Differentiating Smartphone Users by App Usage

Pascal Welke
University of Bonn
welke@uni-bonn.de

Ionut Andone
University of Bonn
andone@cs.uni-bonn.de

Konrad Błaszkiewicz
University of Bonn
blaszkie@cs.uni-bonn.de

Alexander Markowetz
markowetz.de
alex@markowetz.de

ABSTRACT

Tracking users across websites and apps is as desirable to the marketing industry as it is unalluring to users. The central challenge lies in identifying users from the perspective of different apps/sites. While there are methods to identify users via technical settings of their phones, these are prone to countermeasures. Yet, in this paper, we show that it is possible to differentiate users via their set of used apps, their *app signature*. To this end, we investigate the app usage of 46 726 participants from the Menthal project. Even limiting our observation to the 500 globally most frequent apps results in unique signatures for 99.67% of users. Furthermore, even under this restriction, the average minimum Hamming distance to the closest other user is 25.93. Avoiding identification would thus require a massive change in the behavior of a user. Indeed, 99.4% of all users have unique usage patterns among the top 60 globally used apps. In contrast to previous work, this paper differentiates between users based on behavior instead of technical parameters. It thus opens an entirely new discussion regarding privacy.

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

Author Keywords

Mobile Devices; Usage Patterns; Privacy

INTRODUCTION

Targeted advertisement and surveillance both aim at tracking as many activities of users as possible. One challenge lies in identifying a user across different services. There exists a large industry that tracks user interactions with various platforms including web pages and applications on smartphones and sells this information to advertisers.

There are various approaches to tracking users on the web. Traditional methods use cookies or invisible images. However, [5] showed in a large empirical study that the information browsers send to websites is rich enough to identify them without the use of cookies. It is thus possible to identify a user by e.g. browser version, browser extensions, installed fonts, timezone, etc. Many users do not want to be tracked and hence try to circumvent some of the fingerprinting techniques by using dedicated tools such as Ghostery[1]. More recently, however, [9] analyzed three commercial browser fingerprinting libraries and showed that they were able to circumvent many countermeasures.

On smartphones, operating systems such as Android provide several means to application developers to uniquely identify phones. The MAC address of Wifi and Bluetooth modules, as well as the IMEI of the GSM module are unique and can be accessed by apps with the right permissions. Additionally, there is a unique "android id" that may only change upon a factory reset of the phone. Several companies offer services and frameworks that can be included in apps to provide tracking capabilities for targeted advertisement. However, little is known on the signals that these companies use to identify and track users. Accessing identifier information (e.g. the IMEI or Android ID) commonly requires explicit permissions and consent from the user during app installation. Again, users who want to avoid tracking on their phones can apply countermeasures. They can e.g. spoof their MAC addresses, block access to identifiers or parts of the phone's hardware.

Sadly, this does not suffice to achieve anonymity. Sensors and information that are freely available to any installed app can serve as signals that differentiate between different users. In a rare publication, Bojinov et al. [4] showed that accelerometer imperfections and distortions in the speaker-microphone system can identify a smartphone. They identified between 12 of 15 and 16 of 17 identical phones correctly playing and recording a sound for three or more seconds. In a larger experiment they identified 8.3% of 3 583 previously seen devices among some 16 000 devices using accelerometer imperfections. Weiss and Lockhart [11] showed on a study with 70 participants that accelerometer data can be used to predict user traits such as sex, height and weight.

