

QUANTIFICATION OF ID SHARING WITHIN THE WEB ECOSYSTEM FROM THE EDGE

MARJAN FALAHRASTEGAR

HAMED HADDADI

STEVE UHLIG

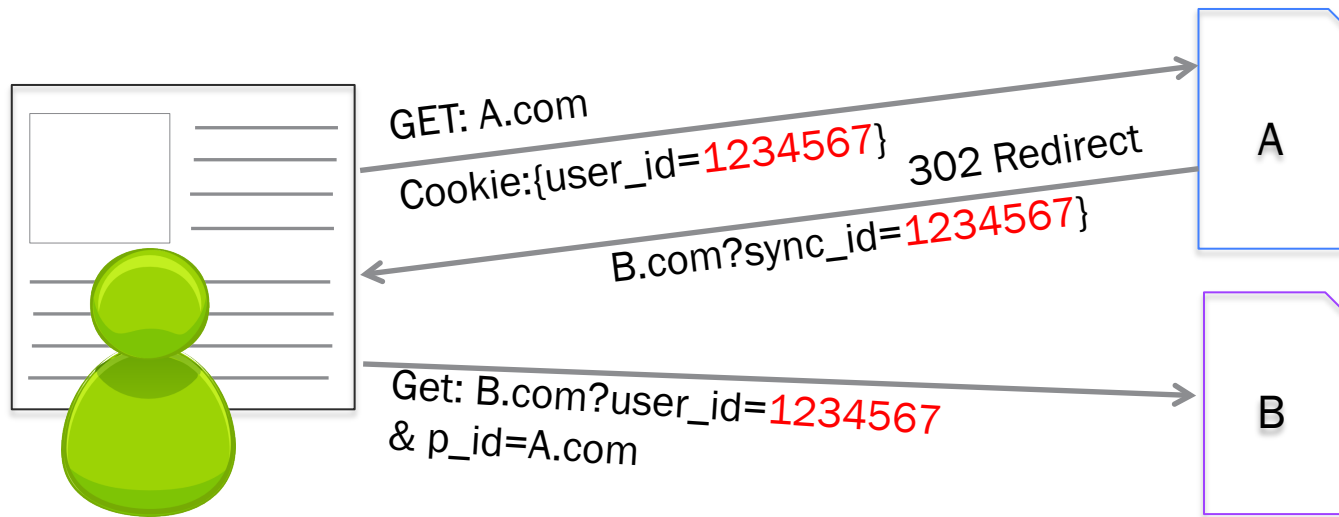
RICHARD MORTIER



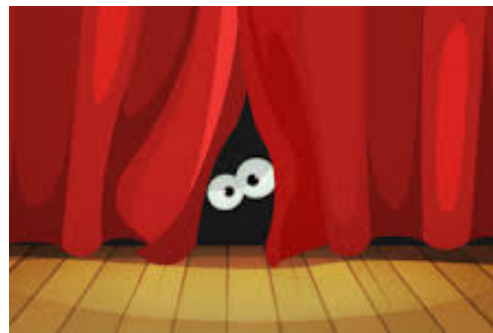
Queen Mary
University of London

USER ID SHARING

User ID sharing is a practice by which two different trackers **share** the IDs that they've assigned to **a specific user**.



The parties participating in ID sharing are able to **aggregate** the data collected by other parties in order to construct a **comprehensive profile** of users.



DATA COLLECTION

- We extended the **Lightbeam Firefox plug-in** to log all headers of HTTP requests and responses.
- In total we had **129 participants** from **22 countries** across the globe. All our data were obtained between 20 February 2015 until 1 April 2015.

