Short: A Measurement Study of Tracking in Paid Mobile Applications

Suranga Seneviratne[†⋆], Harini Kolamunna[†⋆], Aruna Seneviratne[†⋆]
[†]NICTA, Australia; [⋆]School of EET, University of New South Wales, Australia
email : first name.last name@nicta.com.au

ABSTRACT

Smartphone usage is tightly coupled with the use of apps that can be either free or paid. Numerous studies have investigated the tracking libraries associated with free apps. Only a limited number of these have focused on paid apps. As expected, these investigations indicate that tracking is happening to a lesser extent in paid apps, yet there is no conclusive evidence. This paper provides the first large-scale study of paid apps. We analyse top paid apps obtained from four different countries: Australia, Brazil, Germany, and US, and quantify the level of tracking taking place in paid apps in comparison to free apps. Our analysis shows that 60% of the paid apps are connected to trackers that collect personal information compared to 85%–95% in free apps. We further show that approximately 20% of the paid apps are connected to more than three trackers. With tracking being pervasive in both free and paid apps, we then quantify the aggregated privacy leakages associated with individual users. Using the data of user installed apps of over 300 smartphone users, we show that 50% of the users are exposed to more than 25 trackers which can result in significant leakages of privacy.

1. INTRODUCTION

Apps are the driving force behind the use of smartphones. Apps can either be obtained free or bought (paid apps). The adoption of free apps is much greater than paid apps. For example, according to recent reports, the free app percentage is as high as 82% and 92% in Google Play Store and Apple App Store respectively [7, 17]. Free apps are usually monetised by offering advertisements and in-app purchasing capabilities (i.e. virtual goods or additional app features).

Monetisation through advertising requires the collection of user's personal information to tailor future advertisements. The collection of personal information raises concerns about users privacy, and even more so when using mobile devices as they can provide access to a range of richer personal information. Personal information in mobile apps is collected by integrating third party advertising and analytics libraries (*trackers*) with the app. There are a number of studies, that have investigated how trackers associated with free apps collect personal information [12, 28, 8, 6, 26]. In contrast, there is only a limited amount of work which focus on tracking in paid apps [11, 14]. These studies indicate that there is less tracking in paid apps [14]. The observation about lesser tracking happening in paid apps can be corroborated by other analysis such as resource consumption [31, 29] as they show that paid apps consume fewer resources, yet there is no conclusive evidence.

Furthermore, majority of the above studies are based on data collected by crawling Android app markets such as Google Play Store. Such data only provides information about the requested permissions by the app and only free app binaries can be downloaded by crawling. The requested permissions in Android environment are abstract [11] and that information is not sufficient to differentiate the actual permissions requested for the functionality of the app from the permissions requested for the operation of third party trackers. Therefore, in order to get a full understanding on tracking in paid apps, it is necessary to pay and download the app binary files. This paper presents the first insight into tracking in paid apps by purchasing top-100 paid apps from four different countries: Australia, Brazil, Germany, and US, and characterising the tracking happening in paid apps in comparison to free apps.

We make the following contributions in this paper.

- We show that approximately 60% of the paid apps have at least one integrated tracker and around 20% of the paid apps have more than three integrated trackers.
- We also show that trackers popular in free apps are also popular among paid apps and thus expose the users to same level of privacy leakages associated with free apps.
- With only a limited number of trackers being popular in the app eco-system, the users can be exposed to trackers via multiple apps. By analysing list of apps installed by over 300 smartphone users, we show that 50% of the users are exposed to over 25 trackers and 20% of the users are exposed to over 40 trackers.

The remainder of this paper is organised as follows. In Section 2 we discuss the related work and in Section 3 we discuss the datasets used. Methodology is presented Section 4 and we present our findings in Section 5 and Section 6. Section 7 concludes the paper.

WiSec'15, June 22 - 26, 2015, New York, NY, USA

DOI: http://dx.doi.org/10.1145/2766498.2766523.