This investigation does not focus on technical parameters, but traces actual behavior. We model users as static vectors of all their used apps during our observation period. Hence, to change their fingerprint, users would need to change their interaction patterns with the phone significantly, i.e., by using different apps, not just alter some technical settings. Furthermore, as we trace usage, it would not suffice to just install

[1] www.ghostery.com

UbiComp '16, September 12-16, 2016, Heidelberg, Germany

ISBN 978-1-4503-4461-6/16/09...$15.00

DOI: http://dx.doi.org/10.1145/2971648.2971707

age	number of users
$0 \leq x < 12$	292
$12 \leq x < 17$	10283
$17 \leq x < 21$	10970
$21 \leq x < 25$	7398
$25 \leq x < 30$	6560
$30 \leq x < 35$	4105
$35 \leq x < 40$	2531
$40 \leq x < 50$	3244
$50 \leq x < 70$	1300
$70 \leq x < 100$	43

Table 1. Age distribution of our user sample.

a few additional apps, the user would need to interact with these new apps to change her fingerprint. Falaki et al. [6] were the first to show that there is an impressive diversity in how people use their smartphones. However, we are not aware of any published attempts to differentiate between users based on this insight. On an abstract level, tracking users via their unique behavior, opens an entirely new discussion regarding privacy.

Indeed, as this paper shows, the sets of used apps are unique for 99.67% of all users in a dataset of over 46000 Android phones. Furthermore, this method is robust under several challenges: First, the usage patterns are quite different from each other even when we only consider the 500 globally most frequently used apps. Second, the Hamming distance between users is so large that starting or stopping to use a few additional apps will not allow the majority of users to disguise themselves. Third, even when the observation is limited to the 60 globally most frequently used apps, it is possible to uniquely identify 99.4% of all users. These findings promise that user identification solely based on interaction with apps might be possible with high accuracy. In our discussion, we outline the next steps towards this exciting though disturbing goal and some alarming implications.

THE MENTHAL DATASET

The Menthal project studies smartphone usage on a very large scale. At its core, it consists of an Android app that logs the users interaction with their smartphone. On the one hand, the app thus provides users with feedback about their phone usage. This use case proved highly attractive, and motivated a great number of people to install our software [1]. On the other hand, the app sends the collected data to our server, for scientific analysis. This data collection process was approved by our ethical review board and required users' informed consent. In addition, some users voluntarily answered questionnaires regarding basic demographic information, life satisfaction and personality [2].

Among others, our app logged meta data on Phone usage, SMS, phone calls and GPS position. Most notably, it recorded the amount of time spent on the phone, the number of unlocks per day, and the frequency and duration of usage per app. Early versions of our app used the Accessibility Service features on Android devices for tracking app usage. Since this required an intricate setup process and extensive permissions, we implemented a method for polling recent tasks from the Android OS. This new method only needs basic permissions

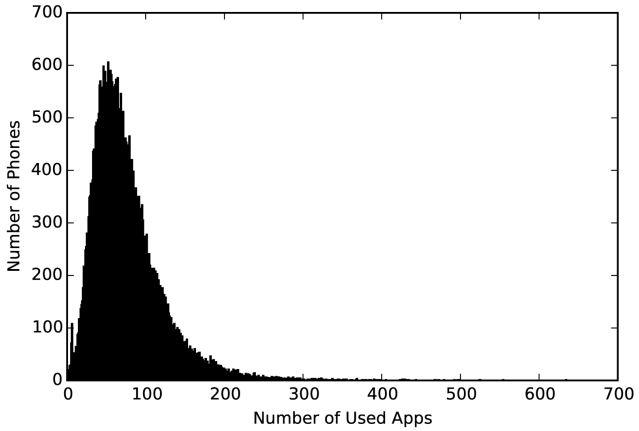

Figure 1. Distribution of the number of used apps per phone. The average is 74.37**.**

and no setup. Besides the polling of recent tasks, we have also used the official Android API for collecting app usage. This, again, requires additional steps from the user and needs extensive permissions. On Android versions newer than 5.0 the polling method does not work and we resort to the API calls. In this work, we do not differentiate between the methods of data collection since they produce equivalent output.

There are other studies that collect similar data in this fashion, e.g. Baeza-Yates et al. [3] who predict the next app some user will start. Ferreira et al. [7] find and investigate "micro-usage" patterns of apps – brief bursts of interactions that last for less than 30 seconds.

