50 Ways to Leak Your Data

By Byron Downey, Jessica Rock and Thomas Petersen

Introduction

Research paper in focus:

50 Ways to Leak Your Data: An Exploration of Apps' Circumvention of the Android Permissions System

Reardon, J.   Feal, A.   Wijesekera, P.   On, A.,   Vallina-Rodriquez, N.,   Egelman, S.

Paper aims to:

- Boost enforcement through regulators
- Improve implementation in industry
- Educate users

Android Permission System

- Principle of least privilege
  - ‘An entity should only have the minimum capabilities it needs to perform its task.’
- Application must:
  - Ask user for permission
  - Not be able to access information without permission

Concerns:

- Access requests are not regulated
- Applications are getting around it

Main Methods of Circumvention

- Side Channel
  - Exposes path to resource outside security mechanism
    - Either flaw in design of security mechanism
    - Or flaw in the implementation of the design
  - Eg. cryptography - power usage of hardware can leak secret key

- Covert Channels
  - More intentional
  - Applications cooperate (1 has access, 1 doesn’t)
  - Eg. alternating CPU load (high/low) to send binary message

Common Targets:

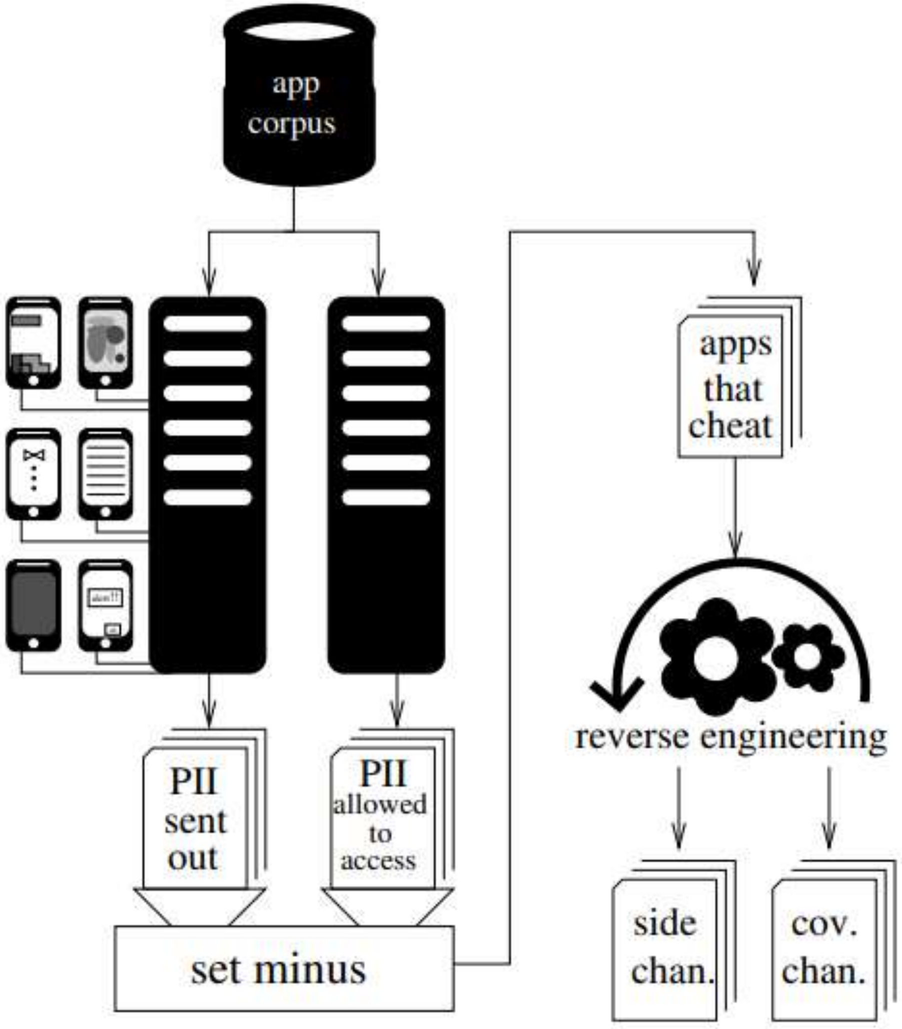
- International Mobile Equipment Identifiers (IMEI’s)
- Device Media Access Control (MAC) addresses

Examples of Circumvention

- Obtaining MAC addresses of WiFi base stations - location data
- Unity was obtaining device MAC addresses - location data
- SD cards as covert channels
- Picture metadata as location data

Testing Environment and Analysis Pipeline

- The authors of the paper set up an automated pipeline to analyse apps
- Evidence of covert- and side-channel usage in 252,864 versions of 88,113 different Android apps
- Analysed at the OS level as well as the network level by decrypting traffic



Analysis Pipeline [1]

