TurtleGuard: Helping Android Users Apply Contextual Privacy Preferences

Lynn Tsai  
Primal Wijesekera  
Joel Reardon  
Irwin Reyes  
Jung-Wei (Jennifer) Chen  
Nathan Good  
**Serge Egelman**  
David Wagner  

UC Berkeley  
University of British Columbia  
UC Berkeley  
ICSI  
Good Research  
Good Research  
ICSI / UC Berkeley  
UC Berkeley
HOW TO SHOW PRIVACY NOTICES ON MOBILE DEVICES?
comprehension study

online study of 308 Android users
- quantitative survey

laboratory study with 24 users
- think-aloud experiment
- semi-structured interviews

suggestions

- many were habituated—**too many requests**
- only show information when necessary

- many were unaware—**too late in the process**
- provide information prior to decision-making
this ignores context
show privacy notices more than once, but

**UNDER WHAT CIRCUMSTANCES?**
how often are resources accessed \textit{in practice}? \\

dynamic analysis \\
\quad – modified Android OS and gave phones to 40 people \\
\quad – hooked all API methods involving access to sensitive data \\

what are users doing when data is accessed (\textit{context})?

the results

36 Android smartphone users
6,048 hours of real-world use
27 million permission requests
more runtime requests?

213 requests per hour!
  – location (10,960/day)
  – reading SMS data (611/day)
  – sending SMS (8/day)
  – reading browser history (19/day)

asking each time is infeasible
  – …but 80% of participants wanted to block at least 1 request
  – on average, they wanted to block 35% of all requests
lessons learned

*visibility* of the application requesting permission is a strong contextual cue.

*frequency* at which requests occur makes it impractical to prompt user on every case.

*ask-on-first-use* can be extended to capture the context.
privacy as contextual integrity

inappropriate data flows violate contextual information norms

contextual information norms are modeled using:

• data subject (i.e., the user)
• data sender
• data recipient
• information type
• transmission principle (constraints)
what does this mean for user-centered design?

*notice should be shown in context*

only provide notice when *reasonable* privacy expectations are likely to be violated

- don’t bother the user about appropriate data flows …to prevent habituation
- do bother the user about unknown or inappropriate flows …to preserve contextual integrity
HOW CAN NOTICES BETTER ACCOUNT FOR CONTEXT?
notice in context

definitely determining context is impossible

…use proxies to approximate, and learn from mistakes

can we use machine learning to detect when context has changed from *expected* data use to *unexpected*?
automatically regulating access

<table>
<thead>
<tr>
<th>Model</th>
<th>Error Rate</th>
<th>Average Prompts/User</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ask-on-first-use (Android/iOS)</td>
<td>15.4%</td>
<td>12.34</td>
</tr>
<tr>
<td>ML Model</td>
<td>3.2%</td>
<td>12.00</td>
</tr>
<tr>
<td>ML Model (low-prompt)</td>
<td>7.4%</td>
<td>8.00</td>
</tr>
</tbody>
</table>
because the classifier isn’t perfect,

WHERE CAN USERS GO TO CORRECT ERRORS?
App Info

ZvW
version 4.4.4004.404664

DISABLE
FORCE STOP

Storage
52.1 MB used in Internal storage

Data usage
3.24 MG used since Dec 14, 2016

Permissions
Location

Notifications

Open by default
No defaults set

Battery
20% used since last full charge

Memory
39 MB avg memory used in last 3 hours
status quo problems

1) no holistic view of what apps have accessed

2) per-app information is many layers down

3) permissions are not contextual
privacy dashboard

classifier reduces error rates five-fold

- *some* errors will always occur
- open question: *acceptable* error rate?

users need to be able to:

- *understand* automated decisions (auditing/error detection)
- *change* incorrect decisions (retraining)
prior work showed <25% of users understand background applications have same abilities as foreground ones

does the dashboard correct this mental model?
does it help users regulate access based on context?

<table>
<thead>
<tr>
<th>Activity</th>
<th>Apps</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Browser</strong></td>
<td></td>
</tr>
<tr>
<td>Denied Location</td>
<td></td>
</tr>
<tr>
<td>Sun Jan 01 05:08:52 PST 2017</td>
<td></td>
</tr>
<tr>
<td><strong>Contacts</strong></td>
<td></td>
</tr>
<tr>
<td>Allowed Read Call Log</td>
<td></td>
</tr>
<tr>
<td>Sun Jan 01 05:08:48 PST 2017</td>
<td></td>
</tr>
<tr>
<td><strong>Browser</strong></td>
<td></td>
</tr>
<tr>
<td>Allowed Location</td>
<td></td>
</tr>
<tr>
<td>Sun Jan 01 05:06:11 PST 2017</td>
<td></td>
</tr>
<tr>
<td><strong>ZvW</strong></td>
<td></td>
</tr>
<tr>
<td>Allowed Location</td>
<td></td>
</tr>
<tr>
<td>Sun Jan 01 05:03:02 PST 2017</td>
<td></td>
</tr>
</tbody>
</table>

- **Location**: Always
- **Read Contacts**: Never
- **Write Contacts**: Never
- **Read Call Log**: Never
TurtleGuard

Permission Manager

<table>
<thead>
<tr>
<th>ACTIVITY</th>
<th>APPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>App Store</td>
<td>4 permissions requested</td>
</tr>
<tr>
<td>Browser</td>
<td>4 permissions requested</td>
</tr>
<tr>
<td>Contacts</td>
<td>4 permissions requested</td>
</tr>
<tr>
<td>Music Player</td>
<td>4 permissions requested</td>
</tr>
<tr>
<td>ZvW</td>
<td>4 permissions requested</td>
</tr>
</tbody>
</table>
TurtleGuard