Since its launch in January 2014, our app was downloaded more than 380000 times. In this work we analyze the data of 46726 users who created a user account, completed the questionnaires, and provided demographic information about themselves. In particular, we analyze the set of *used* apps. We thus consider all apps that were started at least once on at least one phone during the presence of our app. Each user can thus be associated with her *app signature*, the set of used apps. The observed time intervals differ from user to user. We excluded users with less than one complete day of observed behavior. On average, an app signature considered in this work collects the used apps in 48.61 ± 46.73 days, the longest observation period was 216 days.

User Statistics

The Google Play Store provides app developers with download statistics of their apps. There is a large participant population in the German speaking countries, due to our app being released in English and German. Of the more than 380000 downloads of our app in total, 68.52% came from Germany, 4.52% from Austria, 2.80% from Brazil, 2.74% from Switzerland, and 1.97% from India. We have no data specific to the subset that we investigate in this study but we assume that the general trend still holds.

Among the 46726 users of our sample dataset, 18550 (39.7%) reported being female and 28176 (60.3%) being male. Table 1 shows the age distribution over the dataset. More than half of

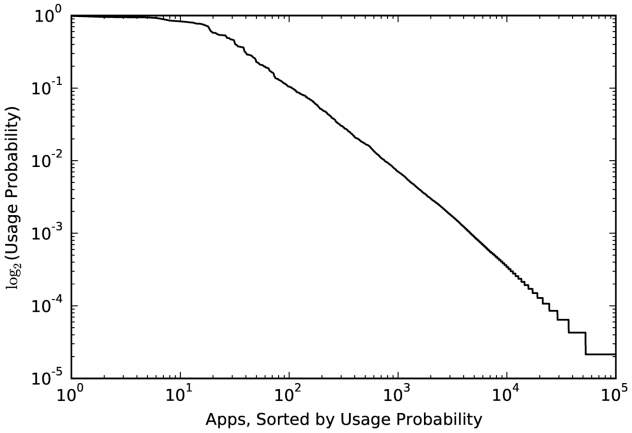

Figure 2. Usage frequencies of top 100 000 apps in our dataset.

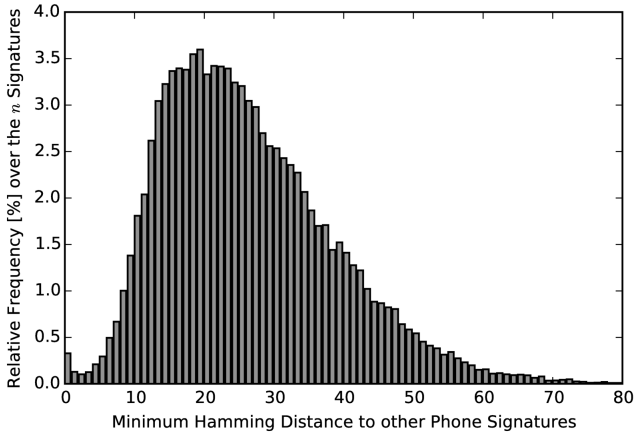

Figure 3. Hamming distance to the closest user's app signature (based on top 500 apps). The average minimum distance is 25.93.

all users are younger than 25 years. Though this information is self-reported by the users and cannot be checked for correctness, we expect it to be rather accurate: Those users willingly answered the optional questionnaire about their personality traits.

App Statistics

As of November 2015, there are more than 1.8 million apps in the Play Store [10]. Among the 46 726 phones in our dataset, there were a total of 146 532 different apps in use. This set opens a vast space of possible app signatures. On our test set, the average number of used apps per phone was 74.37, with a fairly large standard deviation of 44.16. Figure 1 shows a histogram of the number of installed apps per phone.

