Should Smartphone Users Mock Apps?

Nick DiRienzo and Geoffrey Challen
Department of Computer Science and Engineering
University at Buffalo
{nvdirien,challen}@buffalo.edu

ABSTRACT

Smartphones represent the most serious threat to user privacy of any widely-deployed computing technology. Unfortunately, existing permission models provide smartphone users with limited protection, in part due to the difficulty users have distinguishing between legitimate and illegitimate use of their data. A mapping app may upload the same location information it uses to download maps (legitimate) to a marketing agency interested in delivering location-based ads (illegitimate). However, armed with the right technology users can turn apps' interest in personal data against them by intentionally manipulating the data that they expose. We refer to the intentional substitution of real data with artificial data intended to alter an apps perception of a user as *mocking* to differentiate this approach from other privacy-motivated techniques that focus on concealing data. In this paper, we explore the desirability and implications of this approach, present results from a survey suggesting that many users are interested in mocking apps, and discuss ethical and practical issues related to widespread app mocking.

1. INTRODUCTION AND MOTIVATION

Recent studies show that privacy is one of today's smartphone users' top concerns with their devices, second only to battery life [2]. A plurality of 43% of users are *not* willing to share any information about themselves with a company in exchange for a free or subsidized app, despite ad-driven free apps being common on mobile app marketplaces. And as smartphone apps get smarter they will be able to determine more about our lives through passive observation. Mapping services already know where we are, sensors reveal whether we walked or drove, social networking apps track our relationship with the person we came to meet, and payment apps reveal what we did together. And as analytical approaches that fuse data from multiple sensors improve, smartphones will likely reveal even deeper facts about us: the strength of our friendships, our devotion to our job, even our level of happiness.

Today's smartphone platforms have attempted to address this loss of privacy through permission mechanisms that limit apps' access to user information. Unfortunately, this approach has many well-documented problems: apps request permissions they don't need [4, 5], and users do not understand the implications of permissions that apps request [6]. The "take-it-or-leave-it" model used by Android provides no options for users uncomfortable with the permissions an app requests other than not to install it. Even a more selective "take-it-or-break-it" approach [10] that could allow users selectively to enable individual permissions is no cure. Users are still forced to make poorly informed tradeoffs between privacy and functionality, as it is unclear how an app might behave if denied access to information or fed random values. As a result, users tend to give apps the permissions they request.

One fundamental problem with all permission-based approaches to protecting privacy on smartphones is that many apps legitimately require access to certain kinds of sensitive information to function properly. When I am lost, my mapping app must know where I am in order to route me to my destination. When I am monitoring my fitness, my pedometer must be able to access the accelerometer to count the number of steps I take each day. It is unreasonable to expect users to be able to distinguish between legitimate and illegitimate information requests, and burdensome and error-prone to ask them to enable data sources only when they feel comfortable with what a particular app is doing.

Obviously users can always choose not to use apps with which they feel uncomfortable, and there are many projects looking at how to make safer app marketplaces by preventing malicious apps that *only* want to steal personal information. However, this focus on malicious apps obscures a harder truth: even legitimate data collected by non-malicious apps represents a privacy risk. For a typical user who wants to read email, browse the web, take pictures, and use social networking and messaging clients, legitimate data collection by legitimate apps still constitutes a significant privacy risk. Even after preventing unnecessary data collection by legitimate apps, a tradeoff between privacy and usability remains. The only options remaining are to remove useful apps or stop using the smartphone entirely—both unattractive.

Rather than trying to limit or control the flow of accurate user data, another approach is to overwhelm the accurate data with inaccurate data. If the inaccurate data is random, this has the effect of obscuring the accurate data. For example, if an app is using location updates to try and determine a users work schedule, inserting random locations would make that more difficult. However, the inaccurate data can also be designed to achieve a specific objective. We call this process *mocking* and the data used *mocked* data. Continuing with the same example, mocking the app would mean injecting data that would cause the app to reach an incorrect conclusion rather than no conclusion.

