Seminar
Understanding Worldwide Private Information Collection on Android

Yun Shen, Pierre-Antoine Vervier, Gianluca Stringhini

Derrick Chen

Objectives

01

Part one

Introduce the reseach

02

Part two

Possible Solutions

03

Part three

Q&A

About the Research

Dataset

21 Months between 2018 – 2019
6 Billion Record
17.3 Million devices in **201** Countries and Regions
2.13 Million Apps (6.5 Million SHA2s)

Information Collection

Mobile app activity data collect from mobile security product and measure in several steps.
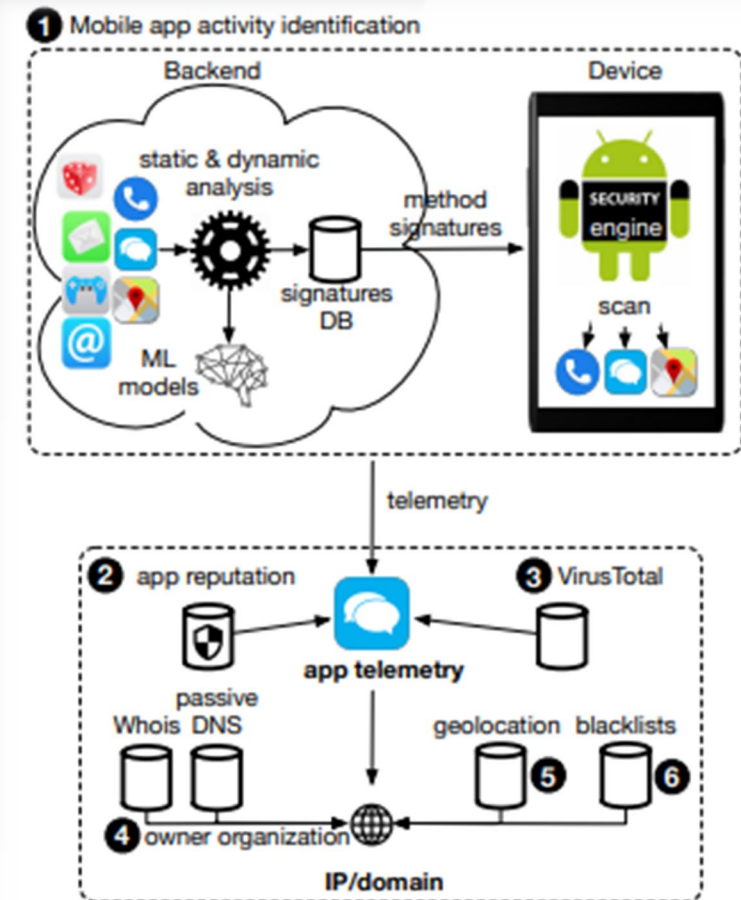

Fig. 1: Workflow of our measurement study.