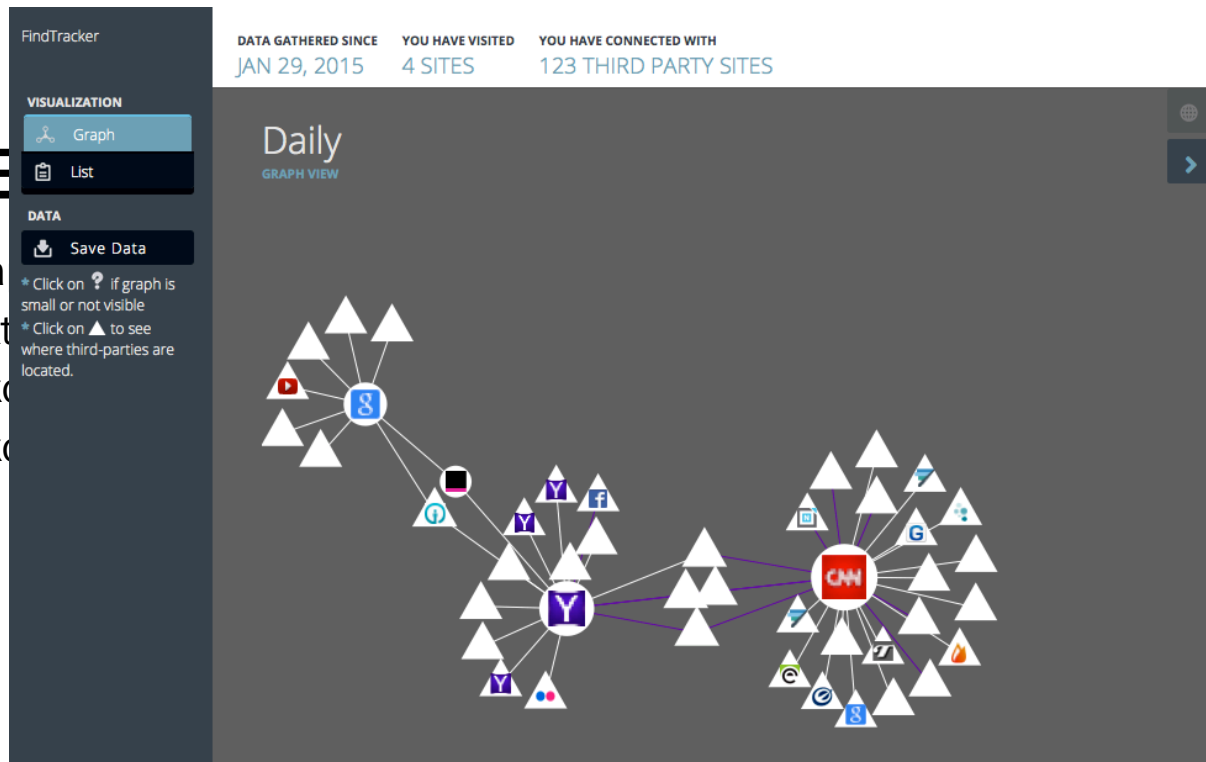
IDE

we a

✓ Ext

✧ Exc

✧ Exc



USER ID SHARING GROUPS

- 3,224 unique user IDs from 806 domains.
- 769 domains sharing unique user IDs with other domains.
- 660 unique ID sharing groups containing 2 to more than 8 domains.

