Do You Get What You Pay For?

Comparing the Privacy Behaviors of Free vs. Paid Apps

Catherine Han

Irwin Reyes

Amit Elazari Bar On

Joel Reardon

Álvaro Feal

Kenneth A. Bamberger

Serge Egelman

Narseo Vallina-Rodriguez

PRIVACYCON

catherinehan@berkeley.edu

@catherinekshan

Geometry Dash Lite

RobTop Games Arcade

★★★★½ 6,340,121

Everyone

Contains Ads

Add to Wishlist Install

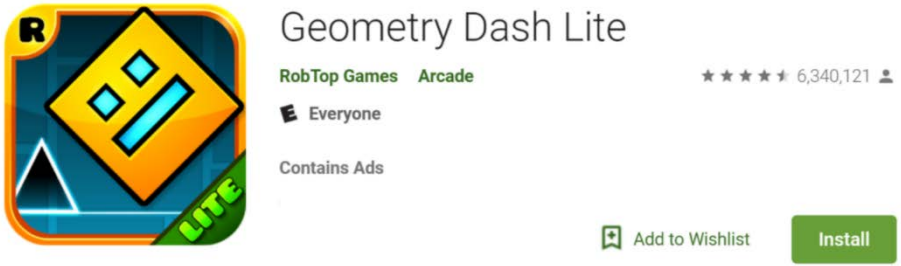
Geometry Dash

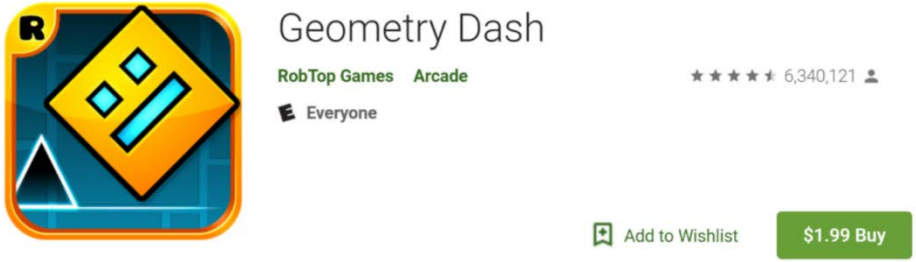
RobTop Games Arcade

★★★★½ 6,340,121

Everyone

Add to Wishlist $1.99 Buy





- In what ways do you expect the given apps to differ?
- Which would you be more likely to install? Why?

Free or Paid… and *Why*?

- Differences: ads (49.4%), security & privacy (1%)
- Preference for paid (20% overall): ad removal (30%), security & privacy (6%)

"I hate spending money on apps… even though there's a higher chance of a virus for [the free app], I would download it to save the money"

"[the paid app] would be less susceptible to security breaches and data mining"

Consumer Expectations

Free

- Have looser user **data sharing** practices
- Have poorer **data retention** practices

Paid

- Follow better **security practices**
- Be **compliant** with laws and regulation
- Have more **transparency** about its privacy behaviors

The New York Times

How to Protect Your Privacy as More Apps Harvest Your Data

By Brian X. Chen

But if it's a for-profit business offering a free product, count on it monetizing your data somehow.

"Follow the money," Mr. Tien said. "If you're not paying for it with money, you're paying for it with data."

Why you should want to pay for apps

by MATTHEW PANZARINO — Apr 23, 2011 in APPS

Sometimes however, especially in the case of apps that are offered without a price, the cost is hidden. Instead of paying for the app up front you may pay for it in one of many other ways over the course of using the app. The developer of a free app has several avenues of making money provided to them by Apple including in-app purchases and of course, advertising.

How do the data collection practices of free apps compare to their paid versions?

We analyze…

permissions declared

third-party packages included in app

domains that receive **sensitive data**

Corpus

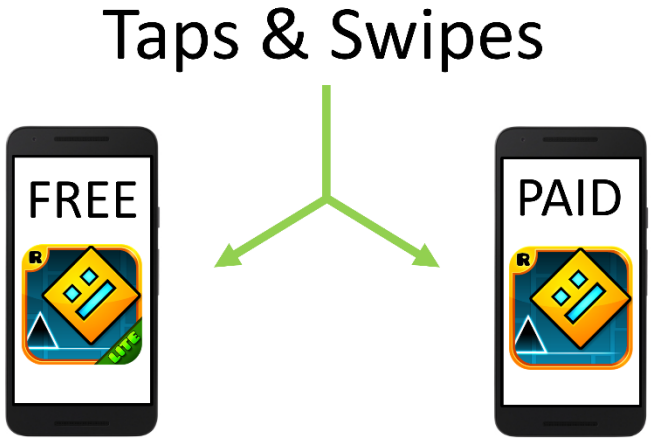
1,505 pairs of apps from the Google Play Store