2. RELATED WORK

A number of work investigated smartphone apps with respect to personal information collection and tracker connectivity [12, 14, 28, 26, 11, 6, 8, 28]. Grace et al. [12] analysed 100,000 free apps collected between March-May, 2011, and identified 100 representative in-app advertisement libraries, embedded in 52.1% of the apps. Various unsafe data collections carried out by advertisement libraries such as user's call logs, account information, and phone number were characterised. Similarly, Leontiadis et al. [14] studied the requested permissions of around 250,000 Android apps (both free and paid) and showed that free apps asked for more dangerous permissions. A similar study was carried out by Viennot et al. [28] and approximately 880,000 free apps were decompiled and integrated ad libraries were presented. Baerera et al. [6] analysed 1,100 free Android apps in order to group their permission patterns. Felt et al. [11] analysed the byte-code of 940 apps that included 100 paid apps.

A number of frameworks for analysing Android applications have also been proposed and implemented [5, 10, 29, 30]. Enck et al. [10] proposed TaintDroid, a framework to track how sensitive smartphone data could leak to the Internet. By monitoring the behaviour of 30 popular Android applications, authors showed that 20 of these applications might misuse users' private information. ProfileDroid proposed by Wei et al. [29], provides a multi-layer system for monitoring and profiling apps. Through evaluation of 27 free and paid Android apps, authors show that there are discrepancies between the app specification and app execution, free versions of apps could end up costing more than their paid counterparts, due to an order of magnitude increase in traffic, and apps communicate with many more sources than users might expect.

These work highlight the over permissions in free apps and associated trackers and provide tools that make such analysis easier. Limited number of work has focused on small number of paid apps. To the best of our knowledge, this is the first systematic study on quantification of the amount of tracking happening in paid apps and how the end users are exposed to aggregated tracking happening in both free and paid apps.

3. DATASETS

3.1 Top-100 Apps

We collected *top-100 free apps* and *top-100 paid apps* from four countries representing four geographical regions. We used Amazon Elastic Computing Cloud [15] to host Squid [2] proxy servers in Sydney (Australia), Sao Paulo (Brazil), Frankfurt (Germany) and North Virginia (United States). Then, we created new Google user accounts for each country via the proxy and associated an Android smartphone to each account. Afterwards, we downloaded top-100 free apps for each country by connecting the smartphone to the Google Play Store via the proxy. We also purchased the top-100 paid apps for each country and downloaded them using the Raccoon automated app download tool [4]. We used the APK files of these apps in the subsequent analysis.

Overall, we had 275 unique free apps (out of 400) and 234 (out of 400) unique paid apps. This is due to the popular apps being common across multiple countries. For example, free apps such as *Facebook, Skype* and *Clash of the Clans* were in top-100 free apps in all four countries and paid apps such as *Minecraft, Tasker, and Poweramp* were in top-100 paid apps across all countries. The number of apps commonly popular between pairs of countries are shown in Table 1.

Table 1: Number of commonly popular apps among countries (Shaded cells represent the paid apps)

	Australia	Brazil	Germany	US
Australia	x	26	32	34
Brazil	33	x	37	32
Germany	39	63	x	32
US	39	43	40	x

3.2 User Installed Apps

We use a dataset containing the lists of *user installed apps* from 338 smartphone users that was used in our previous work [24], under a different context. The users were voluntary researchers or paid workers recruited via Amazon Mechanical Turk [16]. For all the apps found in users' app lists, we downloaded the APK files from Google Play Store again using the Raccoon APK downloader. As it is expensive to purchase all the paid apps users have installed, for paid apps we only considered the paid apps that we purchased described in Section 3.1 and present in users' app lists.

Out of the 5,857 unique apps found among all the users, we were able to obtain APK files of 3,605 apps. The difference was due to numerous reasons, such as user downloading the app from a different app market than Google Play Store, app being no longer available in Google Play Store, and app not being in the set of paid apps that we purchased etc.

4. METHODOLOGY

We decompiled the downloaded APK files using two APK analysis tools to identify the embedded trackers and the API calls executed by the trackers as described below.

4.1 Tracker Identification

Trackers usually provide their SDKs as *jar* files to app developers so that they can be easily embedded into apps. Thus, the decompilation of Android APK files allows the identification of the included third party libraries. Using *apktool*[1] we decompiled the APK files in our dataset to obtain the Java class hierarchy of the app. Then for each app, we manually inspected the class hierarchy and identified the included third party libraries. Using existing literature [27, 12, 28], market reports [18, 13, 23], and searching online for library names, we determined whether or not an integrated third party library is a tracking library. Through this process we were able to identify 124 third party tracker libraries.