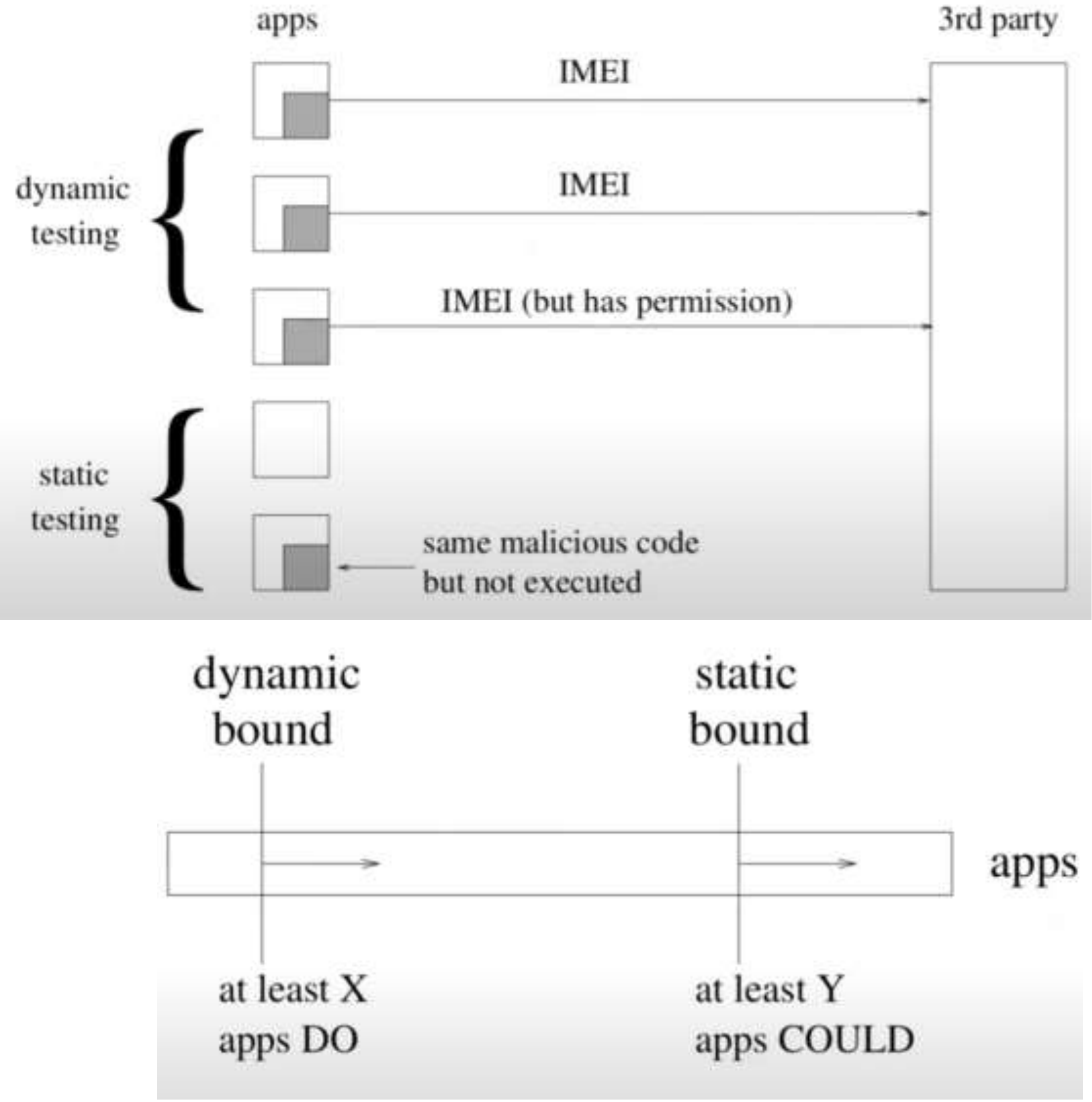
Static Analysis

- Static analysis scans for execution flows in apps' code to see where privacy might be violated
- Issues with static analysis:
  - Some apps may choose not to violate user privacy even when they could
  - Code obfuscation could cause missed instances
  - Dynamically loaded code cannot be analysed this way

Static analysis [2]

Dynamic Analysis

- Gather collection of phones
- Setup the pipeline to:
  - Automatically download the apps
  - Interact with them via the UI/Application Exerciser Monkey for 10 minutes
  - Monitor via the kernel when certain resources are Accessed
  - Intercept and decrypt network traffic



Dynamic Analysis [3]

Results

- IMEI
- Network MAC
- Router MAC
- Geolocation

IMEI (International Mobile Equipment Identity)