PART 1

About Privacy Information Collection on Android

Types of Private Information

Private Information Collection on Android

Tracking

- SIM card, Operator info
- Device Info
- Setting Info
- Location

Activity and Social Profile

- Account Info
- Contact Info
- Call/SMS Log
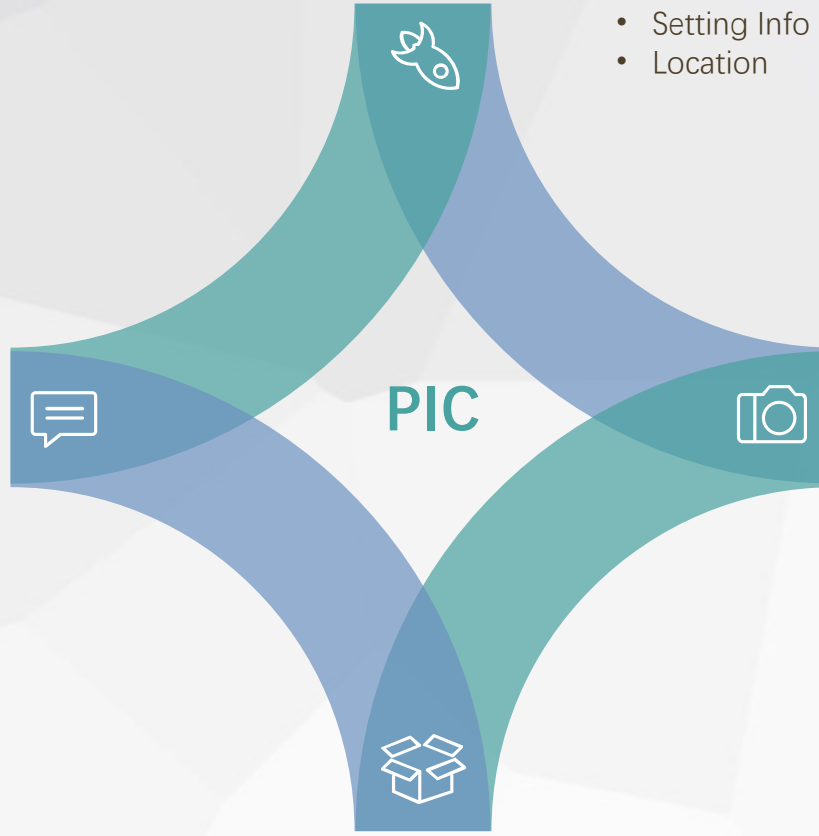- Calendar Info

PIC

Audio/ Video

- Audio Record
- Gallery
- Camera Info

Usage Preference

- Installed App
- Running App
- Browser Info

Pervasiveness of Private Information Collection

Private Information Collection on Android

Android Devices

87.2%

Send to 5 or more domains

United State collect

62%

Private information in the research

Over 156K
APPs

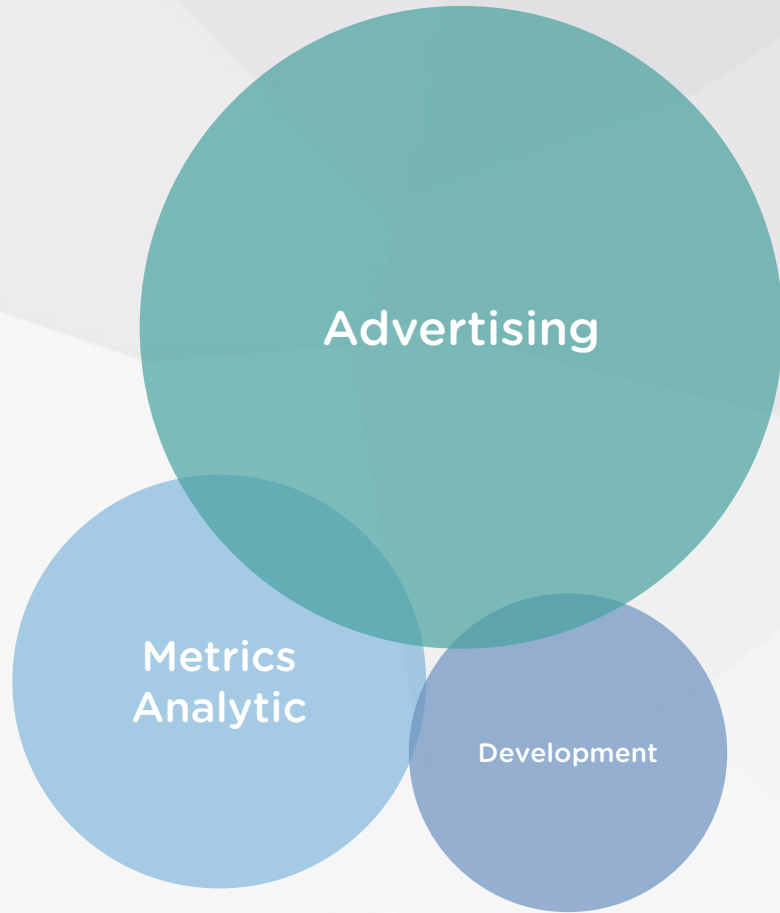Collect at least 5 unique categories of private information

13M
Devices

These apps cover 13 million devices which cover 74.9% of the dataset

Over 10K
Domain

There are 10,736 domains belong to 9,593 organisations

Top 20 Private Information Collection Domain

Private Information Collection on Android

Advertising

Metrics Analytic

Development

Purpose

Most of the domains (15/20) are use for advertising purpose

Data Interest

On average, a domain collect over 8 types of private information, like call/SMS log, identity info, location info and social network info

Geography

Certain PIC domains have a high reginal app presence .