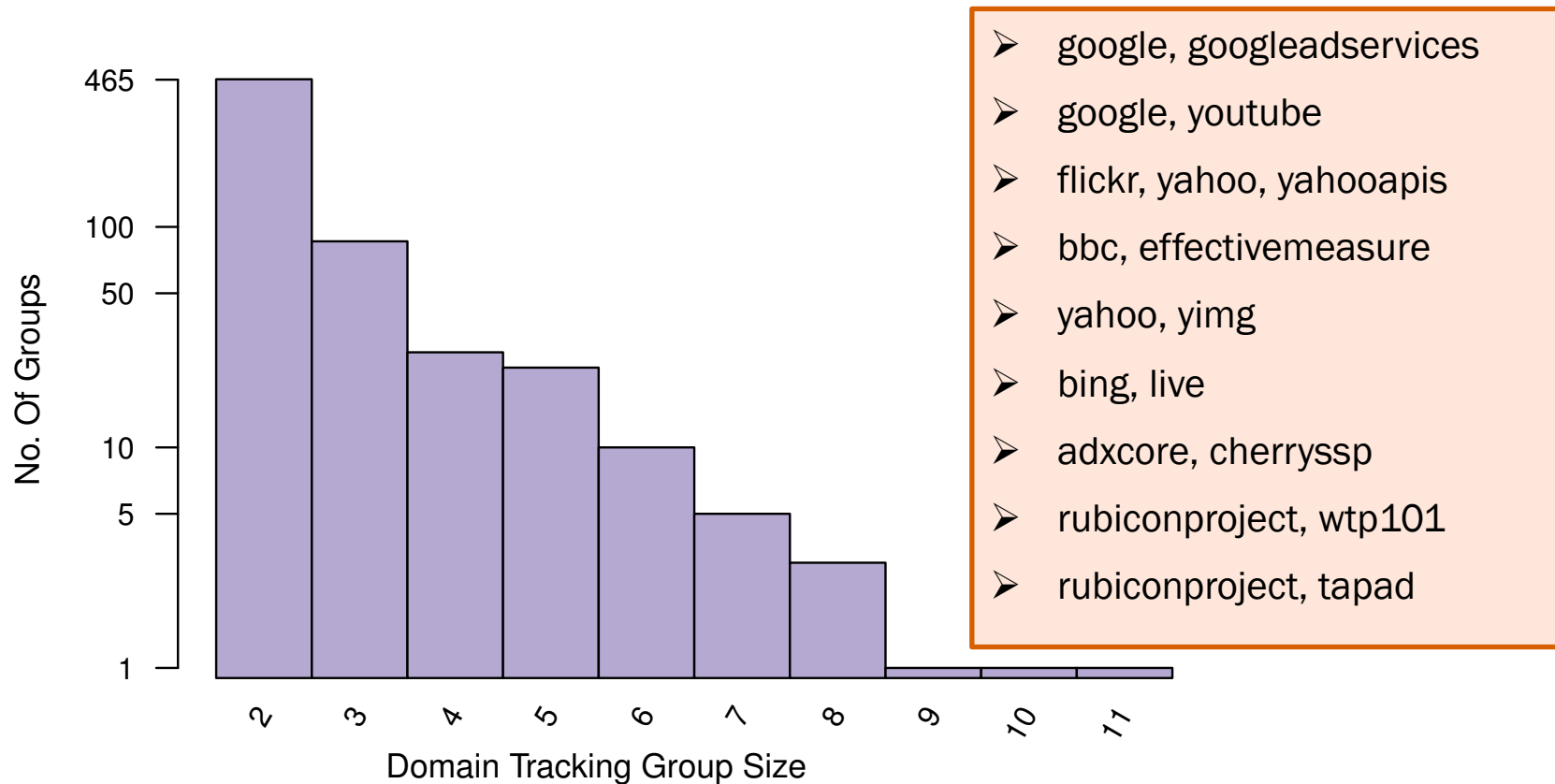


Fig 1. Size of sharing groups

ORGANISATIONAL SHARING

- ID sharing groups consist of multiple domains that may actually belong to the same organisation
- To identify the **parent organisations** of the domains, we used the **e-mail** addresses of the domains using *ns-lookup* Linux utility tool (querying their SOA (Start of Authority) record)