We make available this list of trackers together with a brief description about the company and the tracker category (i.e. Advertising, Analytics, Utilities etc.) to the research community [1].

4.2 Personal Information Access

Access to user's personal information in Android is governed by *permissions* and the users need to grant these permissions to the apps at the time of installation. Nonetheless, these permissions are abstract and may not necessarily

[1]https://code.google.com/p/android-apktool/

represent the full implications with respect to leakage of private information associated with granting permissions [11]. A more accurate means of quantifying personal information leakage is to study the API calls executed by the tracker libraries. To this end, we leveraged the capabilities of a commercial malware detection platform, *Joe Sandbox Mobile* [21] that decompiles the source code and provide the code segments which execute Android API calls. We then checked whether these methods are called inside a tracker library or not by comparing the prefixes of the code segments with the previously mentioned library names. For example, when we see that API call *getLastKnownLocation* is called inside the Java class *com.flurry.android.FlurryAgent*, we conclude that it is a call by a tracker as we know that *com.flurry* is a tracker.

5. CHARACTERISATION

In this section we provide a characterisation of the privacy leakages associated with the third party tracking libraries found in paid apps in comparison to free apps.

5.1 Integrated Trackers

Figure 1a and Figure 1b show of the number of trackers integrated to free and paid apps respectively. Approximately 85%–95% of the free apps had at least one tracker integrated in all countries. US had the lowest percentage of apps with zero trackers and that was approximately 4%. Approximately 60% of the paid apps had least one embedded tracker and 20% had more than three trackers. Though the number of trackers in paid apps is lower compared to free apps, it's surprising to see such volume of tracking happening in paid apps, despite the main means of monetisation not being advertising.

Overall, the free app *Swamp Attack*[2] which is an action game, had the highest number of trackers connected (21). From paid apps, the arcade game *Vector (Full Version)*[3] was connected to the highest number of trackers (10).

5.2 Popular Trackers

We then investigated the popular trackers in free and paid apps to understand whether different trackers are popular among these two app categories. For each country, we identified the top-10 trackers by frequency of occurrence for free apps and paid apps and combined those to create two sets of trackers, one that consists of trackers frequently used in free apps and another that is frequently used in paid apps (15 trackers from free apps and 17 from paid apps).

Figure 2a and Figure 2b show the popularity of these trackers as a percentage of availability in top-100 apps. For both free and paid apps the *Google Ads* and *Flurry* were the two most popular trackers and were integrated with over 25% of the apps. There were 10 common trackers between the union of popular trackers (total of 22 trackers) in free apps and paid apps indicating the major players in tracking are equally popular in both free and paid apps.

Overall, out of 124 trackers we identified (cf Section 4.1), 119 trackers were present in free apps and 57 were present in paid apps. Despite top trackers being common in free and paid apps it is possible that tracking objectives in paid apps may not necessarily be similar to free apps. We further investigate this by categorising the trackers according to their functionality and calculating the category popularity in the two types apps as described in the next section.

[2]http://outfit7.com/other/swamp-attack/

[3]http://nekki.com/vector/

Table 2: Tracker categories and examples

Tracker Type	**Description**	**Examples**
Advertising (∼65%)	Libraries mainly serving in-app advertisements and during that process may collect personal information with the objective of providing more personalised advertisements.	Google Ads, Millennial Media, Inmobi, Mopub
Analytics (∼24%)	Libraries providing various analytics such as attributes of the audience of the app and revenue performance of integrated advertisement libraries etc. Usually these libraries act as aggregators who can track users across apps.	Flurry, Google Analytics, Comscore, Amazon Insights, Localytics, Kontagent, Apsalar
Utilities (∼11%)	Libraries assisting developers to track bugs and crashes in apps by providing additional information collected from the smartphones.	Crashlytics, Bugsense