PIC Domain Summary

Private Information Collection on Android

- Some of the actors who managed to get their libraries installed in many apps failed to have many users running them.

- Certain PIC domains consistently collect multiple types of private information from the devices

- Device penetration could be a better way to analyse than app presence

- Different regional domains targeting users in different continents, and collecting different types of private information.

- United State and China are the largest two countries hosting the PIC domains.

- Other countries host significantly fewer PIC domains

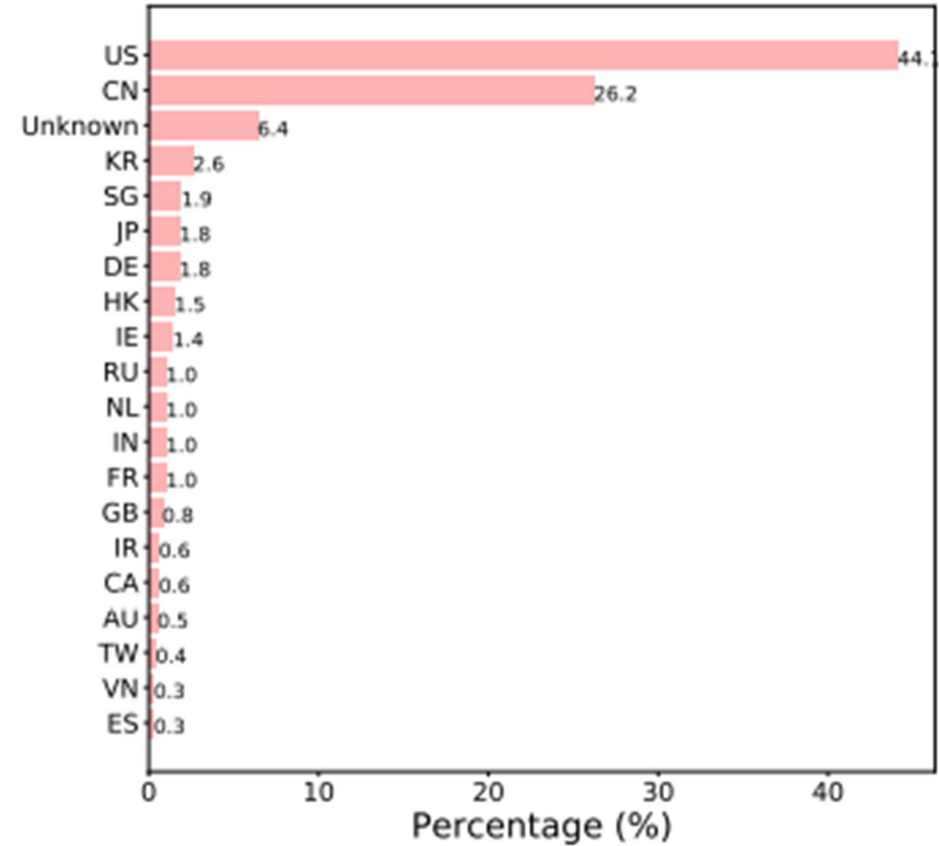- There are also 6.4% of the domain could not be recognised the destination