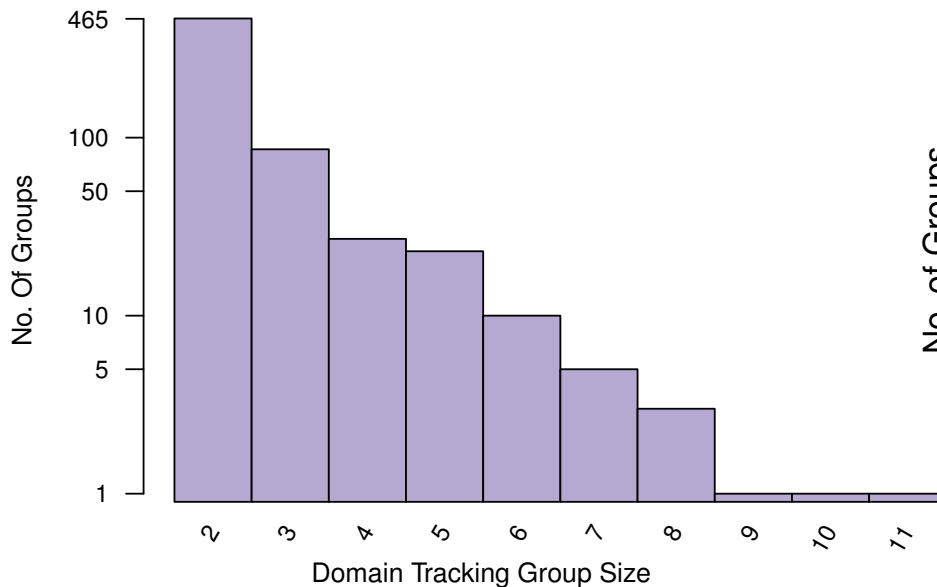


Fig 1. Domain ID sharing groups

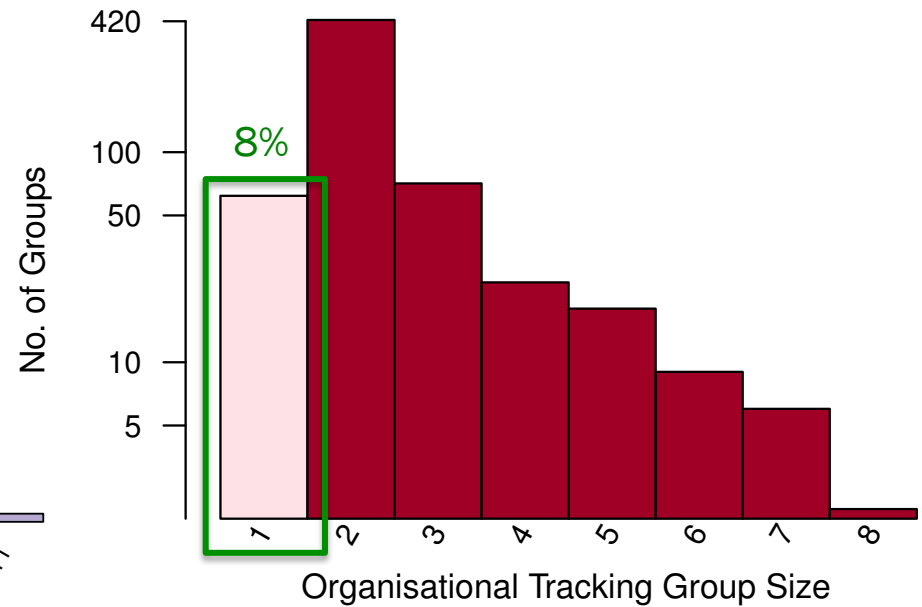


Fig 2. Organisational ID sharing groups

CROSS CATEGORIES SHARING

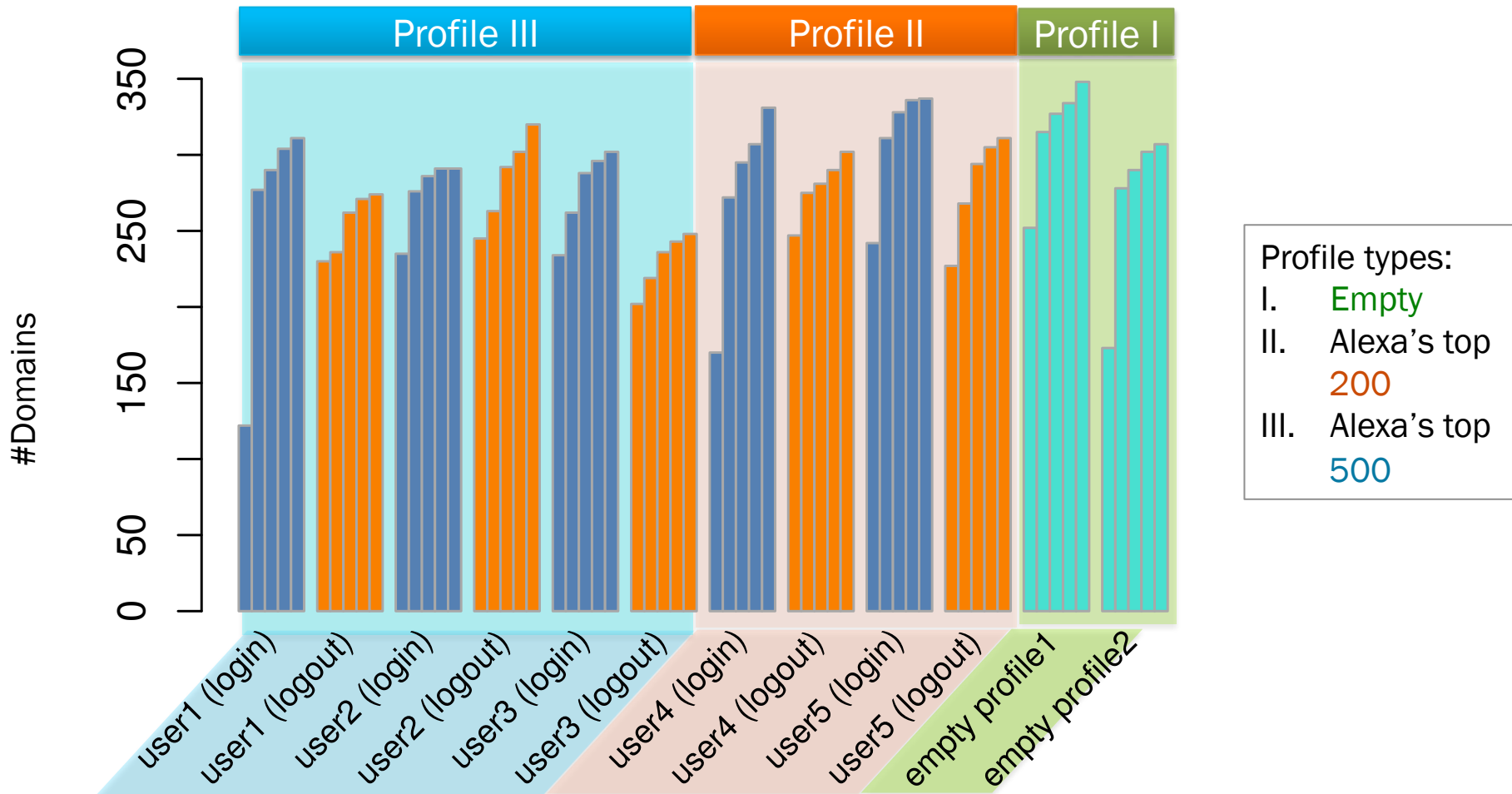
- Identified domain categories using the [Trend Micro Site Safety](#) service
- Categories related to [the advertisement ecosystem](#) (e.g, ad networks, analytics, ad exchanges) have the highest presence
- [Sensitive](#) domain categories such as health related ones participate in the ID sharing with domains related to [advertisement trackers](#) and [search engines](#).

search engines, web advertisements	c-cat.
search engines, streaming media	c-cat.
ad-tracker	w-cat.
search engines	w-cat.
ad-tracker, web advertisements	c-cat.
ad-tracker, internet infrastructure	c-cat.
ad tracker, photo searches, search engines	c-cat.
media, news	c-cat.
ad tracker, news	c-cat.

Table 2. Top 10 categories of ID sharing groups

EFFECT OF USER PROFILE

- 5 users with accounts on Google, Amazon, Twitter & eBay. 2 users without any account



CONCLUSION AND FUTURE WORK

- We identified a significant amount of ID sharing **across** different **organisations** and **categories**
- Further explore the relation between ID sharing groups and users' profile information.
- We cannot **directly observe** whether domains use these IDs to **merge** collected data from different sources. Considering the possibility of such practice, it is important to get additional insight about what ID sharing groups actually do through the user IDs.

MORE ON THIRD-PARTY TRACKERS:

- The Murky World of Third Party Web Tracking [MIT Technology Review]
- The Rise of Panopticons: Examining Region-Specific Third-Party Web Tracking [TMA2014]
- <http://www.eecs.qmul.ac.uk/~marjan/>