Stay Off My Private Data:
A Framework to Examine Mobile App Privacy Claims

Bai Guangdong@UQ | Zhang Qing@ByteDance | Wang Zeyu@ByteDance
01 About Us
02 Background
03 The Framework
04 Summary & Recommendations
05 Q&A
About Us
Bai Guangdong
Senior Lecturer from The University of Queens & Australia
Research on mobile security & protocol analysis

Zhang Qing
Senior security researcher from ByteDance
Research on Mobile security & payment security

Wang Zeyu
Senior security researcher from ByteDance
Research on Android application security & privacy protection
Background
This video clip is from the film Jexi(https://en.wikipedia.org/wiki/Jexi).
Jexi is a 2019 romantic comedy film written and directed by Jon Lucas and Scott Moore.
Applications are collecting your data

### A variety of data types
- Location
- Device ID
- SIM info
- Browsing history

### A variety of purposes
- Analytics
- App functionality
- Product personalization
- Third-party advertising

---

#### APP A
**“Data Linked To You”**
- **Analytics**
  - Purchases
  - Financial info
  - Location
- **Identifiers**
  - User ID
  - Device ID
- **Usage Data**
  - Product interaction
  - Diagnostics
- **Other Data**
  - Other Data Types

#### APP B
**“Data Linked To You”**
- **Analytics**
  - Purchases
  - Financial info
  - Location
- **Identifiers**
  - User ID
  - Device ID
- **Usage Data**
  - Product interaction
  - Diagnostics
- **Other Data**
  - Other Data Types

#### APP C
**“Data Linked To You”**
- **Analytics**
  - Purchases
  - Financial info
  - Location
- **Identifiers**
  - User ID
  - Device ID
- **Usage Data**
  - Product interaction
  - Diagnostics
- **Other Data**
  - Other Data Types

#### APP D
**“Data Linked To You”**
- **Analytics**
  - Purchases
  - Financial info
  - Location
- **Identifiers**
  - User ID
  - Device ID
- **Usage Data**
  - Product interaction
  - Diagnostics
- **Other Data**
  - Other Data Types

---

### Third-Party Advertising

- **APP A**
- **APP B**
- **APP C**
- **APP D**
Data protection regulations are global
European General Data Protection Regulation (GDPR)

- GDPR legislation came into force on May 2018 in European Union (EU) countries

- A major update to the EU Data Protection Directive (95/46/EC) (DPD-95) introduced in 1995

- Covers six aspects

GDPR Principles

1. Personal information shall be collected for specified, explicit and legitimate purposes.
2. Personal information shall be processed in an appropriate manner to maintain security.
3. Personal information shall be adequate, relevant, and limited to what is necessary.
4. Personal information shall be accurate and, where necessary, kept up to date.
5. Personal information shall be processed in an appropriate manner to maintain security.
6. Personal information shall be processed lawfully, fairly, and in a transparent manner.

US California Consumer Privacy Act (CCPA) & US Children's Online Privacy Protection Rule (COPPA)

Similar to GDPR, with minor differences in scope, key definitions, legal basis, rights & enforcement.

**Scope**
- GDPR: “an identified or identifiable natural person”
- CCPA: “a natural person who is a California resident.”

**Region**
- GDPR: both “entities or organizations” “established” or “not established” “in the EU”
- CCPA: “doing business in California.”
Google Play's Requirements for Privacy Policies

4.8
• You agree that if You make Your Products available through Google Play You will protect the privacy & legal rights of users. If the users provide You with, or Your Product accesses or uses, usernames, passwords, or other login information or personal information, You agree to make the users aware that the information will be available to Your Product, & You agree to provide legally adequate privacy notice & protection for those users. Further, Your Product may only use that information for the limited purposes for which the user has given You permission to do so. If Your Product stores personal or sensitive information provided by users, You agree to do so securely & only for as long as it is needed. However, if the user has opted into a separate agreement with You that allows You or Your Product to store or use personal or sensitive information directly related to Your Product (not including other products or applications), then the terms of that separate agreement will govern Your use of such information. If the user provides Your Product with Google Account information, Your Product may only use that information to access the user's Google Account when, & for the limited purposes for which, the user has given You permission to do so.

Data Collection & Storage
• What data is collected?
• What the data is used for?
• Who is the data shared with?
• How is the data is protected?