spans 1,159 unique developers

Side by side runs…

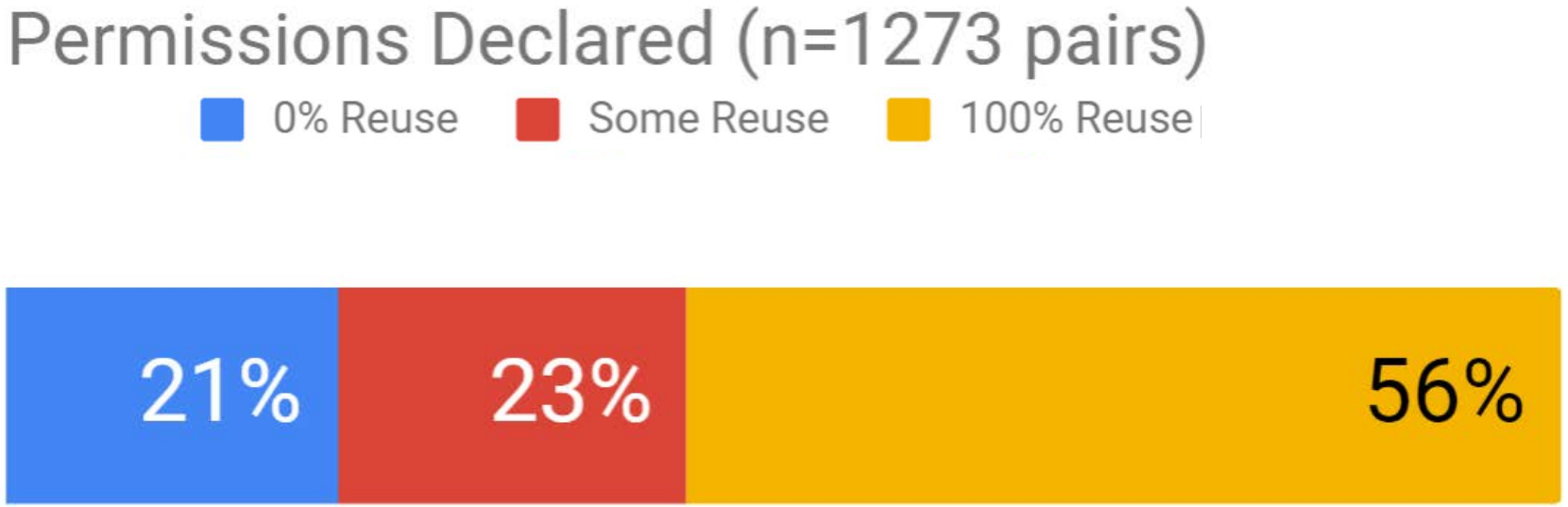
Each free app and its paid counterpart were ***installed simultaneously*** on two identical Nexus 5X phones

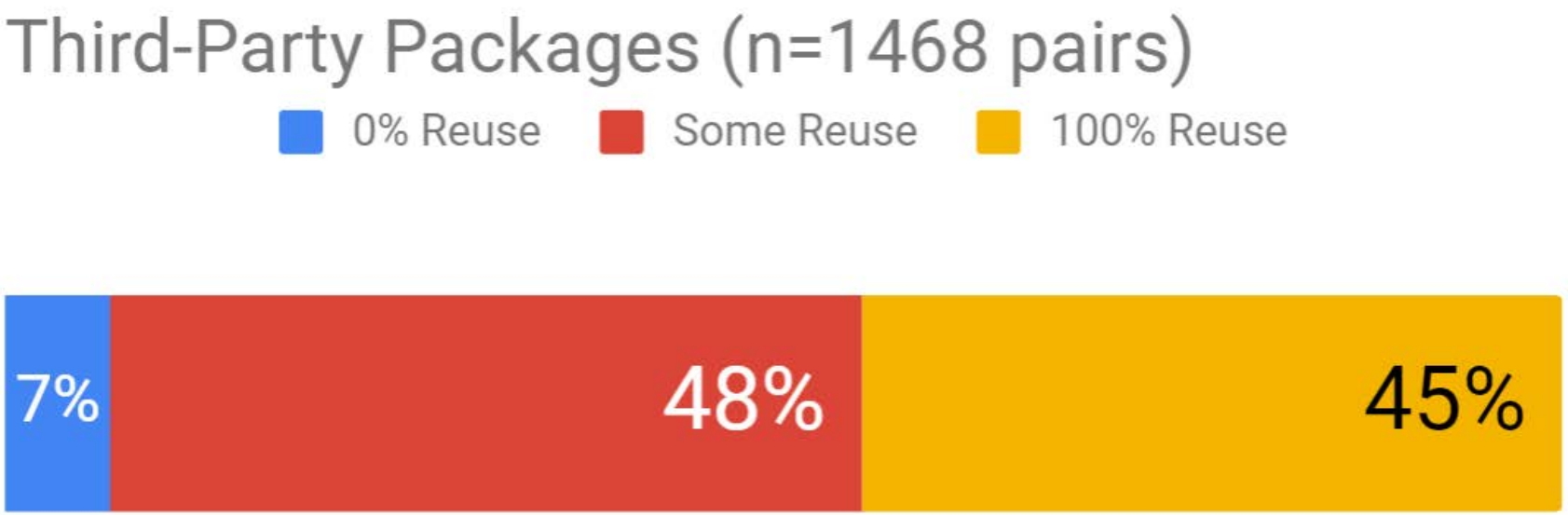
Controlling differences in app behavior by using the ***same random input stream*** at the same time