- Salmonads
    - 5 applications were using a covert channel to access the IMEI
    - The malicious apps were downloaded at least 17.6 million times
- Baidu
    - Stored the IMEI in an AES-encrypted json file
        - The AES key was baidu cid 2012 12 03
    - This json file was encrypted in base-64 on the SD card
    - 8 applications were using a covert channel to access the IMEI
    - The malicious apps were downloaded at least 700 million times

Device IMEI [4]

Network MAC Address

- Unity was able to obtain the MAC address of the device
- Normally protected by the ACCESS_NETWORK_STATE permission
- The exploit used the IOCTL in the native Unity library
- 42 applications were exploiting this, but 12, 408 apps had the capability to

Unity Logo [5]

Router MAC Address

- ARP table
  - /proc/net/arp was not protected properly, and anyone could read the cache file
  - OpenX is once such company, and utilised the exploit through a function called ‘getDeviceMacAddressFromArp’
    - Would first ask for the correct permission
    - If denied, it would acquire the router MAC through the aforementioned exploit
- UPnP protocol
  - Requested the igd.xml (The internet gateway device configuration) file
  - Three Peel’s smart-remote control apps utilised this

“The fact that the router is providing this information to devices hosted in the home network is not a flaw with Android per se. Rather it is a consequence of considering every app on every phone connected to a WiFi network to be on the trusted side of the firewall.”

Geolocation

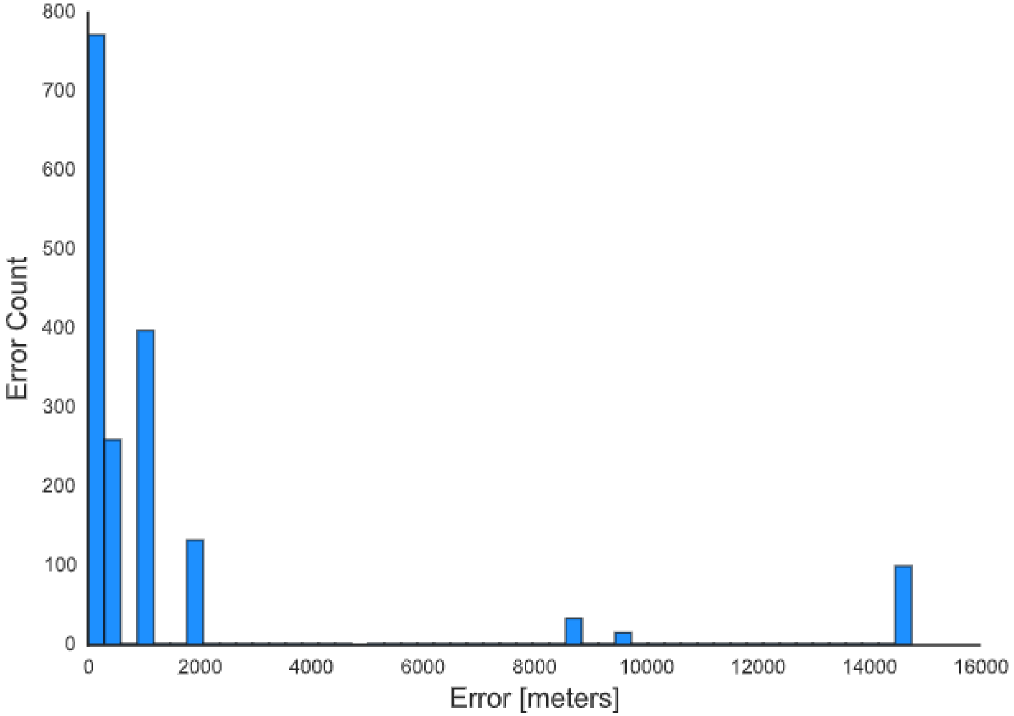
- 70 applications that would send its location to 45 different domains
  - Hypothesised that this was by ad mediation services that would provide the
  - E.g., IP geolocation
- Shutterfly
  - Would send images including the EXIF metadata to its own server
  - Metadata included the latitude and longitude of where the photo was taken
  - Any application with photo library read permissions would be able to read this metadata

Geolocation [6]

Related Work

Related works:

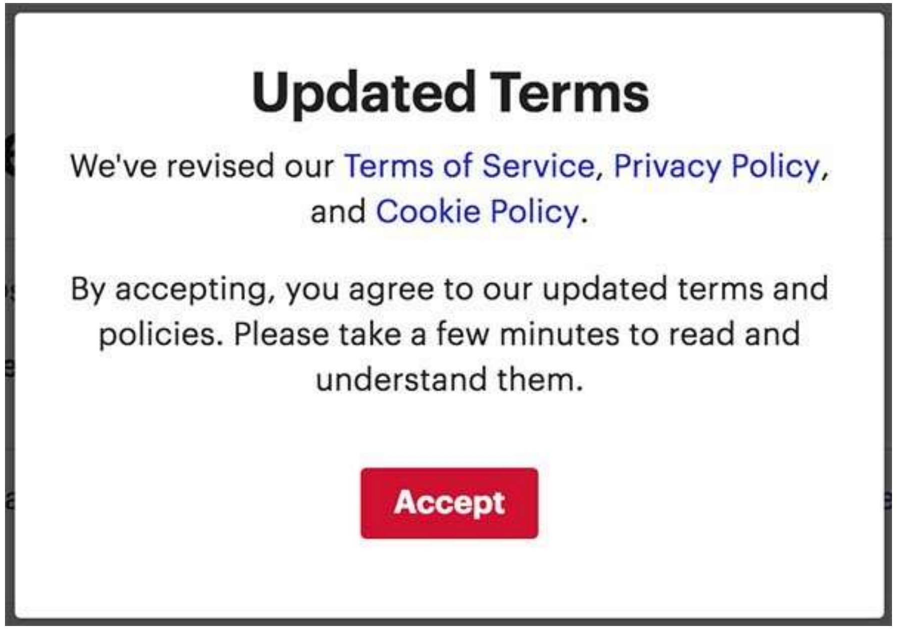
- Vibrations could be used as a covert channel between devices
- Data could be shared through UNIX sockets and external storage - demonstrated in the 50 Ways to Leak Your Data paper
- Users can be located by monitoring the power consumption of their phones
- Devices could be fingerprinted using the factory calibration of their sensors



Error showing that 90% of errors are within 1km - tracking users based on their phone's power consumption. [7]

Privacy Expectations

- The notice and consent framework governs a lot of countries' laws on privacy such as:
  - U.S law
  - The EU's General Data Protection Regulation
  - Australia's Privacy Act
- French privacy commission, CNIL, fined Google 50 million Euros for breaching the General Data Protection Regulation
- Australia's Privacy Act - companies must "manage personal information in an open and transparent way".



Updated Terms

We've revised our Terms of Service, Privacy Policy, and Cookie Policy.

By accepting, you agree to our updated terms and policies. Please take a few minutes to read and understand them.

Example Privacy Policy dialog box [8]

Limitations, Future Work, Recommendations

- Paper only shows the lower bound for privacy violations due to issues with:
  - Obfuscated data transmissions
  - Certificate pinning by apps
  - Android Exerciser Monkey only similar to a human 60% of the time
  - Android Exerciser Monkey could not proceed if apps required login or specific input like a CAPTCHA
- Even at the time of the paper, Google had been working on fixing some of the issues raised
- In the 3 years since, Google has likely fixed even more issues

Mobile App Testing [9]

References

[1]J. Reardon, A. Elazari Bar On, Á. Feal, N. Vallina-Rodriguez, P. Wijesekera and S. Egelman, "50 Ways to Leak Your Data: An Exploration of Apps’ Circumvention of the Android Permissions System", *USENIX*, 2019. Available: https://www.usenix.org/system/files/sec19-reardon.pdf. [Accessed 21 May 2022].

[2]R. Bellairs, *Static Analysis*. 2020.

[3]USENIX, *USENIX Security '19 - 50 Ways to Leak Your Data: An Exploration of Apps'*. 2019.

[4]Wondershare, *track-phone-using-imei-1*. 2020.

[5]1000Logos, *Unity Logo*. 2022.

[6]Ulistic, *geolocation*. 2022.

[7]Y. Michalevsky, G. Nakibly, A. Schulman, G. Arumugam Veerapandian and D. Boneh, "PowerSpy: Location Tracking using Mobile Device Power Analysis", *USENIX*, p. 9, 2015. Available: https://arxiv.org/abs/1502.03182. [Accessed 21 May 2022].

[8]Privacy Policy Generator, *Creating Agreement Checkboxes*. 2020.

[9]Freestocks, *Mobile app testing*. 2021.

[10]"Dynamic Linking vs. Dynamic Loading", *Baeldung*, 2021. [Online]. Available: https://www.baeldung.com/cs/dynamic-linking-vs-dynamic-loading#:~:text=Dynamic%20loading%20is%20the%20technique,used%20to%20implement%20software%20plugins. [Accessed: 21- May- 2022].

[11]"French Highest Administrative Court Upholds 50 Million Euro Fine against Google for Alleged GDPR Violations", *Privacy & Information Security Law Blog*, 2020. [Online]. Available: https://www.huntonprivacyblog.com/2020/06/23/french-highest-administrative-court-upholds-50-million-euro-fine-against-google-for-alleged-gdpr-violations/. [Accessed: 21- May- 2022].

[12]"Part 5: Notice and consent", *Home*, 2021. [Online]. Available: https://www.oaic.gov.au/privacy/the-privacy-act/review-of-the-privacy-act/privacy-act-review-issues-paper-submission/part-5. [Accessed: 21- May- 2022].