Fig. 7: Global top 20 countries ranked by the number of PIC domains hosted.
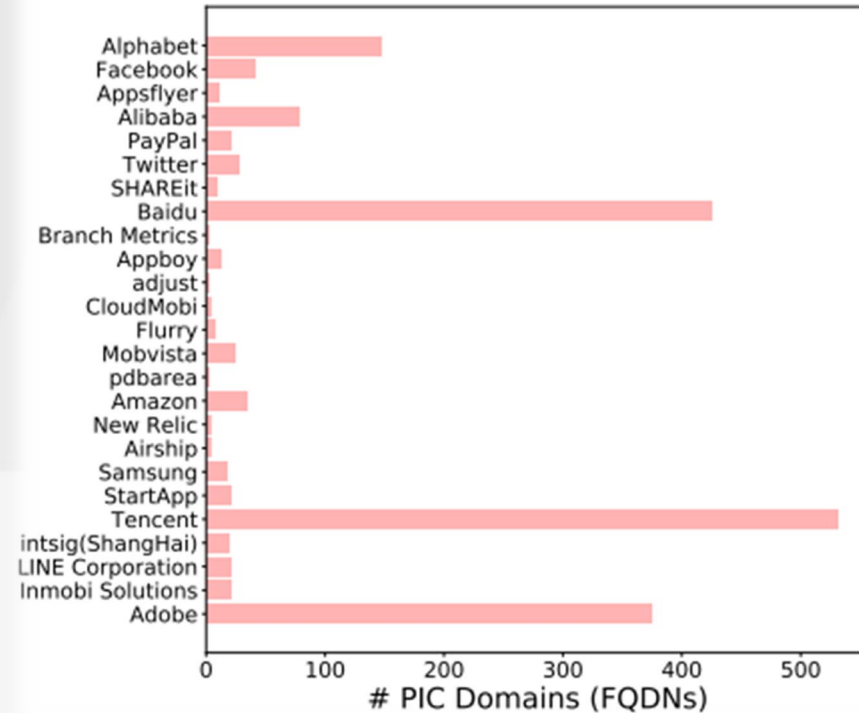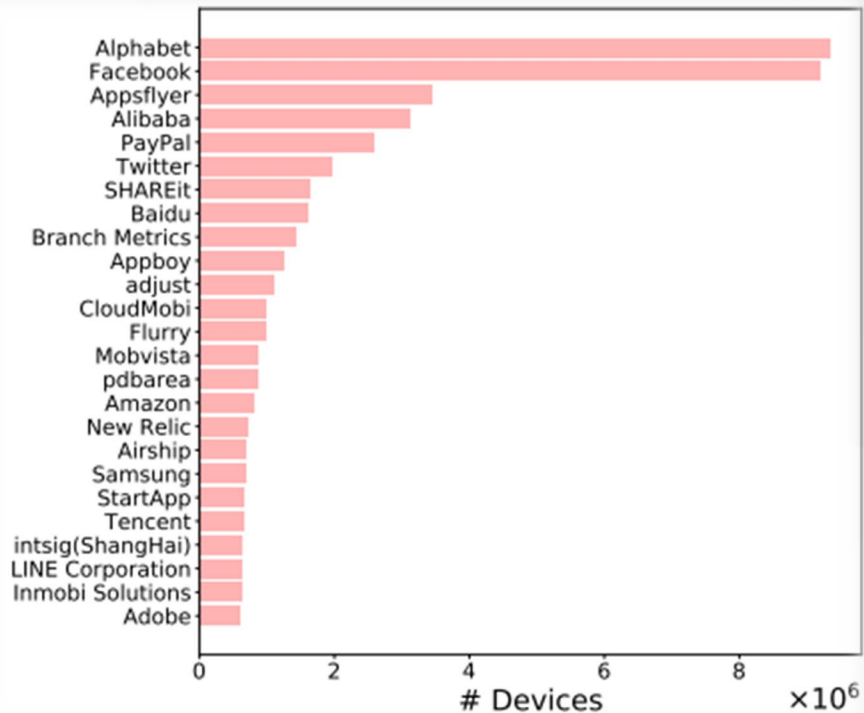
Private Information Destinations

Private Information Collection on Android

- United State and China are the largest two countries hosting the PIC domains.

- PIC domains hosted in the United States collect 62% of global private information flows.