Most of the apps are used by very few users; only very few apps are highly popular. Figure 2 shows usage frequencies of the 100 000 most frequent apps in our dataset as a log-log plot. There are 20 apps (including ours, by selection bias) that are highly popular among our user base. The most popular app (apart from ours) in our dataset was used on 95.49% of all phones, while the 20th most popular app was used by 57.69% of users. After these 20 apps, the popularity starts to drop even more drastically. E.g., the 50th most popular app was already only used on 22.88%, the 100th most frequent app on 10.44% of all phones. The line of the plot is slightly slanted (in any range), hence implying a super-polynomial relation and between rank and frequency, hence ruling out a power law distribution. Considering the huge number of available apps this is an extreme form of a winner-takes-all market.

FINDINGS

Considering the large number of used apps in our dataset, it should not be surprising that almost every user has a unique app signature. Each user can almost be expected to use a unique app. Hence, to avoid this problem, we restrict ourselves to the set of 500 most frequently used apps for the rest of this paper.

If we consider only the 500 globally most frequently used apps, the signatures remain almost unique. Furthermore, this measure proves rather robust: First, the average Hamming distance to the nearest other user is rather large. Second, even the 60 most frequently used apps suffice to identify most users.

Unique App Signatures

We define the *app signature* of a user to be the set of apps that were started at least once on her phone. If we thus consider the 500 globally most frequently used apps, we can represent each user as 500 dimensional binary vector with the *i*th bit set if and only if she used the *i*th most frequent app. The restriction to the 500 most frequent apps is merely motivated by the fact that the 500th most frequent app was only used on 1.69% phones.

If the app signatures of two users are identical, they cannot be differentiated and we call them *anonymous*. However, this hardly ever happens. Among the 46 726 users there were only 153 anonymous users leaving 99.67% of users with a unique signature. The anonymous users shared a total of 20 app signatures; the number of users with one of those non-unique signatures were below ten for all but three of these signatures. Even for those 153 anonymous users, it is thus possible to differentiate them from almost all other users.

Anonymity Unachievable

The signatures based on the top 500 apps are so robust that it is highly difficult to achieve anonymity through deliberate (non-) usage of apps. In order to quantify the similarity between two users, we define their distance as the Hamming distance between their app signatures. That is, the number of apps used exclusively by either of the two users. Hence, a user is anonymous in our dataset, if there is some other user with distance 0. Furthermore, this measure captures 'how unique' a certain user is by considering the distance to the nearest neighbor.

Figure 3 shows the distance to the closest other user for each user in our dataset. We can see the 153 users that are not uniquely identifiable as a spike for distance 0. Overall, the average minimum distance is 25.93, and for 95.27% of the users the closest different user has Hamming distance at least 10. These users would thus have to change their behavior regarding at least ten different apps, in order to achieve anonymity.

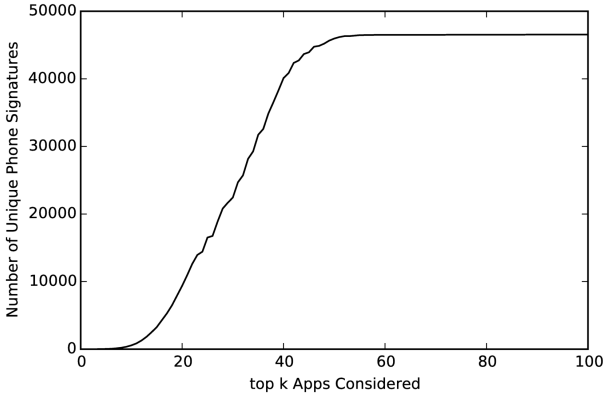

Figure 4. Number of unique app signatures when only the top-k most frequently used apps are considered. For 10 or less, there are only a few different signatures, while for 60 and more, almost all users have a different signature.