Table 3: Category-wise popularity of trackers

	Advertising		**Analytics**		**Utilities**	
	Free	Paid	Free	Paid	Free	Paid
Australia	59%	54%	27%	30%	14%	16%
Brazil	65%	49%	24%	34%	11%	17%
Germany	59%	48%	28%	33%	13%	20%
US	65%	54%	22%	30%	12%	16%
Overall	64%	58%	25%	28%	11%	14%

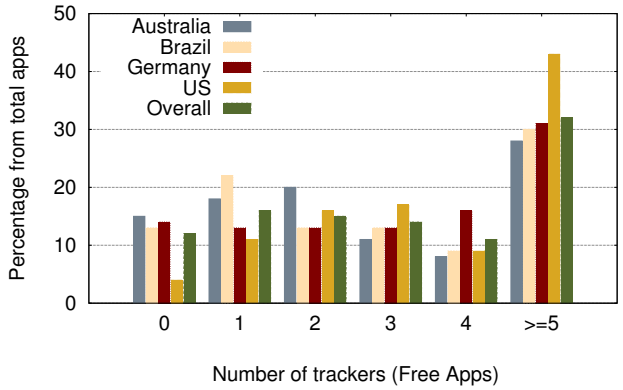
5.3 Tracker Category-wise Popularity

As mentioned in Section 4.1 we categorised the trackers according to existing literature and market reports. Table 2 shows the categories we found and some examples for each category. Table 3 shows the category-wise distribution of the trackers. For both app types, advertising libraries are the most popular followed by analytics and utilities. Overall, 64% of trackers present in free apps were advertising trackers, 25% were analytics trackers, and 11% were utility trackers. Paid apps have a similar composition with 58% of advertising trackers, 28% of analytics trackers, and 14% of utility trackers. This shows that tracking behaviours of paid apps are almost the same as those of free apps and thus all the privacy related issues applicable for free apps are also applicable for paid apps.

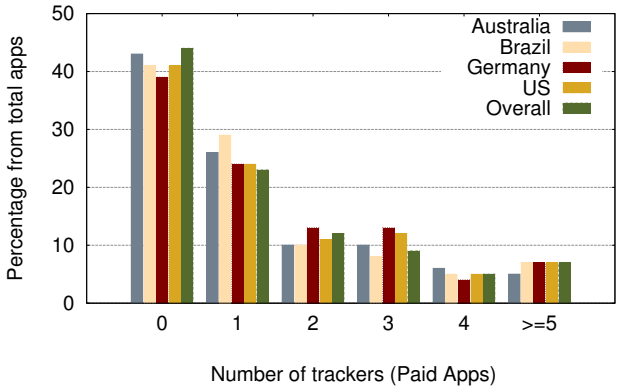
5.4 Accessible Personal Information

In Table 4 we show the personal information accessed by the top 22 tracker libraries identified in Section 5.2. Each row represents one of more related API calls executed by the library. For example, the *Location* row contains the API calls *getLatitude*, *getLongitude*, *getLastKnownLocation*.

Interestingly, most of the libraries did not access crucial personal information such as phone book, SMS content or browser history etc. However, some information accessed by the trackers may result in privacy leakages. For example, eight out of the top 22 trackers collected user location which may be considered as a privacy threat by some users. Four trackers collected either running processes or the list of installed applications. As shown in [24, 25] this information

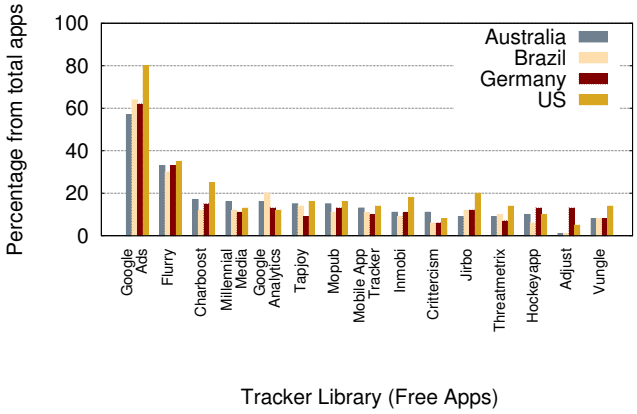


(a) Trackers frequency (Free Apps)