Google Play’s Policy on User Data

User Data
You must be transparent in how you handle user data (e.g., information collected from or about a user, including device information). That means disclosing your app’s access, collection, use, & sharing of the data, & limiting the use of the data to the purposes disclosed. In addition, if your app handles personal or sensitive user data, please also refer to the additional requirements in the "Personal & Sensitive Information" section below. These Google Play requirements are in addition to any requirements prescribed by applicable privacy & data protection laws.
Personal & Sensitive Information

Personal & sensitive user data includes, but isn't limited to, personally identifiable information, financial & payment information, authentication information, phonebook, contacts, device location, SMS & call related data, microphone, camera, & other sensitive device or usage data. If your app handles sensitive user data, then you must:

• Limit your access, collection, use, & sharing of personal or sensitive data acquired through the app to purposes directly related to providing & improving the features of the app (e.g., user anticipated functionality that is documented & promoted in the app's description in the Play Store). Apps that extend usage of this data for serving advertising must be in compliance with our Ads Policy.
• Post a privacy policy in both the designated field in the Play Console & within the app itself. The privacy policy must, together with any in-app disclosures, comprehensively disclose how your app accesses, collects, uses, & shares user data. Your privacy policy must disclose the types of personal & sensitive data your app accesses, collects, uses, & shares & the types of parties with which any personal or sensitive user data is shared.
• Handle all personal or sensitive user data securely, including transmitting it using modern cryptography (for example, over HTTPS).
• Use a runtime permissions request whenever available, prior to accessing data gated by Android permissions.
• Not sell personal or sensitive user data.
<table>
<thead>
<tr>
<th>Privacy policy common themes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Our policy on children’s data</td>
</tr>
<tr>
<td>What we do with the data we collect</td>
</tr>
<tr>
<td>Choices you have about your data</td>
</tr>
<tr>
<td>How &amp; when we share your data</td>
</tr>
<tr>
<td>How long we retain your data</td>
</tr>
<tr>
<td>How to download your data</td>
</tr>
</tbody>
</table>
The Framework
Motivation

When conducting an application privacy compliance audit, how do we ensure that testing covers the majority, if not all privacy compliance issues?
Challenges

From regulations to best practice:
- Regulations are abstract, theoretical, & complex.
- Regulators mainly focus on enacting the laws & demand reliable regulation/auditing technologies.

From passive to active:
- Privacy challenges may often be remediated in a passive way of patching when vulnerabilities are discovered, like a whack-a-mole game.
- May lack a systematical approach to actively enforce the protection proactively.

From research to application:
- Many privacy incidents are exposed by researchers, where manual efforts are needed.
- We summarize the complex privacy security audit methodology in a systematic framework based on our experience with a large number of app privacy security audits.
Global view & guidelines

• Global view: take the entire data lifecycle into consideration.
• Guideline: break it down into an enforceable check list.
Our framework

- **Logic level**: allows decision makers to have a big picture view of potential scenarios where privacy issues may occur.
- **Enforcement level**: guides auditors/analysts to make inspections in an efficient & comprehensive way.
Logic level: data lifecycle

Including whether apps provide users choice & control over how/when their data is acquired, accessed, retained, stored, erased, & localized.
# Logic level break-down

<table>
<thead>
<tr>
<th>Type</th>
<th>Item</th>
<th>Description</th>
<th>Test method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Illegal collection of user data</td>
<td>Collection or use of data is beyond scope of user authorization.</td>
<td>Actual collection of users' personal data by app goes beyond types of data listed in the privacy statement.</td>
<td>Manual work</td>
</tr>
<tr>
<td></td>
<td>Collection of personal data without user consent or authorization.</td>
<td>Collection personal data without the user's consent, or after the user has expressly refused to do so.</td>
<td>Automation tools &amp; manual work</td>
</tr>
<tr>
<td></td>
<td>Purpose &amp; scope of collection &amp; use of personal information are not stated.</td>
<td>Rules for personal data collection &amp; use of app do not list the purpose, method, &amp; scope of personal data collection &amp; use or relevant permissions individually.</td>
<td>Manual work</td>
</tr>
<tr>
<td></td>
<td>Compulsory granting permissions related to user data, or requesting permission related to user info without using it.</td>
<td>Continued request for permissions, or requesting permissions but not using related functions.</td>
<td>Manual work</td>
</tr>
<tr>
<td>Illegal use of user data</td>
<td>Force users to use directed push.</td>
<td>No option to turn off directional push is provided. The collected user privacy information which is not marked in a significant way &amp; without the user's consent is used for targeted push or targeted advertising marketing.</td>
<td>Manual work</td>
</tr>
<tr>
<td></td>
<td>Providing personal data to others without users' consent.</td>
<td>Providing users' personal data to others or third party SDKs without users' consent.</td>
<td>Automation tools &amp; manual work</td>
</tr>
<tr>
<td>Illegal transfer of user data</td>
<td>Transmission of sensitive data in clear text.</td>
<td>Using http or insecure https to transfer private user data.</td>
<td>Automation tools &amp; manual work</td>
</tr>
<tr>
<td></td>
<td>Storage &amp; requesting of resources across borders.</td>
<td>Storing &amp; requesting resources across borders.</td>
<td>Automation tool &amp; manual work</td>
</tr>
<tr>
<td>Illegal storage of user data</td>
<td>Storage of sensitive data in SD card.</td>
<td>Storing sensitive data in SD card leading to the data to be obtained by other third-party application.</td>
<td>Automation tool &amp; manual work</td>
</tr>
<tr>
<td></td>
<td>Deletion of user data without permission.</td>
<td>Deleting ser data in the SD card without permission.</td>
<td>Manual work</td>
</tr>
</tbody>
</table>
## Enforcement level

