The Price is (not) Right: Comparing Privacy in Free and Paid Apps

Catherine Han,[1] Irwin Reyes,[2] Alvaro Feal,[4] Joel Reardon,[3]

Narseo Vallina-Rodriguez,[1,4] Serge Egelman[1,2]

[1]UC Berkeley, [2]ICSI, [3]University of Calgary, [4]IMDEA Networks Institute

PI: Serge Egelman

Contact: catherinehan@berkeley.edu @catherinekshan | **Project URL**: https://appcensus.io

Do consumers actually expect better privacy practices from free versions of applications, as suggested by expert advice?

- Designed and conducted a consumer expectation survey of 1,000 participants on Prolific
- Found that consumers have definite expectations of paid apps having better security and privacy behaviors
- These expectations, however, do not align with the observed ground truth

To what extent do the data collection practices of free apps differ from their paid counterparts?

Geometry Dash Lite
RobTop Games   Arcade
★★★★★ 6,340,121
Everyone
Contains Ads
Add to Wishlist   Install

Geometry Dash
RobTop Games   Arcade
★★★★★ 6,340,121
Everyone
Add to Wishlist $1.99 Buy
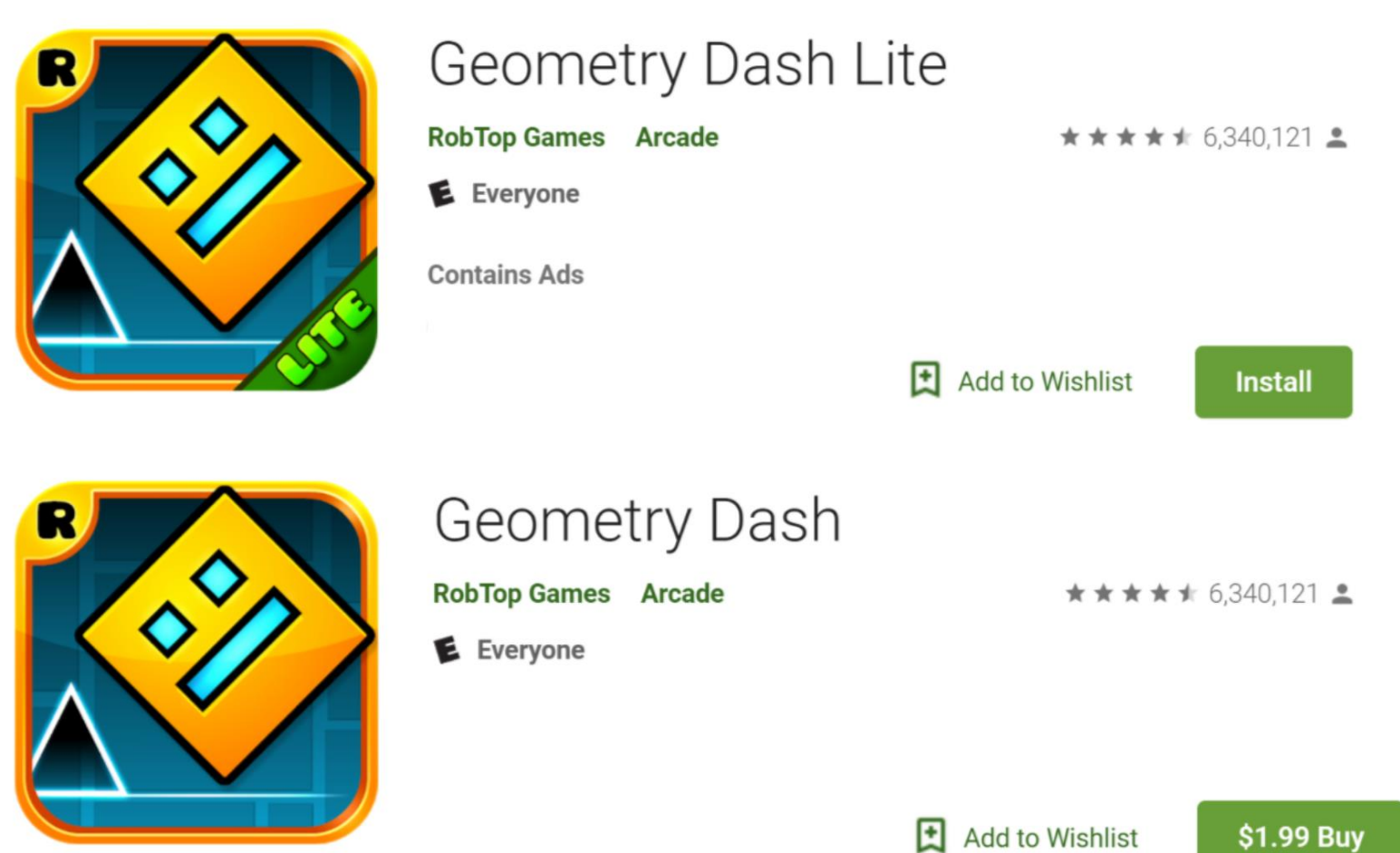
Methodology