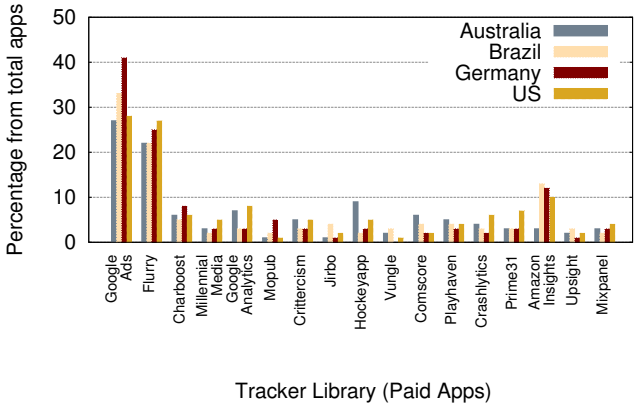


(b) Trackers frequency (Paid Apps)

Figure 1: Tracker Frequency



(a) Top trackers in free apps



(b) Top trackers in paid apps

Figure 2: Tracker Popularity

can be used to easily infer numerous user attributes such as gender, marital status, and religion. 15 out of 22 trackers accessed the connected network information. It was shown by Achara et al. [3] that this information can be exploited to infer users attributes such as travel history, friend information, and location.

Collection of unique identifiers such as Android ID, WiFi MAC address, or device identifiers such as IMEI, allows trackers to identify users across applications enabling behavioural analysis. 17 trackers out of top-22, have access to this type information and it shows the popularity of this method of tracking. This is an interesting observation since in August 2014 Google requested developers to stop using persistent unique identifiers and adapt to the non-persistent advertising identifier for advertising purposes [20].

Collecting Android build information, SIM provider, and network operator information appears to be harmless and related to functionality of some analytic services. For example, for *Crashlytics*, a tracking library that build crash reports for apps, it might be useful to know the Android build versions for debugging purposes. However, this information has the potential to be used for fingerprinting devices, since the persistent unique identifier usage is continuously being discouraged by smartphone OS manufactures due to regulatory concerns [9, 22, 19].

6. DATA AGGREGATION

With tracking being pervasive in both free and paid apps and a limited number of trackers being dominant in the eco-system, individual users can be connected to the same tracker through multiple apps. The use of unique identifiers or device fingerprinting methods, enables these trackers to identify the users across apps and thus can effectively collect more information about those users. Moreover, most of the apps are connected to more than one tracker and therefore the user can be exposed to significant number of trackers despite having a limited number of apps installed in her phone. We checked how users are connected to trackers by analysing the dataset of user installed apps of 338 users that was described in Section 3.2.

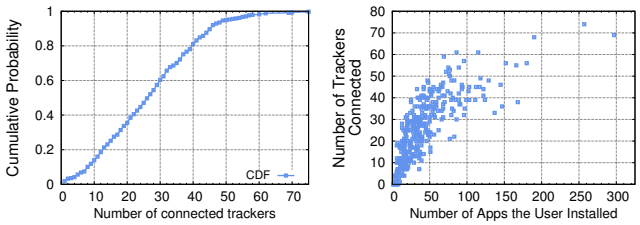
6.1 Trackers Per User

For each app user has installed, we identified the connected trackers using the methodology described in Section 4.1. Then we identified the number of unique trackers the user has connected to. Figure 3a shows the CDF of the number unique trackers connected to each user. As can be seen, *50% of the users are connected to more than 25 trackers and 20% of the users are connected to more than 40 trackers.* Figure 3b shows a scatter plot of number of unique trackers the user has connected to, against the number of apps installed by the user. It can be seen that the