<table>
<thead>
<tr>
<th>Type</th>
<th>Item</th>
<th>Description</th>
<th>Test method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network request</td>
<td>Uploading user data</td>
<td>Statistics &amp; log service uploads user data such as messages, credit card data etc. to third-party servers. Requesting third party links with cookies.</td>
<td>Manual work</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Video data such as drafts &amp; screenshots are automatically uploaded to the server without authorization.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>HTTP, WS plaintext transmission or unencrypted financial data transferred to server</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Overseas data communication</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Certificate validation concerns</td>
<td>Manual work</td>
</tr>
<tr>
<td></td>
<td>Elevation of Privilege to get users’ sensitive data</td>
<td>Using add, get, set, &amp; change APIs to obtain users’ sensitive data</td>
<td>Manual work</td>
</tr>
<tr>
<td>Privacy policy</td>
<td>Collect personal information without user consent or authorization.</td>
<td>Such as MAC address, SSID, BSSID, Android_ID, openUID, packageList, clipboard, serial number, etc.</td>
<td>Automation tools</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Other items in the logical classification about supervision requirements.</td>
<td>Manual work</td>
</tr>
<tr>
<td>Application permissions</td>
<td>Using permissions before user agreement</td>
<td>Requesting the permission when not using the related functions, such as location, SMS, etc.</td>
<td>Manual work</td>
</tr>
<tr>
<td></td>
<td>Requesting the permission when not using the related functions, such as location, SMS, etc.</td>
<td>If user does not agree to enable other permissions beside the minimum necessary permissions for app, &amp; the App won't work.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Keeping requesting permissions after rejection</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hard code</td>
<td>oversea domains &amp; IP addresses, userinfo, test codes hard code in the application</td>
<td></td>
<td>Automation tools &amp; manual work</td>
</tr>
<tr>
<td>App log</td>
<td>App log prints users’ private information</td>
<td></td>
<td>Manual work</td>
</tr>
<tr>
<td>Excessive data request</td>
<td>Frequently reading call logs, locations, clipboard</td>
<td></td>
<td>Automation tools</td>
</tr>
</tbody>
</table>
## Enforcement level: Example

<table>
<thead>
<tr>
<th>Network request</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uploading user data</td>
<td>Statistics &amp; log service to upload user data such as messages, credit card data to third-party servers. Requesting third party link with cookie.</td>
</tr>
<tr>
<td></td>
<td>Video data such as drafts &amp; screenshots automatically uploaded server without authorization.</td>
</tr>
<tr>
<td>HTTP, WS plaintext transmission or unencrypted financial data transferred to server</td>
<td></td>
</tr>
<tr>
<td>Elevation of privilege</td>
<td>Using add, get, set, &amp; change APIs to obtain users’ sensitive data</td>
</tr>
</tbody>
</table>
Privacy principles in different countries

US & other countries

China
## Automated tools used in the framework

<table>
<thead>
<tr>
<th>Name</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain test tool</td>
<td>Obtain all IP addresses &amp; domains from application to check if they belong to trustworthy entities.</td>
</tr>
<tr>
<td>Privacy pop-up window test tool</td>
<td>Test if the application obtains user’s privacy information before user agrees to privacy policy using hook.</td>
</tr>
<tr>
<td>Application scan tool</td>
<td>Check all the APIs related to privacy &amp; validate if the application has certificate validation concerns</td>
</tr>
<tr>
<td>Traffic detection tool</td>
<td>Obtain all network traffic of the application when running.</td>
</tr>
</tbody>
</table>
## Common private data permissions

