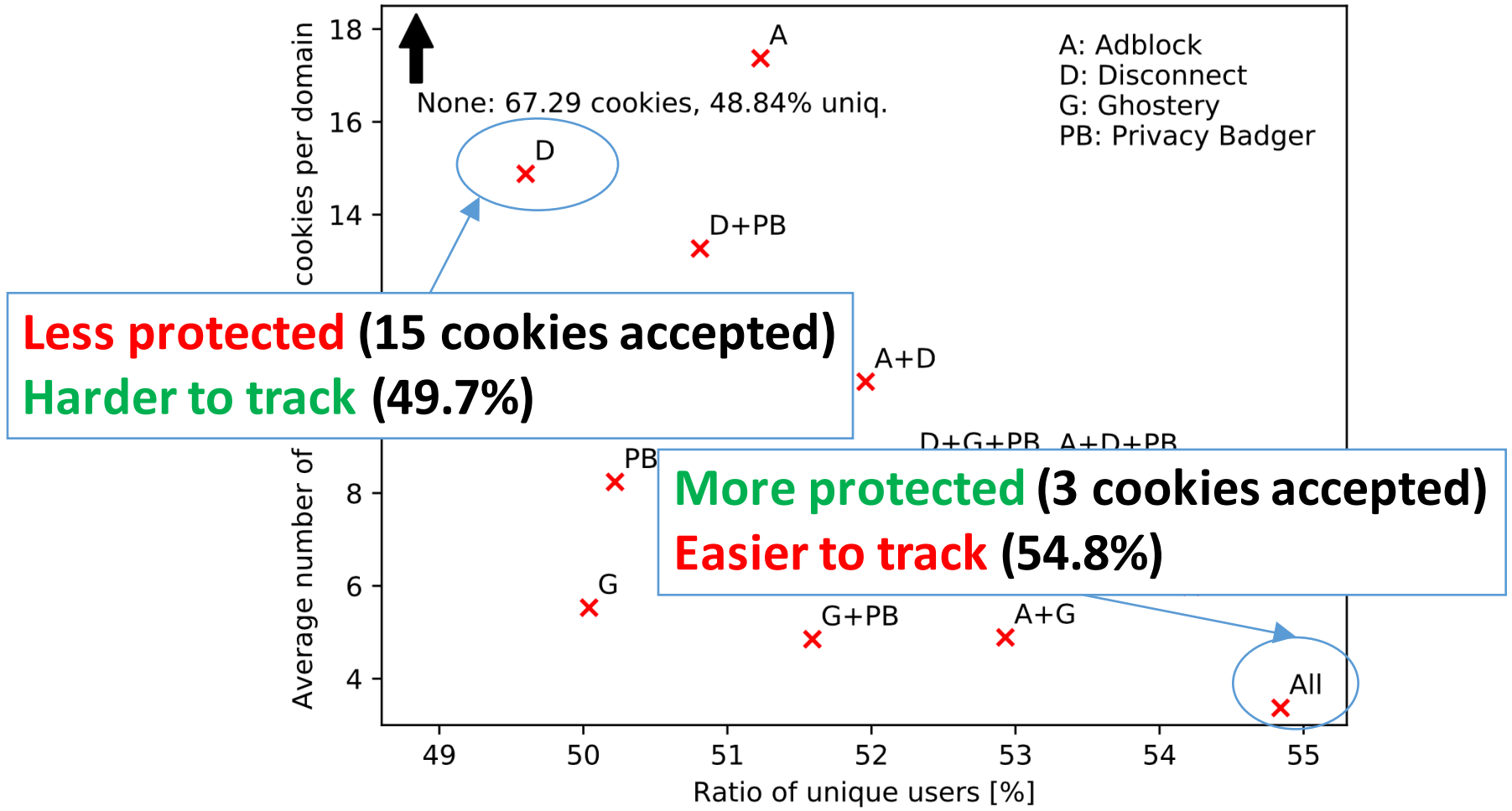
The dilemma of privacy extensions

- Privacy extensions **block some trackers**
- Privacy extensions **make a user more unique**

- What is the trade-off between **privacy gain** (some trackers are blocked) and **privacy loss** (user is more unique)?

G.G. Gulyás, D. F. Some, **N. Bielova** and C. Castelluccia. To Extend or not to Extend: on the Uniqueness of Browser Extensions and Web Logins. *WPES@ACM CCS 2018.*

Uniqueness of users vs. number of accepted third-party cookies



*4,000 pages crawled