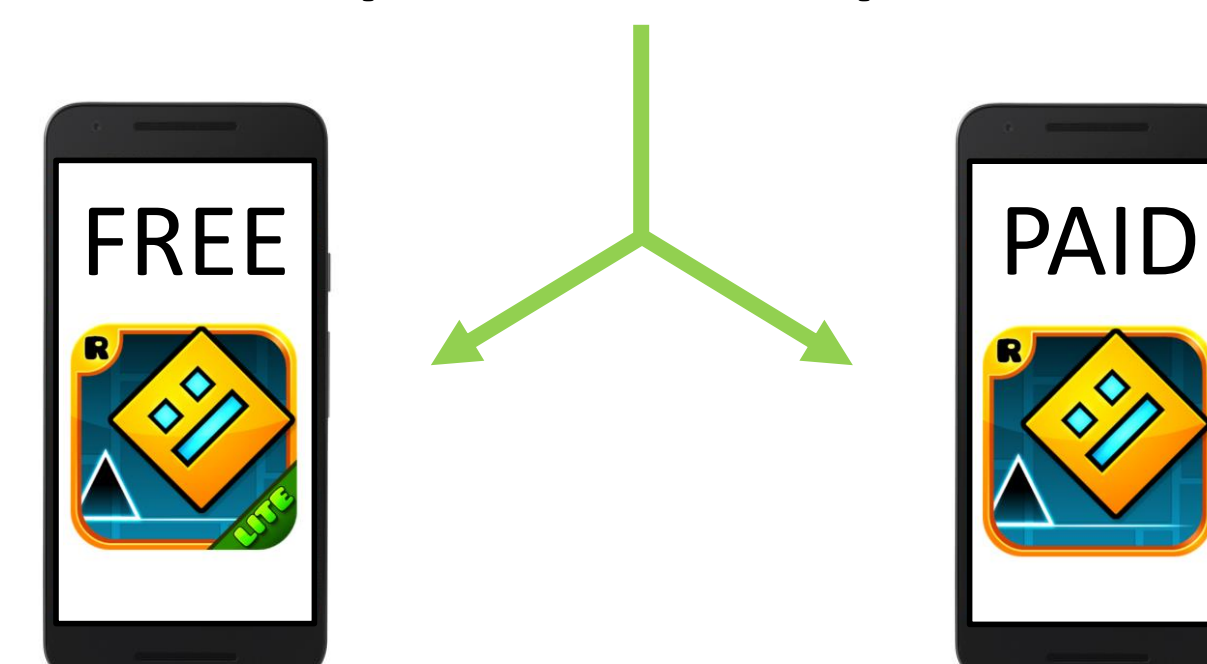After identifying a free app and its paid version, we employed static and dynamic analysis to:

- Measure prospective differences in third-party packages such as advertising and analytics libraries
- Compare the level of access to sensitive resources via the Android permission system
- Identify the different remote domains that receive users' sensitive information like location data and Android Advertising ID, among others
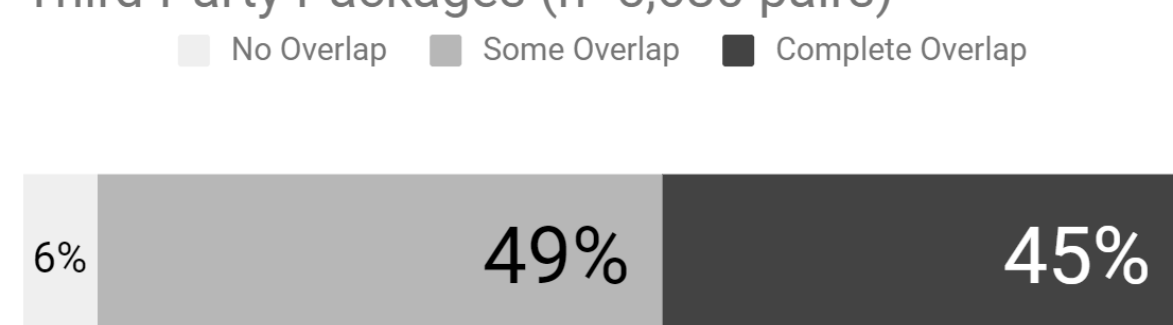
Pipeline