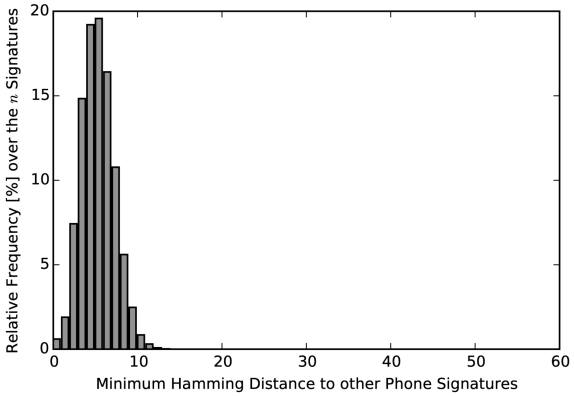

Figure 5. Distribution of the Hamming distance to the closest app signature (based on top 60 apps) for all 46 726 phone signatures. The average minimum distance is 4.9.

(This assumes they know the exact app signature of the most similar user.) For 99.54% of the users, the closest other user has Hamming distance at least 2.

Top-k Apps

Observing the top 500 most common apps is rather generous. In fact, a much smaller set of most common apps suffices to differentiate between almost all users in our dataset.

Figure 4 shows the number of unique app signatures based solely on the top-k most frequent apps for all $k < 100$. For $k < 60$, the number of unique app signatures drops dramatically. Astonishingly however, the signatures based on the top 60 most frequent apps suffice to almost perfectly differentiate between the users in our dataset. The number of anonymous users increases from 153 to 281, while the number of signatures that are not unique increases from 20 to 72. Figure 5 shows the distance to the closest different phone based on the top 60 signatures. As was to be expected, this distance drops, now that we use only 60 apps, its average is now 4.9.

DISCUSSION

This short paper investigates the usage of the 500 most frequently used apps on a sample of 46 726 Android phones. For this, we model a user as the static set of apps that she used at least once during our observation period of 48.61 days on average. This measure differentiates between 99.67% of all users, i.e. there are a total of only 153 anonymous users in our dataset. Furthermore, the resulting signatures are surprisingly robust, the average minimum Hamming distance to a different user being 25.93.

As this paper shows, it is rather challenging to assume the app usage pattern of even the nearest neighbor, by mimicking her behavior. One would have to avoid a fair number of previously used apps, while adding others. This, of course, would require the knowledge of the used apps of your neighbor and the willingness (or ability) of the user to change her actual behavior. Even more so, there is no "John Doe", no average smartphone user whose identity we could try to mimic in order to seek anonymity. Behavior, as it turns out, is highly individual.

Our findings raise a serious privacy concern: In older versions of Android (< 5.0), polling active apps requires relatively few permissions. This method has been used as an attack vector by "spyware-like" apps such as flashlight apps that collect data about their users [8]. Even on recent Android versions, though, identifying currently used apps is easily possible, as most apps generate some network traffic (due to their intended functionality, tracking, or to display ads). Xu et al. [12], among others, have shown that it is possible to infer the used apps based on their network traffic. Hence, service providers or anyone hosting free wifi can possibly identify even those users that spoof their hardware addresses or change numbers or phones repeatedly, as long as they do not change their behavior. This can also be used by "proxy-like" apps that monitor network traffic of the phone to obtain these insights. Furthermore, our results imply that one can not anonymize a dataset as in [12] by removing phone identifyers like IMEI or MAC addresses.

Using only the top 60 most frequently used apps, we were still able to differentiate between 99.4% of users in our dataset with a unique fingerprint based on these apps. If, due to some bug in a mobile browser, it is possible to gather information about the presence or absence of particular apps (e.g. using deep links) and assuming that our results extend to installed apps, a small number of tests may very well suffice to uniquely identify a user from the mobile browser. In fact, this was possible in Chrome for Android until the beginning of 2015. Hence, it is paramount for the future development of browsers and the Android SDK to keep this issue in mind.

Next, we intent to investigate how well the sets of *installed* apps discriminate between different phones, in contrast to used apps. This question is of high practical relevance, as any installed app can easily obtain a list of all installed apps on the phone. Furthermore, we plan to investigate how used and installed apps for any one particular user change over time. In the case of app usage we would like to find the minimal observation period (e.g. 1 hour, 1 day, 1 week) before we can identify users, i.e. differentiate between different users as well as decide if we have seen a particular user before. In the case of installed apps we plan to use only snapshots at

different moments in time and not the full tracking period as a cumulative set of installed apps.

