ReCon: Revealing and Controlling PII Leaks in Mobile Networks
Jingjing Ren, David Choffnes, Northeastern U, Ashwin Rao, U of Helsinki, Martina Lindorfer, SBA Research, Arnaud Legout, INRIA Sophia-Antipolis

MOTIVATION
- Mobile devices: ubiquitous and connected to Internet
- Personally Identifiable Information (PII) leaks are pervasive

Key questions:
- What information is leaked?
- How is it sent out?
- Who receives this information?
- What can users do to control it?

How to detect PII leaks?
- At the OS, e.g. information flow analysis (IFA)
  - Doesn't cover everything, hard to scale
- Simpler approach: Focus on network traffic
  - Independent of OS, app store

Our approach: Find PII in network traffic
- Machine Learning classifiers to detect leaks
- Software middleboxes to control leaks
- Works today on all major platforms

Flow contains
- /getImage.php
- auid
- conn
- /getImage.php
- uid

An example decision tree classifier

SYSTEM
- Detect PII Leaks
- Allow user feedback
- Block/modify PII
  - Coarsen locations
  - Anonymize

USER STUDY
- IRB-approved
- 213 iOS, 225 Android (9/2017)
  - 30,289 PII leaks
  - 200 credential leaks, 168 verified

- Identified 30 apps exposing passwords in plaintext or sending to third parties
  - Used by millions (Pinterest, Grubhub, Match, Epocrates etc.)
  - Responsibly disclosed
  - 17 have fixed the problem

SUMMARY
- Need for improved transparency/control over PII
- ReCon approach addresses this
  - Learn what information is being leaked
  - Crowdsourcing to determine correctness/importance
  - Allow users to block/change what is leaked

http://www.harvest-documentary.com/

Sponsored by

Data Transparency Lab
Homeland Security
Northeastern University
Cybersecurity and Privacy Institute