Top trackers in free and paid apps

	Goo. Ads	Flurry	Chartboost	Mille. Media	Goo. Anlytics	Tapjoy	Mopub	Crittercism	M. Apptracker	Inmobi	Jibro	HockeyApp	Threatmetrix	Vungle	Comscore	Playhaven	Crashlytics	Adjust	Prime31	Ama. Insights	Upsight	Mixpanel
User apps																						
Installed apps	-	-	-	-	-	✓	-	-	-	✓	-	-	-	-	-	-	-	-	-	-	-	-
Running apps	-	-	-	-	-	-	-	✓	-	-	-	-	-	-	-	-	-	-	✓	-	-	-
Location	✓	✓	-	✓	-	✓	✓	-	-	✓	-	-	✓	✓	-	-	-	-	-	-	-	-
Calendar Entries	-	-	-	-	-	✓	-	-	-	✓	-	-	-	-	-	-	-	-	-	-	-	-
Unique IDs																						
Android ID	✓	✓	✓	✓	-	✓	✓	-	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	-	-	-	✓
WiFi MAC Address	-	-	✓	-	-	✓	-	-	✓	-	✓	-	-	✓	-	-	✓	✓	-	-	-	-
IMEI	-	✓	-	✓	-	✓	-	-	✓	-	✓	-	✓	-	-	-	✓	-	-	-	-	-
Device Info.																						
OS Build Info.	✓	✓	✓	✓	✓	✓	✓	-	✓	✓	✓	✓	✓	✓	✓	✓	✓	-	✓	✓	-	✓
Phone Info.	✓	✓	✓	-	-	-	-	-	✓	-	-	-	-	✓	-	✓	✓	-	-	✓	-	✓
Network Info.																						
Connectivity Info	✓	✓	✓	✓	✓	✓	✓	-	✓	✓	✓	-	-	✓	✓	-	✓	✓	-	✓	-	-
Connectivity State	-	-	-	-	✓	✓	-	-	-	✓	-	-	-	-	-	-	-	-	-	-	-	-
WiFi Scan	-	-	-	-	-	-	-	-	-	✓	-	-	-	-	-	-	-	-	-	-	-	-
Operator Info.	✓	-	✓	✓	-	✓	✓	✓	✓	✓	✓	-	✓	✓	-	-	-	-	-	✓	-	-
SIM Provider Info.	-	-	✓	-	-	-	✓	-	✓	✓	-	-	✓	-	-	-	-	-	-	-	-	-
Contacts	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	-	-	-
Query emails	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	-	-	-
Reads Logcat	-	-	-	-	-	-	-	✓	-	✓	-	-	-	-	-	-	-	-	-	-	-	-

Table 4: Data accessed by trackers



(a) Trackers per User (b) Trackers vs. Apps

Figure 3: User tracker connectivity

number of connected trackers increased significantly with the number of apps up to approximately 50 apps, and the number of exposed trackers is increased at a lower rate after that.

6.2 Tracker Penetration

Table 5 shows the top trackers, which had a presence among at least 50% of the users. It shows that in addition to Google trackers, over 80% users are connected to other trackers such as Flurry, Millennial Media, Crashlytics, and MoPub.

Trackers can collect more data about the user when they are present in more than one app among the apps user has installed as the rate of receiving data samples increases. In Figure 4 we show the percentage of users giving access to a tracker via more than one app. For example, out of 326 users who had at least one app connected to Google Ads, approximately 78% had more than five apps connected to Google Ads. Corresponding value for Flurry was 55%. This analysis shows top-trackers cover a significant fraction of users across multiple apps and thus receive much richer data about the users.

In Figure 5 we show how an example user from our dataset who is having only 11 apps, is exposed to 26 different trackers and how personal information is flown to the trackers through these apps (For clarity only important personal information are labeled).

Table 5: Trackers who had presence among at least 50% of the users

Tracker	Frequency (Percentage)
Google Ads	326 (96%)
Flurry	307 (91%)
Google Analytics	295 (87%)
Millennial Media	290 (86%)
Crashlytics	287 (85%)
Mopub	274 (81%)
Inmobi	267 (79%)
Hockeyapp	240 (71%)
Comscore	238 (70%)
Crittercism	230 (68%)
Admarvel	201 (59%)
Tapjoy	185 (55%)
Appsflyer	177 (52%)
Chartboost	173 (51%)