Permission Manager

<table>
<thead>
<tr>
<th>ACTIVITY</th>
<th>APPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>App Store</td>
<td></td>
</tr>
<tr>
<td>Location</td>
<td>When in use</td>
</tr>
<tr>
<td>Read Contacts</td>
<td>When in use</td>
</tr>
<tr>
<td>Write Contacts</td>
<td>Never</td>
</tr>
<tr>
<td>Read Call Log</td>
<td>Never</td>
</tr>
<tr>
<td>Browser</td>
<td></td>
</tr>
<tr>
<td>Contacts</td>
<td></td>
</tr>
<tr>
<td>Music Player</td>
<td></td>
</tr>
<tr>
<td>Activity</td>
<td>Apps</td>
</tr>
<tr>
<td>----------</td>
<td>------</td>
</tr>
<tr>
<td>Location</td>
<td></td>
</tr>
<tr>
<td>Read Contacts</td>
<td>Never</td>
</tr>
<tr>
<td>Write Contacts</td>
<td>Never</td>
</tr>
<tr>
<td>Read Call Log</td>
<td>Never</td>
</tr>
</tbody>
</table>

**ZvW: Location**

- Always
- When in use
- Never
dashboard user study

functioning prototype implemented in HTML5:

- control (Android M settings panel)
- study conducted online (n=400)

tasks designed to test:

- understanding resource usage
- fixing misaligned settings
task 1

what are the two most recent apps that accessed the device’s location?

• open-ended response
task 1: control
task 1: control
task 1: control
task 1: TurtleGuard
task 1: TurtleGuard
task 1

results:
- control: 84% correct (167/198)
- TurtleGuard: 68% correct (132/194)

observations:
- confusion from presenting all data types together
- a quarter never opened TurtleGuard
task 2

currently, which of the following data types can be accessed by the ZvW app?

• multiple choice
task 2: control
<table>
<thead>
<tr>
<th>Activity</th>
<th>Apps</th>
</tr>
</thead>
<tbody>
<tr>
<td>App Store</td>
<td>4 permissions requested</td>
</tr>
<tr>
<td>Location</td>
<td>When in use</td>
</tr>
<tr>
<td>Read Contacts</td>
<td>When in use</td>
</tr>
<tr>
<td>Write Contacts</td>
<td>Never</td>
</tr>
<tr>
<td>Read Call Log</td>
<td>Never</td>
</tr>
<tr>
<td>Browser</td>
<td>4 permissions requested</td>
</tr>
<tr>
<td>Contacts</td>
<td>4 permissions requested</td>
</tr>
<tr>
<td>Music Player</td>
<td>4 permissions requested</td>
</tr>
</tbody>
</table>
task 2

results:

• control: 70% correct (140/198)
• TurtleGuard: 59% correct (116/194)

observations:

• several confounding factors
tasks 3 & 4

is the ZvW app able to access location data even when it is not actively being used?

• multiple choice

prevent it from doing so…or explain whether it is even possible

• open-ended response
tasks 3 & 4: control

• answering question relies on knowledge of Android!

• not possible to regulate foreground vs. background data access!
tasks 3 & 4: control

• answering question relies on knowledge of Android!

• not possible to regulate foreground vs. background data access!
tasks 3 & 4: TurtleGuard
tasks 3 & 4

results:
• control:
  • 43% understood background access would occur
  • 23% correctly understood they could not regulate it
• TurtleGuard:
  • 78% understood background access would occur
  • 75% were able to limit it

observations:
• strong effects ($\phi_{\text{task 3}} = 0.36$, $\phi_{\text{task 4}} = 0.51$)
ITERATIVE DESIGN
<table>
<thead>
<tr>
<th>ALLOWED</th>
<th>DENIED</th>
<th>APPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>App Store</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Browser</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contacts</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dater</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Downloads</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gamerz</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maps</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Music Player</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZvW</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
VALIDATION STUDY
methodology

repeat of pilot study
  • same 4 tasks

new Mechanical Turk participants (n=298) compared with Prolific Academic sample (n=300)
  • no differences in behavior
validation study results

task 1: understanding recent location access
- control: 83% correct (of 287)
- TurtleGuard: 83% correct (of 293)

task 2: finding granted permissions
- control: 77% correct (of 287)
- TurtleGuard: 81% correct (of 293)
validation study results

task 3: understanding privacy settings (background access):
- control: 38% correct (of 287)
- TurtleGuard: 79% correct (of 293)
- medium-to-large effect size: $\phi=0.41$

task 4: changing privacy settings (background access):
- control: 28% said it was impossible (of 287)
- TurtleGuard: 77% correctly did it (of 293)
- large effect size: $\phi=0.49$
conclusions

demonstrated how to use iterative user-centered design to construct a new interface for managing app permissions:

- without prior training, performs as well as the permissions interfaces users are accustomed to
- performs significantly better at communicating foreground vs. background data access
NEXT STEPS
validation field study

we have a working prototype…

…how will real users use it?

will users make the same decisions when permission are actually denied?

will they rectify errors with the dashboard?
validation **field study**

**methodology**

- recruit 40 local Android users
- use our phones as their own for a week
- measure decision-making and interactions
- post-study interviews/tasks
questions?