

Bug Fixes, Improvements,... and Privacy Leaks: A Longitudinal Study of PII Leaks Across Android App Versions

Jingjing Ren*, Martina Lindorfert†, Daniel J. Dubois*,
Ashwin Rao‡, David Choffnes*, Narseo Vallina-Rodriguez§

*Northeastern University †UC Santa Barbara

‡University Of Helsinki §IMDEA Networks Institute And ICSI

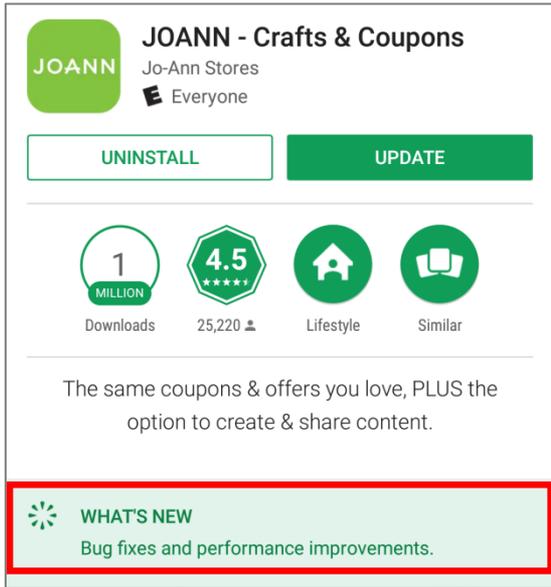
Sponsored by:



DATA
TRANSPARENCY
LAB

PRIVACYCON

Motivation



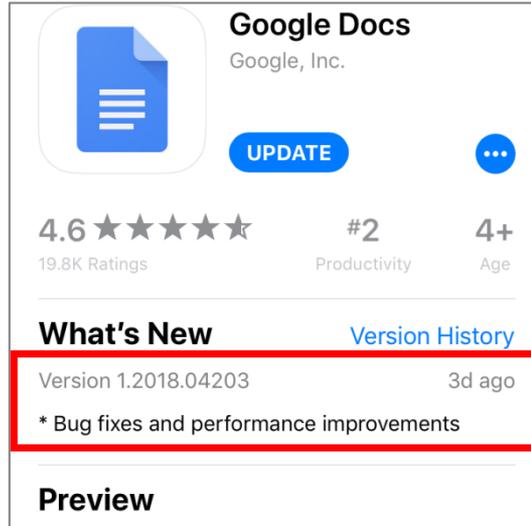
JOANN - Crafts & Coupons
Jo-Ann Stores
Everyone

UNINSTALL UPDATE

1 MILLION Downloads 4.5 25,220 Lifestyle Similar

The same coupons & offers you love, PLUS the option to create & share content.

WHAT'S NEW
Bug fixes and performance improvements.



Google Docs
Google, Inc.

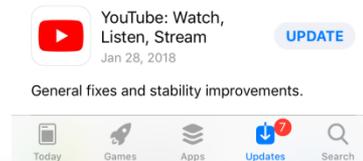
UPDATE

4.6 19.8K Ratings #2 Productivity 4+ Age

What's New [Version History](#)

Version 1.2018.04203 3d ago
* Bug fixes and performance improvements

Preview



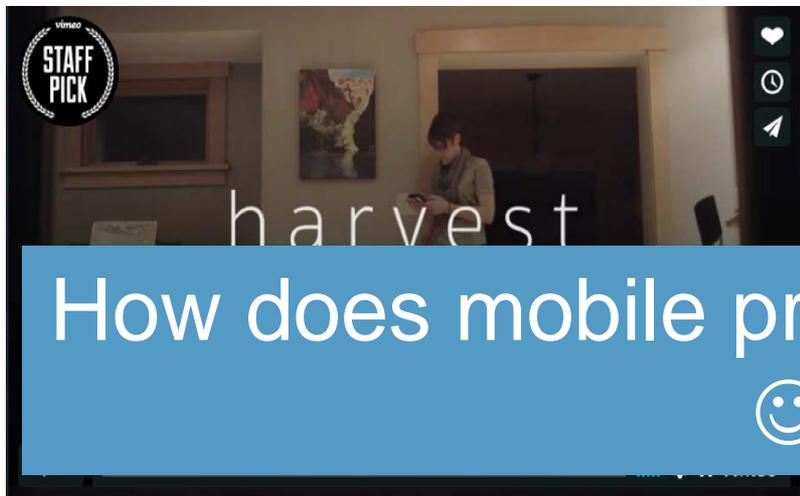
YouTube: Watch, Listen, Stream
Jan 28, 2018

UPDATE

General fixes and stability improvements.

Today Games Apps Updates Search

Evolution of Privacy in Mobile Devices



3,545 times in a week
JOANN

LOCATION LEAK FREQUENCY

How does mobile privacy evolve over time?

😊 or 😞

Nur
2016-01-09
2016-01-10
2016-07-25
2016-08-01
2016-08-13
2016-12-12
2017-06-06
Version

What Do I Mean by “Privacy” in This Work?

What information is shared?



**Personally Identifiable Information
(PII)**

Tracking ID, User information
Location, Contact, ...