<table>
<thead>
<tr>
<th>Dangerous Permission</th>
<th>Normal Permission</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>serial_number</td>
<td>serial_number (No any permission needed)</td>
<td>Build.getSerial()</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ro.serialno</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ro.boot.serialno</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Build.SERIAL</td>
</tr>
<tr>
<td>Imei/Meid</td>
<td>Mac address/Bssid/Ssid</td>
<td>Mac:/sys/class/net/wlan0/address</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mac:getNetworkInterfaces()</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Router Mac:/proc/pid/net/arp</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DEVICE_ID≈IMEI≈MEID≈ESN</td>
</tr>
<tr>
<td>Imsi/Iccid/Phone number</td>
<td>Android_id/Openudid (No any permission needed)</td>
<td>Android_id=Openudid</td>
</tr>
<tr>
<td>Location</td>
<td>Package list (No any permission needed)</td>
<td>List:Pm list</td>
</tr>
<tr>
<td>SMS/Call Log</td>
<td>Clipboard (No any permission needed)</td>
<td>getPrimaryClip()</td>
</tr>
</tbody>
</table>
How can I tell if a static variable is being called?

Static variable can **not** be hooked by Frida or Xposed.

New instrumentation technology:

1. **Add the code to these places after Build.SERIAL.**
2. **Build all the dexes.**
3. **Sign the APK using your signature.**

- **Dex2smali & add log & self signature hook modules to the last dex.**
- **Add the signature code to the application class.**
- **Use the new dexes to replace the dexes in the APK.**
Automated tools in the framework
## Statistical results

<table>
<thead>
<tr>
<th>Description</th>
<th>API</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task</td>
<td>1. getRunningAppProcesses 2. getRunningTasks 3. getRunningServices 4. getRecentTasks</td>
</tr>
<tr>
<td>Wi-fi</td>
<td>1. getSSID 2. toString 3. getScanResults 4. getEx fratInfo 5. getBSSID 6. /net/arp 7. getHardwareAddress 8. /sys/class/net/wlan0/address</td>
</tr>
<tr>
<td>Applist</td>
<td>1. getInstalledPackages 2. queryIntentActivities 3. getInstalledApplications 4. list package</td>
</tr>
<tr>
<td>SystemProperty</td>
<td>getprop</td>
</tr>
<tr>
<td>Clipboard</td>
<td>1. setPrimaryClip 2. getPrimaryClip</td>
</tr>
<tr>
<td>Sensor</td>
<td>1. registerListener 2. requestTriggerSensor 3. registerDynamicSensorCallback</td>
</tr>
<tr>
<td>Location</td>
<td>1. getLatitude 2. getLongitude 3. getCellLocation 4. getAllCellInfo</td>
</tr>
<tr>
<td>IMEI</td>
<td>1. getImei 2. getDeviceId 3. getMeid</td>
</tr>
<tr>
<td>SimcardInfo</td>
<td>1. getSimSerialNumber iccid 2. getSubscriberId 3. getLine1Number</td>
</tr>
<tr>
<td>Calendar</td>
<td>1. content://com.&amp;roid.calendar/events 2. content://com.&amp;roid.calendar/reminders</td>
</tr>
</tbody>
</table>
### Statistical data

<table>
<thead>
<tr>
<th>Name</th>
<th>Times</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target App</td>
<td>1,490</td>
</tr>
<tr>
<td>Effected App</td>
<td>1,229</td>
</tr>
<tr>
<td>Effected App (Except task and SystemProperty)</td>
<td>1041</td>
</tr>
<tr>
<td>Using unnormal API App</td>
<td>579</td>
</tr>
</tbody>
</table>

**Scattergram**

Top 1,490 Apps from Google Play Store & other app stores.
Statistical data

API call statistics

Percentage

- Dangerous Permission: 1970
- Normal Permission: 4544
Summary & Recommendations
Advantages

- Cover as many app privacy audit scenarios as possible.
- Provide the analyst a big picture on items audited.

Recommendations to mobile device users

- Update Android OS to Android 10 or Android 11.
- Update iOS device to iOS14.
- Latest versions of both Android & iOS enforce strict regulations on apps.
- Review & update native privacy settings on iOS & Android
- Review & update app privacy settings.
Thank You!