- PIC domains hosted in the United States dominate the private information collection in the EU

- India generate 40% of data but only 1.28% terminating at the same country
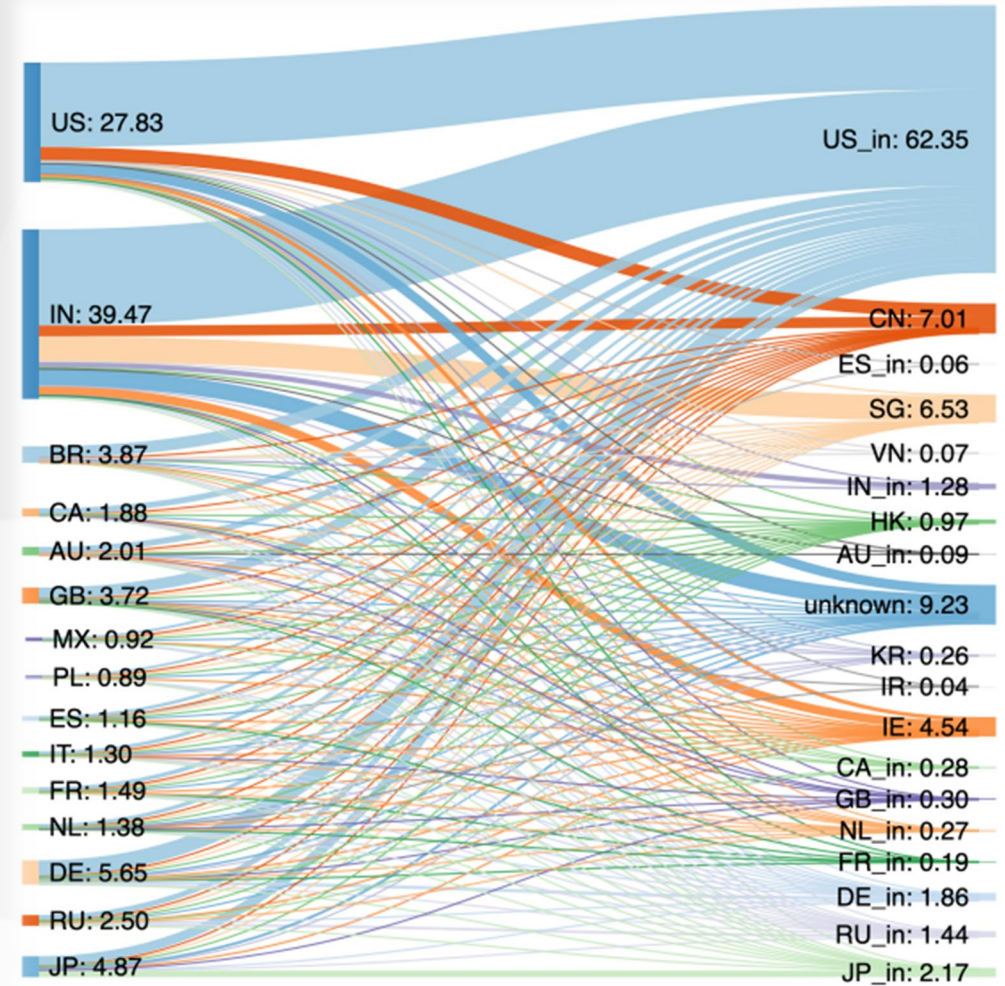
Data Processors & Controllers

Private Information Collection on Android

Global top 25 data controllers ranked by the fraction of devices they collect private information from.
These 25 data controllers collect private information from a total of 13.9M devices covering 80.2% of all devices used in this study.

PART 2

Possible Solution

Solution

Private Information Collection on Android

Authority

- EU announced GDPR
- Policymakers need further regulate how private information is used by and shared among the companies and
- Accountability

Protect Private Information

End-User

- Awareness of privacy protection

Android

- Since Android API 23, dangerous permissions added that request permission when app access restricted data
- Android API 29 Add further security measures to protect privacy

PART 3

Q&A

THANKS!

Reference

Shen, Yun, Pierre-Antoine Vervier, and Gianluca Stringhini, Understanding Worldwide Private Information Collection on Android, In Proceedings of Network and Distributed Systems Security Symposium, 2021.
https://www.ndss-symposium.org/wp-content/uploads/ndss2021_3B-3_24076_paper.pdf