What We Found

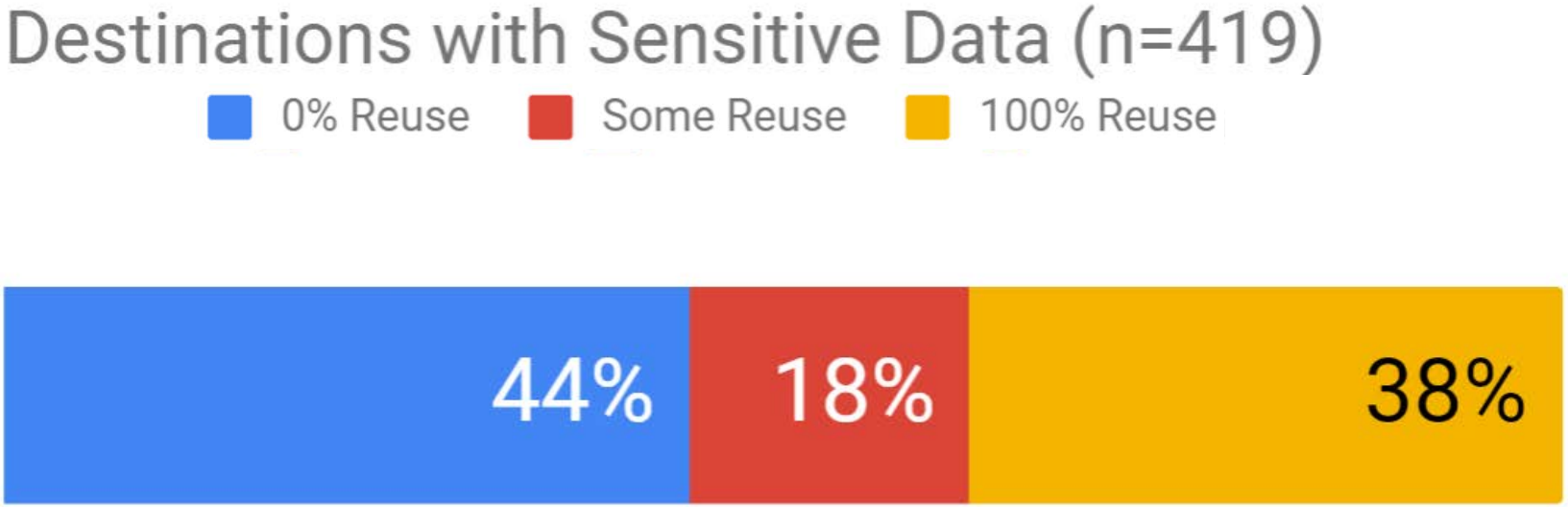
PRIVACYCON


Permissions Declared (n=1273 pairs)
0% Reuse
Some Reuse
100% Reuse
21%
23%
56%


Third-Party Packages (n=1468 pairs)
0% Reuse
Some Reuse
100% Reuse
7%
48%
45%

Third-Party Library Categories

- Categorization of libraries via **LibRadar**
- 831 of the pairs had at least one advertising library across the versions
  - **96.5%** of these had ad libraries in the *free* version
  - **49.1%** of the *paid* versions contained ad libraries


Destinations with Sensitive Data (n=419)
0% Reuse
Some Reuse
100% Reuse
44%
18%
38%

Takeaways

Measurable privacy benefits of paying for an app are **tenuous at best.**

Assumptions surrounding **“pay-for-privacy”** as a model are *misleading*.

PRIVACYCON

catherinehan@berkeley.edu
@catherinekshan

More Information & References

[1] Corpus Pairs: https://github.com/io-reyes/play-store-purchase/blob/master/data/pairs-conpro.csv

[2] LibRadar: http://sei.pku.edu.cn/~maziang14/papers/Ma-ICSE-16.pdf

[3] Dynamic Analysis Tooling: https://arxiv.org/pdf/1510.01419.pdf