Introducing Privacy Threats from Ad Libraries to Android Users Through Privacy Granules

Anand Paturi  Patrick Gage Kelley  Subhasish Mazumdar

New Mexico Tech

The University of New Mexico

Anand Paturi
anand@cs.nmt.edu
The Problem

- Resources in the mobile device are grouped as different permissions
- List of permissions is displayed to the user before installation
- User accepts and installs the app

What about third parties?
Third Parties

Ad Networks (Libraries)
- AdMob by Google
- mopub
- Tapjoy
- InMobi
- millennialmedia

Social SDKs
- facebook
- WeChat
- heyZap

Development Tools
- Google Analytics
- ThreatMetrix
- Flurry
Data Accessibility to Third Parties

- Custom SDK implementation
  - Mandatory permissions
  - Optional permissions
  - Developer discretion

What data is accessible to third parties?

Android Smartphone

Location
Contacts
SMS

APK (App code + Third party SDK)
Determining Data Access to Third Parties

Static Analysis
- APK
- Manifest.xml
- Permissions list (data accessible to app provider)
- Third party details
- Search queries

Dynamic Analysis
- TaintDroid-based Emulator
- Network traffic analysis
- Data accessible to third parties
Static Analysis

- Permissions list (data accessible to app provider)
- Third party details
- Search queries
Dynamic Analysis

TaintDroid-based Emulator

Network traffic analysis

Data accessible to third parties
How to Educate Users?

1. Determine critical information
2. Identify threats posed to critical information
3. Design a usable solution
4. Evaluate and enhance the usable solution
Determine Critical Information

*Privacy granules*

- Units of privacy that might be critical to user privacy
  - Location ($L$): Geographical location
  - Identity ($I$): Personal identifying attributes. *E.g.* Social accounts
  - Query ($Q$): Search query preferences. *E.g.* Nearest restaurants

$L + I + Q$!
# Identifying Threats Posed to Privacy Granules

Top 50 free apps as of Feb. 2014

<table>
<thead>
<tr>
<th>Privacy Granule</th>
<th>Threat from App Provider</th>
<th>Threat from Third Parties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location</td>
<td>Static and Dynamic Analysis</td>
<td>Dynamic Analysis</td>
</tr>
<tr>
<td>Identity</td>
<td>Static and Dynamic Analysis</td>
<td>Dynamic Analysis</td>
</tr>
<tr>
<td>Query</td>
<td>Static Analysis</td>
<td>Dynamic Analysis</td>
</tr>
</tbody>
</table>
Design a Usable Solution

- Convey comprehensive threat
- Avoid text jargon
- Threat representation
- Decisive

Icon-based threat interface
Designing Icons for Privacy Granules

<table>
<thead>
<tr>
<th></th>
<th>Icon 1</th>
<th>Icon 2</th>
<th>Icon 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location</td>
<td><img src="image" alt="Location Icon" /></td>
<td><img src="image" alt="Location Icon" /></td>
<td><img src="image" alt="Location Icon" /></td>
</tr>
<tr>
<td>Identity</td>
<td><img src="image" alt="Identity Icon" /></td>
<td><img src="image" alt="Identity Icon" /></td>
<td><img src="image" alt="Identity Icon" /></td>
</tr>
<tr>
<td>Query</td>
<td><img src="image" alt="Query Icon" /></td>
<td><img src="image" alt="Query Icon" /></td>
<td><img src="image" alt="Query Icon" /></td>
</tr>
</tbody>
</table>
User Evaluation for Icons

- Amazon Mechanical Turk (AMT) - Human Intelligence Task (HIT)
- Questionnaire
  - "How likely are you to select image 1 to represent a person’s identity?"
  - Likert scale responses: Likely (5) - Highly Unlikely (1)
  - 137 valid responses

<table>
<thead>
<tr>
<th></th>
<th>Icon 1</th>
<th>Icon 2</th>
<th>Icon 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location</td>
<td><img src="image" alt="Location Icon" /></td>
<td><img src="image" alt="Location Icon" /></td>
<td><img src="image" alt="Location Icon" /></td>
</tr>
<tr>
<td></td>
<td>76%</td>
<td>43%</td>
<td>30%</td>
</tr>
<tr>
<td>Identity</td>
<td><img src="image" alt="Identity Icon" /></td>
<td><img src="image" alt="Identity Icon" /></td>
<td><img src="image" alt="Identity Icon" /></td>
</tr>
<tr>
<td></td>
<td>56%</td>
<td>29%</td>
<td>33%</td>
</tr>
<tr>
<td>Query</td>
<td><img src="image" alt="Query Icon" /></td>
<td><img src="image" alt="Query Icon" /></td>
<td><img src="image" alt="Query Icon" /></td>
</tr>
<tr>
<td></td>
<td>41%</td>
<td>62%</td>
<td>26%</td>
</tr>
</tbody>
</table>
Adding Threat Representation

Threat to privacy granule

No threat to privacy granule (safe)

Potential threat to privacy granule
Final Interface

Angry Birds needs access to:

- In-app purchases
- Identity
- Photos/Media/Files
- Wi-Fi connection information
- Device ID & call information