We analyzed 5,877 pairs of free and paid apps from Google Play Store

- We controlled for observed differences
- We used *apktool* and *aapt* to extract bundled third-party packages and declared permissions
- Using a pipeline from a prior work [1], we evaluated runtime differences in network transmissions
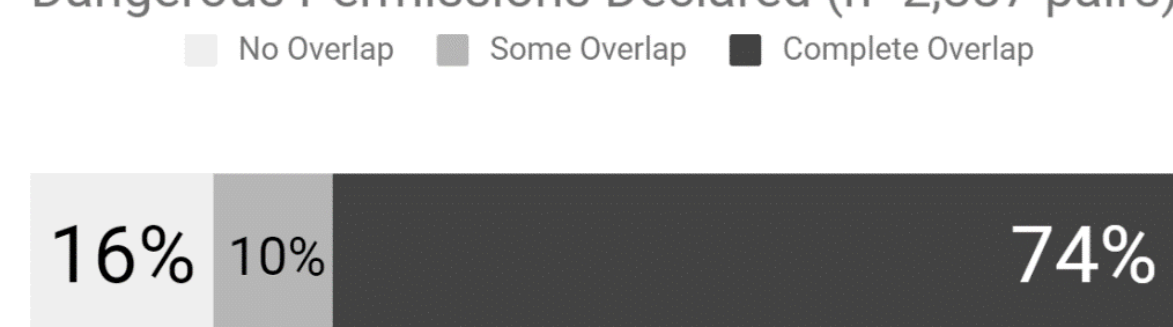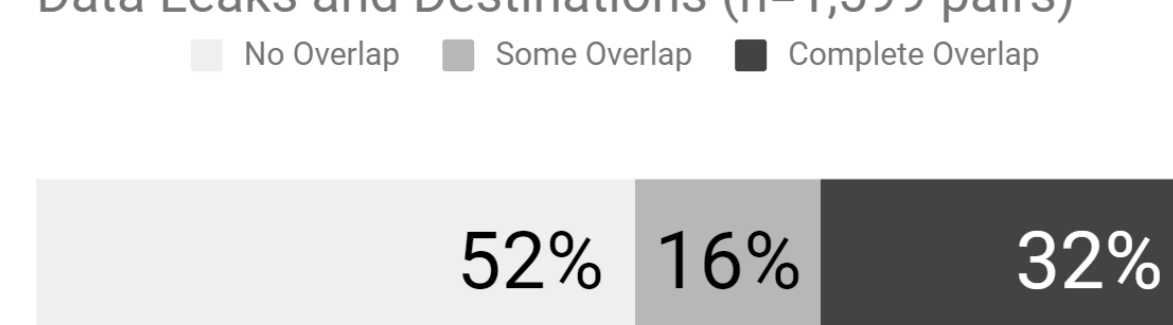
Taps & Swipes

FREE → PAID

Third-Party Packages (n=5,680 pairs)
No Overlap   Some Overlap   Complete Overlap
6% | 49% | 45%

Dangerous Permissions Declared (n=2,887 pairs)
No Overlap   Some Overlap   Complete Overlap
16% | 10% | 74%

Data Leaks and Destinations (n=1,599 pairs)
No Overlap   Some Overlap   Complete Overlap
52% | 16% | 32%

The Ground Truth

Pairs of the same app tend to share the same:

- Dangerous permissions the vast majority of the time
- Third-party libraries about half of the time
- Data transmissions about one third of the time

Contradicts the belief that paying will better protect the consumer

Takeaways

- First large-scale analysis of free and paid apps in direct comparison
- The measurable privacy benefits of paid apps are **tenuous at best**
- The assumptions surrounding the "pay for privacy" monetization model of mobile applications is **misleading**

ICSI   BLUES   BERKELEY LABORATORY FOR USABLE AND EXPERIMENTAL SECURITY   iMdea networks