Initial experiments showed promising results in trying to infer the gender of users based on used apps. We also try to extend this to predicting personality traits or user interests.

ACKNOWLEDGEMENTS

IA is funded by the Federal Ministry of Education and Research (BMBF), project number 00160280. KB has been supported by DFG and Collaborative Research Center SFB 876. IA and KB were funded by the University of Bonn and through numerous short term grants.

REFERENCES

1. Ionut Andone, Konrad Błaszkiewicz, Mark Eibes, Boris Trendafilov, Christian Montag, and Alexander Markowetz. 2016a. Menthal - Running a Science Project as a Start-Up. In *Computing in Mental Health, Workshop at CHI 2016*. ACM.
2. Ionut Andone, Konrad Błaszkiewicz, Mark Eibes, Boris Trendafilov, Christian Montag, and Alexander Markowetz. 2016b. Menthal: A Framework for Mobile Data Collection and Analysis. In *Ubicomp/ISWC 2016 Adjunct*. ACM. DOI: http://dx.doi.org/10.1145/2968219.2971591
3. Ricardo A. Baeza-Yates, Di Jiang, Fabrizio Silvestri, and Beverly Harrison. 2015. Predicting The Next App That You Are Going To Use. In *Proc. of WSDM 2015*. 285–294. DOI: http://dx.doi.org/10.1145/2684822.2685302
4. Hristo Bojinov, Yan Michalevsky, Gabi Nakibly, and Dan Boneh. 2014. Mobile Device Identification via Sensor Fingerprinting. *CoRR* abs/1408.1416 (2014). http://arxiv.org/abs/1408.1416
5. Peter Eckersley. 2010. How Unique Is Your Web Browser?. In *Proc. of PETS 2010*. 1–18. DOI: http://dx.doi.org/10.1007/978-3-642-14527-8_1
6. Hossein Falaki, Ratul Mahajan, Srikanth Kandula, Dimitrios Lymberopoulos, Ramesh Govindan, and Deborah Estrin. 2010. Diversity in smartphone usage. In *Proc. of MobiSys 2010*. 179–194. DOI: http://dx.doi.org/10.1145/1814433.1814453
7. Denzil Ferreira, Jorge Gonçalves, Vassilis Kostakos, Louise Barkhuus, and Anind K. Dey. 2014. Contextual experience sampling of mobile application micro-usage. In *Proc. of MobileHCI 2014*. 91–100. DOI: http://dx.doi.org/10.1145/2628363.2628367
8. Tom Fox-Brewster. 2015. Check the permissions: Android flashlight apps criticised over privacy. (2015). https://www.theguardian.com/technology/2014/oct/03/android-flashlight-apps-permissions-privacy.
9. Nick Nikiforakis, Alexandros Kapravelos, Wouter Joosen, Christopher Kruegel, Frank Piessens, and Giovanni Vigna. 2013. Cookieless Monster: Exploring the Ecosystem of Web-Based Device Fingerprinting. In *SP 2013*. 541–555. DOI:http://dx.doi.org/10.1109/SP.2013.43
10. Statista. 2015. Number of Available Applications in the Google Play store. (2015). www.statista.com/statistics/266210/number-of-available-applications-in-the-google-play-store/, accessed 2016-02-04.
11. Gary M. Weiss and Jeffrey W. Lockhart. Identifying User Traits by Mining Smart Phone Accelerometer Data. In *Proc. of SensorKDD 2011*.
12. Qiang Xu, Jeffrey Erman, Alexandre Gerber, Zhuoqing Morley Mao, Jeffrey Pang, and Shobha Venkataraman. 2011. Identifying diverse usage behaviors of smartphone apps. In *Proc. of SIGCOMM 2011*. 329–344. DOI: http://dx.doi.org/10.1145/2068816.2068847