APP PROVIDER

- Your location is revealed
- Your search queries are safe
- Your identity is revealed

THIRD PARTIES

- Your location is revealed
- Your search queries are safe
- Your identity might be revealed

Accept
Reject
## User Evaluation for Icon-based Interface

**Selected 4 popular and privacy invasive apps**

<table>
<thead>
<tr>
<th>App Provider</th>
<th>Third Parties</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Angry Birds" /></td>
<td>I, L</td>
</tr>
<tr>
<td><img src="image" alt="Cut The Rope" /></td>
<td>I, L</td>
</tr>
<tr>
<td><img src="image" alt="Deer" /></td>
<td>I</td>
</tr>
<tr>
<td><img src="image" alt="Walmart" /></td>
<td>I, L, Q</td>
</tr>
</tbody>
</table>
User Evaluation for Icon-based Interface

- AMT- HITs
- 4 Questionnaires - 1 for each app
  - Privacy granule explanation
  - Current permissions list evaluation
  - Icon-based interface evaluation
    - Ease of finding information
    - Impact during app installation
- 272 valid responses
Ease of Finding Information

Threats from App Provider

"How easy it is to find if your location information is being accessed by reading the permissions list?" - Very Difficult (1) - Very Easy (5)
Ease of Finding Information

**Threats from Third Parties**

"How easy it is to find if your identity information is shared with third parties?" - Very Difficult (1) - Very Easy (5)

<table>
<thead>
<tr>
<th>*p&lt;0.05</th>
<th>Location</th>
<th>Identity</th>
<th>Query</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Z-score</td>
<td>p</td>
<td>Z-score</td>
</tr>
<tr>
<td>Angry Birds</td>
<td>2.11</td>
<td>*</td>
<td>2.01</td>
</tr>
<tr>
<td>Fruit Ninja Free</td>
<td>2.04</td>
<td>*</td>
<td>2.07</td>
</tr>
<tr>
<td>Deer Hunter</td>
<td>2.04</td>
<td>*</td>
<td>2.29</td>
</tr>
<tr>
<td>Yelp</td>
<td>1.96</td>
<td>*</td>
<td>1.98</td>
</tr>
</tbody>
</table>

Analysis of Likert Responses Received For Easiness of Finding Threats From Third Parties
Effectiveness of Including Threats from Third Parties

*How much do you agree with following statements:*

S1. This interface would help you to understand privacy risks posed to your location, identity and search query from both the app and other third parties.

S2. This interface would help you to decide whether or not to install the app because third parties accessing data is important to know.
Effectiveness of Including Threats from Third Parties

<table>
<thead>
<tr>
<th></th>
<th>S1</th>
<th>S2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\mu$ &amp; $\sigma$</td>
<td>Z-score</td>
</tr>
<tr>
<td>Angry Birds</td>
<td>4.63(0.54)</td>
<td>2.08</td>
</tr>
<tr>
<td>Fruit Ninja Free</td>
<td>4.52(0.50)</td>
<td>2.02</td>
</tr>
<tr>
<td>Deer Hunter</td>
<td>4.55(0.50)</td>
<td>2.09</td>
</tr>
<tr>
<td>Yelp</td>
<td>4.49(0.50)</td>
<td>1.97</td>
</tr>
</tbody>
</table>

Analysis of User Responses For Questions On Effectiveness Of The Proposed Interface
Self Reported Installation Decision

"Have you used this app before?" Yes/No

"How likely are you to install this app after seeing the above interface?"
Not likely at all (1) - Very likely (5)

How likely are you to install this app after seeing the above interface?

<table>
<thead>
<tr>
<th>App name</th>
<th>Before viewing icon interface</th>
<th>After viewing icon interface</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Angry Birds</td>
<td>38</td>
<td>30</td>
</tr>
<tr>
<td>Fruit Ninja Free</td>
<td>24</td>
<td>45</td>
</tr>
<tr>
<td>Yelp</td>
<td>33</td>
<td>42</td>
</tr>
<tr>
<td>Deer Hunter</td>
<td>4</td>
<td>56</td>
</tr>
</tbody>
</table>

Users' self reported app installation decisions
How Do Users Perceive Privacy Granule?

"In the proposed interface, identity has information related to your email address, phone number, IMEI number, phone call details, contacts details, SMS, calendar details, microphone, camera and photos. Do you think that this grouping should be split?"

- "I think this group is good", - 62
- "I think this group is good but a couple of those things should get their own icons", - 150
- "I need all this information split up into different groups” - 38
- "I want to see every permission separately”. - 22
How Do Users Perceive Privacy Granule?

"Which information that might reveal your identity would you like to see in a separate group represented by an icon for itself?"

- Camera
- Contacts
- Calendar
- Photos
- SMS
User Comments

• Majority appreciated the novel interface 😊
• Presenting threat from third parties before installation was effective
• More granular approach and usage of data

"knowing what information is being shared with third parties will affect my judgement on weather or not to install an app."

"it's useful to see, but still unknown *why* they want access/permission at all."
Limitations

- No device-based testing
- Real world decisions might be different
- Limited data set

Future Work

- Inference levels of third parties using data across different apps
- Category wise analysis
- Recommendations
Questions

anand@cs.nmt.edu