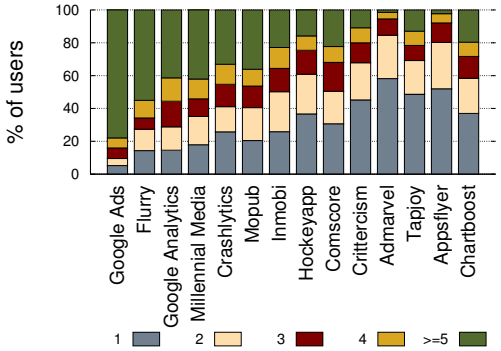


Figure 4: User percentage with more than one app connected to a tracker

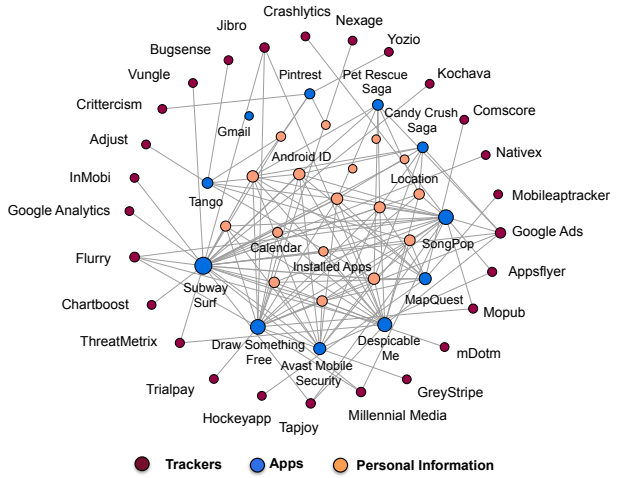


Figure 5: An example user who is only having 11 apps yet connected to 26 trackers

7. CONCLUSION

In this paper we presented a measurement study of tracking in paid apps by analysing top-100 paid apps from four different countries representing four geographical regions. We showed that despite having a different business model, paid apps also collect significant amounts of personal information, and can lead to the same level of privacy leakage as when using free apps. We found that 20% of the paid apps had more than three embedded trackers. Also we showed that 17 out of top-22 trackers collected some form of persistent unique identifiers that allows them to track users across apps.

By analysing apps installed by over 300 smartphone users, we showed that 50% of the smartphone users are connected to more than 25 trackers. The results indicate that it is important to see the overall personal information flow by all the apps installed by a user in addition to evaluating the individual apps' privacy leakages. To overcome this, we are working on an app recommendation system which not only individually evaluates the application's privacy leakages, but also considers how it impacts the overall privacy of the user when the set of applications already installed in the user's smartphone is taken into consideration.

8. ACKNOWLEDGEMENTS

Authors would like to thank Stefan Bühlmann, at Joe Security LLC, Christoph Merian-Ring 11, 4153 Reinach, Switzerland, who kindly provided the access to the *Joe Sandbox Mobile* malware analysis platform, which was used for code level analysis in this paper.

9. REFERENCES

[1] Tracker list. `http://www.privmetrics.org/publications`.

[2] squid-cache.org - Optimising Web Delivery. `http://www.squid-cache.org`, 2015.

[3] J. P. Achara, M. Cunche, V. Roca, and A. Francillon. WifiLeaks: Underestimated Privacy Implications of the ACCESS_WIFI_STATE Android Permission. In *Proc. of the 7th ACM WiSec*, 2014.

[4] P. Ahlbrecht. Raccoon - Google Play desktop client. `http://www.onyxbits.de/raccoon`, 2015.

[5] S. Arzt, S. Rasthofer, C. Fritz, E. Bodden, A. Bartel, J. Klein, Y. Le Traon, D. Octeau, and P. McDaniel. Flowdroid: Precise context, flow, field, object-sensitive and lifecycle-aware taint analysis for android apps. In *Proc. of the 35th ACM SIGPLAN*. ACM, 2014.

[6] D. Barrera, H. G. Kayacik, P. C. van Oorschot, and A. Somayaji. A methodology for empirical analysis of permission-based security models and its application to Android. In *Proc. of the 17th ACM CCS*. ACM, 2010.

[7] C. Bonnington. More iOS apps are free than ever before. http://www.wired.com/2013/07/more-free-ios-apps/, 2013.

[8] P. H. Chia, Y. Yamamoto, and N. Asokan. Is this app safe?: A large scale study on application permissions and risk signals. In *Proc. of the 21st WWW*. ACM, 2012.

[9] D. E. Dilger. Apple adds new "Limit Ad Tracking" feature to iOS 6. `http://appleinsider.com/articles`, 2012.

