A Large Scale Study of User Behavior, Expectations and Engagement with Android Permissions

Weicheng Cao (U. Toronto)    Chunqiu Xia (U. Toronto)    Sai Teja Peddinti (Google)

David Lie (U. Toronto)    Nina Taft (Google)    Lisa M. Austin (U. Toronto)
Controlling private data sharing with Android Permissions

Users choose what **private data** to share with app via **Android permission system**
Many factors affect user’s decision to deny a permission

Behaviors

Demographic

Expectations

Attitudes

Explanations

Goal: Study the *interplay* of all these factors; study the *effect of one factor while controlling for others*

Challenge: collect these disparate types of data from the same individuals

Challenge: collect data from large, international set of participants
PrivaDroid as experiment tool

Become a participant by installing Android App
Stay for 30 Days
$10 Reward

Online mobile Ad
Platforms
Google Play Store

PrivaDroid App

Terms
This experiment is entirely voluntary and you can join/exit the experiment at any given time. Participants must be over 18 years old. The purpose of this experiment is to study the reasons for the decisions users make when answering runtime permission requests and when installing and uninstalling applications. Survey questions will appear at some app install and uninstall events, and sometimes at the moment of a permission request. We collect demographic information once at the beginning and ask for your opinion on a few items in an exit survey. The app collects no personally identifiable data except for your Google Advertising ID and we will not use this ID for any purpose. Your data:

- [ ] Agree to terms above

CONTINUE
What do we collect

**Demographics:**
Gender, age, education, country/region

**Behavior:**
Grant/Deny decisions
Apps installed

**Rationales:**
Why participants granted or denied a permission

**Expectations:**
Whether participants expected the permission request

**Explanations:**
Apps’ explanations in pre-prompts, for permissions

**Attitudes:**
Privacy sensitivity scores
Permission data summary

Study ran from Nov 2019 to May 2020
10 countries and regions, 1,719 participants
~36K permission decision events (30% surveyed) and overall 16.7% deny rate
Explanations

Explaination must have:
- A keyword about data collection, e.g. access, collect, etc.
- A keyword about a permission/resource type, e.g. camera, photos, etc.

Deny rate 15.4% without explanation -> 7.1% with explanation

Mixed effects logistic regression (MELR) shows presence of explanation reduces deny rate
Expectations

Unexpected requests deny rate: 26.9%
Expected requests deny rate: 12.2%

MELR model shows unexpected runtime requests significantly increase likelihood that a user denies a permission. Model shows this is true even when controlling for other factors.
**Expectations**

Unexpected requests deny rate: 26.9%
Expected requests deny rate: 12.2%

MELR model shows unexpected runtime requests significantly increase likelihood that a user denies a permission. Model shows this is true even when controlling for other factors.
Cross country analysis

- **Challenging to understanding country to country comparison**
  - Privacy attitudes, cultural values, regulatory frameworks, etc.
  - Only observations about the participants in our study

- **Deny rates and distribution**
  - 2 distinct cliques of countries found via pairwise ANOVA tests on the deny rate distributions
  - Participants from countries in the same clique are drawn from populations with the same mean deny rates

HK is excluded because of not enough female participants
Factors influencing deny rate

- Mixed effects logistic regression model with 12 features
  - Privacy sensitivity (4)
  - Explanation (1)
  - Runtime expectation (1)
  - Whether permission decision is in Settings menu or runtime (1)
  - Demographic variables (4)
  - Permission type (1)
- Participant and app are included as random effects
<table>
<thead>
<tr>
<th>Variable</th>
<th>Values</th>
<th>β Coefficient (p-value)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>control awareness</td>
<td>[-2, 2]</td>
<td>-0.044 (0.109)</td>
<td></td>
</tr>
<tr>
<td>collection</td>
<td>[-2, 2]</td>
<td>0.404 (***)</td>
<td></td>
</tr>
<tr>
<td>secondary_use</td>
<td>[-2, 2]</td>
<td>-0.264 (*)</td>
<td></td>
</tr>
<tr>
<td>has_explanation settings_menu</td>
<td>Binary</td>
<td>-0.725 (***)</td>
<td></td>
</tr>
<tr>
<td>country/region (reference: US)</td>
<td>Canada</td>
<td>0.870 (***)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Argentina</td>
<td>0.555 (***)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>UK</td>
<td>0.567 (***)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>France</td>
<td>0.795 (***)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Spain</td>
<td>0.883 (***)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>South Africa</td>
<td>0.068</td>
<td></td>
</tr>
<tr>
<td></td>
<td>India</td>
<td>0.118</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Singapore</td>
<td>0.42 (.s)</td>
<td></td>
</tr>
<tr>
<td>gender (reference: Male)</td>
<td>Female</td>
<td>0.299 (***)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Values</th>
<th>β Coefficient (p-value)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>age (reference: Below 30 years)</td>
<td>Between 30 and 50</td>
<td>-0.104</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Above 50</td>
<td>-0.006</td>
<td></td>
</tr>
<tr>
<td>education (reference: Bachelor’s degree)</td>
<td>Less than high school</td>
<td>-0.249 (*)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>High school or equivalent</td>
<td>-0.193</td>
<td></td>
</tr>
<tr>
<td>permission (reference: Location)</td>
<td>Calendar</td>
<td>0.259</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Camera</td>
<td>0.011</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Contacts</td>
<td>0.258 (**)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Microphone</td>
<td>0.606 (***)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Phone</td>
<td>-0.093</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SMS</td>
<td>-0.265</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Storage</td>
<td>-0.379 (***)</td>
<td></td>
</tr>
<tr>
<td>runtime_expected (reference: Yes)</td>
<td>No</td>
<td>1.216 (***)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Not surveyed</td>
<td>0.306 (***)</td>
<td></td>
</tr>
</tbody>
</table>

Random Effect

<table>
<thead>
<tr>
<th>Variable</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>App (intercept)</td>
<td>1.889</td>
</tr>
<tr>
<td>User (intercept)</td>
<td>1.785</td>
</tr>
</tbody>
</table>

Significance codes:

- p < 0.001 (***)
- p < 0.01 (**)
- p < 0.05 (*)
- p < 0.1 (.).
Limitations

- **Selection Bias:** Participants more likely to
  - Respond to mobile advertising
  - Be tolerant to data collection by a mobile app
  - Be incentivised by financial rewards

- **Incomplete visibility:**
  - Can’t see events for apps before study period, such as pre-installed or popular apps
  - Not enough data to analyze behaviors of individual apps
Conclusions

- Mobile advertising effective in recruiting participants
- Including rationales for permissions benefits the apps by reducing deny rate by more than half (7.1% vs 15.4%)
- Both install-time and runtime expectations affect users permission decisions
  - this is true regardless of demographics and permission type
- Participant demographics, their privacy attitudes, expectations, explanations and permission types all play a role in permission denial decision
Thank you!