

An Empirical Study of Third-Party Tracking by Mobile Applications in the Wild

Seungyeop Han
syhan@cs.washington.edu
University of Washington

Jaeyeon Jung
jjung@microsoft.com
Microsoft Research

David Wetherall
djw@cs.washington.edu
University of Washington

As with the web, a rich tracking ecosystem is developing around mobile applications. However, tracking on mobile phones is less well understood and potentially more invasive than tracking on the web. Tracking on mobiles may be done in a variety of ways (as we discovered) by various parties including third parties (e.g., advertising companies), and cannot be controlled by the user beyond the decision of whether to install the application (and grant it the required permissions) in the first place. The privacy threat is heightened on mobiles because they contain a wealth of sensor data and personal information that may be used to profile users. Previous research including our own has shown that some apps do collect and send this information to third parties [1, 2].

We report the first field study of real-world tracking via mobile apps in which we measured how 20 participants were tracked over three weeks as they exercised their Android smartphone apps. We instrumented the phones with dynamic taint tracking to record communications that exposed identifying information, and inspected web cookie databases. We find that 36% of the sites (655 out of 1824) that our study apps are programmed to contact are tracking users. Of these sites, 37% track users with persistent identifiers (mostly AndroidId and IMEI) derived from an identifying string unique to the user’s device. This is privacy risk as

these IDs are long-lived and enable cross-application and cross-site profiling of the user; they are often sent without encryption and with geo-location too.

Advertising and analytics services are widely used, being embedded in 57% of apps and tracking every single participant in our study. Most of these sites are heavy trackers: 25% of their tracking is done with persistent identifiers and geo-location is gathered by half of the top 10 advertisers. Table 1 highlights data on the advertising and analytics tracking domains that were most frequently contacted by apps during the study.

In the poster if accepted, we plan to include a number of interesting analysis results from the measurement data, which we could not fit in the abstract due to space constraint. We will bring an instrumented smartphone to demonstrate how popular Android applications such as Angry birds and NYTimes app allow third parties to track users’ location. The first author of this abstract is a student and he will present the poster at the conference.

References

- [1] W. Enck et al. TaintDroid: An information-flow tracking system for realtime privacy monitoring on smartphones. In *OSDI*, 2010.
- [2] P. Hornyack et al. “These aren’t the droids you’re looking for”: Retrofitting android to protect data from imperious applications. In *CCS*, 2011.

Table 1: Top 5 advertising and top 5 analytics tracking domains

Rank	Advertiser	Apps	Participants	Tracking Identifier	Location	SSL	Cross-App
1	doubleclick.net	65	20	md5(AndroidId)/Cookie		Sometimes	×
2	admob.com	46	19	md5(AndroidId)	Y	Never	×
5	mydas.mobi	9	10	IMEI	Y	Never	×
6	jumptap.com	9	9	IMEI	Y	Never	×
7	atdmt.com	8	9	Cookie		Sometimes	
Rank	Analytics	Apps	Participants	Tracking Identifier	Location	SSL	Cross-App
1	google-analytics.com	49	18	Cookie (separate DB)		Sometimes	
2	flurry.com	28	17	AndroidId	Y	Sometimes	×
3	scorecardresearch.com	13	14	md5(Serial, salt)†		Sometimes	
4	quantserve.com	7	7	Cookie		Sometimes	
5	medialytics.com	6	6	md5(AndroidId)	Y	Never	×

†Uses AndroidId instead of android.os.Build.SERIAL if the system is running a version prior to Android 2.3.