[10] W. Enck, P. Gilbert, S. Han, V. Tendulkar, B.-G. Chun, L. P. Cox, J. Jung, P. McDaniel, and A. N. Sheth. Taintdroid: an information-flow tracking system for realtime privacy monitoring on smartphones. *ACM Transactions on Computer Systems (TOCS)*, 32(2):5, 2014.

[11] A. P. Felt, E. Chin, S. Hanna, D. Song, and D. Wagner. Android permissions demystified. In *Proc. of the 18th ACM CCS*. ACM, 2011.

[12] M. C. Grace, W. Zhou, X. Jiang, and A.-R. Sadeghi. Unsafe exposure analysis of mobile in-app advertisements. In *Proc. of the 5th ACM WiSec*. ACM, 2012.

[13] A. Gulyani. Extensive list of mobile ad network companies. `http://gulyani.com/complete-list-of-mobile-ad-networks-companies/`, 2014.

[14] I. Leontiadis, C. Efstratiou, M. Picone, and C. Mascolo. Don't kill my ads!: Balancing privacy in an ad-supported mobile application market. In *Proc. of the 12th Workshop on Mobile Computing Systems & Applications*. ACM, 2012.

[15] Amazon Inc. Amazon EC2. `http://aws.amazon.com/ec2/`, 2015.

[16] Amazon Inc. Amazon Mechanical Turk. `https://www.mturk.com/`, 2015.

[17] Appbrain Inc. Distribution of free vs. paid Android apps. http://www.appbrain.com/stats/, 2014.

[18] Appbrain Inc. Android library statistics. `http://www.appbrain.com/stats/libraries`, 2015.

[19] Google Inc. Advertising ID. `https://developer.android.com`, 2014.

[20] Google Inc. Google Play developer program policies. `https://play.google.com/about/developer-content-policy.html`, 2014.

[21] Joe Security LCC. Joe Sandbox Mobile. http://www.joesecurity.org/joe-sandbox-mobile, 2015.

[22] S. Oliver. MAC address randomization joins Apple's heap of iOS 8 privacy improvements. http://appleinsider.com/articles, 2014.

[23] C. Reynolds. A list of mobile advertising networks. `http://www.mobyaffiliates.com/blog/a-list-of-mobile-advertising-networks/`, 2013.

[24] S. Seneviratne, A. Seneviratne, P. Mohapatra, and A. Mahanti. Predicting user traits from a snapshot of apps installed on a smartphone. *ACM SIGMOBILE Mobile Computing and Communications Review*, 18(2):1–8, 2014.

[25] S. Seneviratne, A. Seneviratne, P. Mohapatra, and A. Mahanti. Your installed apps reveal your gender and more! *ACM SIGMOBILE Mobile Computing and Communications Review*, 18(3):55–61, 2015.

[26] S. Shekhar, M. Dietz, and D. S. Wallach. Adsplit: separating smartphone advertising from applications. In *Proc. of the 21st USENIX*, 2012.

[27] N. Vallina-Rodriguez, J. Shah, A. Finamore, Y. Grunenberger, K. Papagiannaki, H. Haddadi, and J. Crowcroft. Breaking for commercials: Characterizing mobile advertising. In *Proc. of the 2012 IMC*. ACM, 2012.

[28] N. Viennot, E. Garcia, and J. Nieh. A measurement study of Google Play. In *Proc. of the SIGMETRICS*. ACM, 2014.

[29] X. Wei, L. Gomez, I. Neamtiu, and M. Faloutsos. Profiledroid: Multi-layer profiling of android applications. In *Proc. of the 18th Mobicom*. ACM, 2012.

[30] L. Weichselbaum, M. Neugschwandtner, M. Lindorfer, Y. Fratantonio, V. van der Veen, and C. Platzer. Andrubis: Android malware under the magnifying glass. *Vienna University of Technology, Tech. Rep. TRISECLAB-0414-001*, 2014.

[31] L. Zhang, D. Gupta, and P. Mohapatra. How expensive are free smartphone apps? *ACM SIGMOBILE Mobile Computing and Communications Review*, 16(3):21–32, 2012.