How is it being shared?



Transport security

Plaintext
or
Encrypted

Where is it going?



Destination party:

First (app owner)
or
Third (advertising & analytics)

Why do we care?

- Data Aggregation
- Eavesdropping Attack

Experiment Methodology

Selecting apps

- 512 Android apps, 7,665 unique versions (APKs), 8 years

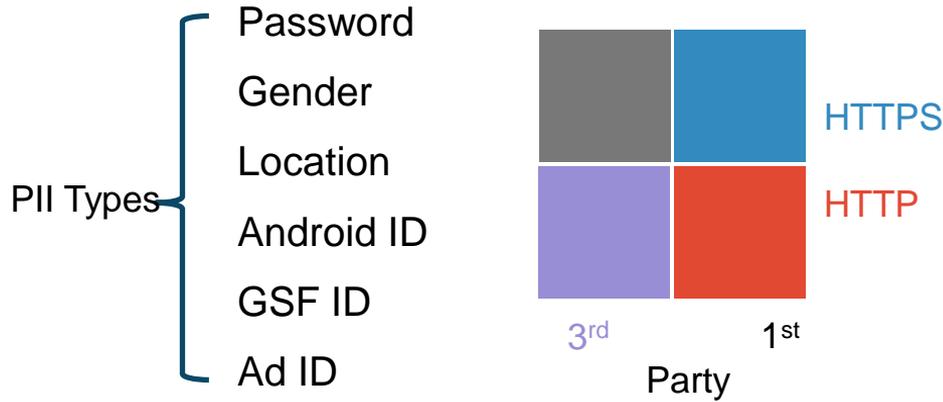
Interaction with apps

- manual is good, ... but not scalable 😞
- Automated and scripted interaction: Monkey!
 - randomly generated events with good coverage
 - ***login and replay*** across the versions

MITM proxy to intercept HTTP(S) traffic

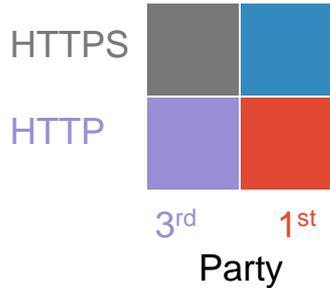
PII detection: ReCon + manual validation

Privacy Attributes

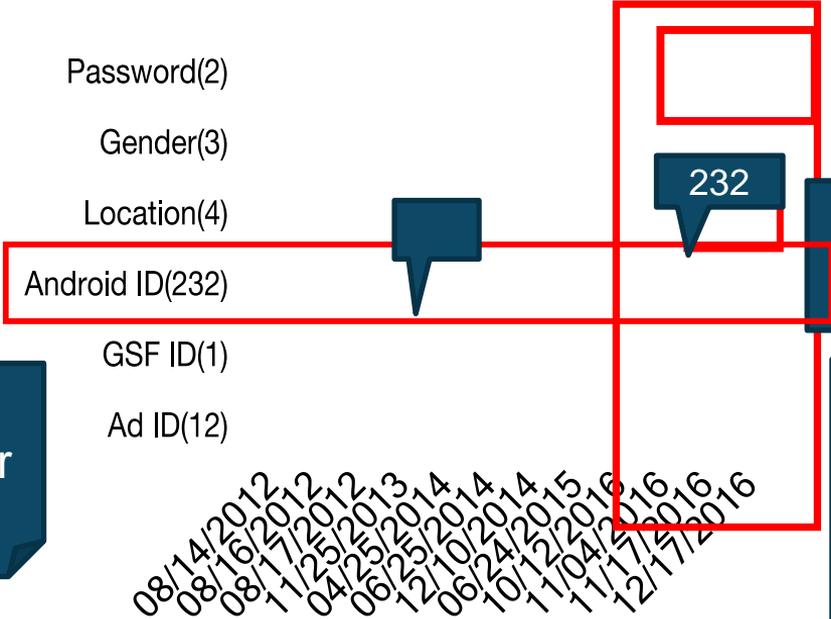


08/14/2012
08/16/2012
08/17/2012
11/25/2012
04/25/2013
06/25/2014
12/10/2014
06/24/2014
10/12/2015
11/04/2016
11/17/2016
12/17/2016

Case study: Pinterest



Increased frequency for Android ID



Sends password to a third party in 2 out of 12 versions

HTTPS is used on Jan. 5, 2017: disclosed Feb. 7, 2017: fixed

More types (gender, location, android ID etc.) are leaked

High variance in privacy risks across versions

Aggregate Trends in Privacy

- PII leaks can *change* substantially across versions
- HTTPS Adoption is *slow*
Takes apps years to adopt HTTPS for supported domains
- Third-party tracking is *pervasive* and *broad*
Evolving tracking systems: hard-coded ID => resettable Ad ID
Permanent linkage between tracking ID and PII: >100 domains

Conclusion

- Privacy has *worsened* over time
 - PII leaks can *change* substantially across versions
 - HTTPS Adoption is *slow*
 - Third-party tracking is *pervasive* and *broad*
- Need for continuous monitoring
 - ReCon, Lumen, AntMonitor etc.