Mocking differs fundamentally from privacy. While privacy aims to limit access to data under the assumption that less data reveals less about users, understanding privacy implications still requires smartphone users to answer difficult questions. If I install and use this app, what will it be able to determine about me? How much of the data that this app is collecting is really necessary? What are the privacy implications of even the legitimate data that this app is collecting? Accurate answers to these questions remain elusive at best. There are billions of dollars at stake for companies in determining how to do more accurate mobile data analytics, and few if any have a business interest in divulging either how their algorithms work or what they know about us. While new tools help smartphone users determine *how much* data smartphone apps collect and even where that data goes, *what it reveals* remains uncertain. In contrast mocking reduces the power of legitimate data by injecting enough mocked data to achieve user-defined objectives and has the potential to change "I don't know what this app knows about me" uncertainty into "I know what this data will cause this app to conclude" certainty. And unlike privacy, which requires hiding data and thus potentially impacting apps' functionality, mocking ensures that apps continue to function normally during each mocking session, making it simpler for users to understand and use.

In this paper we explore the desirability and implications of widespread app mocking. We begin by presenting mocking scenarios in the next section to help make our discussion more concrete and further illustrate the differences between mocking and privacy. Section 3 continues with a brief discussion of the feasibility of this approach, noting that for Android devices mocking can be performed either with our without modifications to the underlying smartphone platform software. In Section 4 we present survey results indicating that users are aware of and concerned by smartphone data collection and willing to utilize mocking techniques. Section 5 continues by raising some of the ethical issues raised by mocking, after which Section 6 concludes the paper.

2. MOCKING SCENARIOS

To make our earlier description of app mocking more concrete, consider these four scenarios:

- Bob wants to appear more active. On Monday he takes a walk to get some exercise. The next day, he doesn't take a real walk, but while he is sitting at his desk his smartphone mocks another walk.
- Alice wants to appear more healthy, but on Monday she visits a fast food restaurant where she enjoys an unwholesome meal. As she eats, however, her smartphone mocks a visit to a nearby organic salad delicatessen.
- Teenager Jerry's parents use a smartphone app to monitor his late-night ventures and to ensure that he returns home by an imposed curfew. One night Jerry remains out on the town later-than-allowed with his friends, but his phone has already mocked him dutifully returning home on time.
- Carol's employer uses her phone to monitor her attendance. During the workday, she surreptitiously slips out for a latte with a friend. Meanwhile, her phone records her apparent continued presence at her desk.

These examples highlight the difference between *privacy* and *mocking*. None of our characters' objectives can be accomplished through privacy, since in each case achieving the objective requires using data to manipulate an app. Of course, Jerry could remove the app that his parents installed, as could Carol, but his parents and her employer would likely notice. In the cases of Alice and Bob, we can assume that they have been asked or incentivized to install these health-monitoring apps but may not feel fully-comfortable with their operation. Similarly, however, removing or disabling the apps may not be an attractive or even feasible option. We include a more in-depth discussion of the practical and ethical implications of mocking in Section 5 after establishing the feasibility and desirability of this approach.

3. FEASIBILITY

A natural question to ask about data mocking is if a system was developed that provided this feature, could it be deployed to a large number of users? We believe that the answer is yes, at least on Android smartphones. Below we briefly discuss two ways to mock apps.

3.1 Platform Support

One way to implement mocking is to add support to smartphone platforms. For example, when an app requested the device's current location from a platform service, the service could either return the user's real location or a mocked location. Given the locked-down nature of major smartphone platforms today, this means

	Address		Social Network		Activity Level		Income		Weight	
Accuracy	exact	81	five best friends	49	**quantity and type**	20	**$1**	6	**1 lb**	8
	street	6	some friends	25	how much	20	**$100**	24	**10 lb**	12
	neighborhood	12	how social	23	activity level	43	$10,000	31	category	39
	nothing	0	nothing	2	nothing	14	nothing	37	nothing	39
High	16		19		25		9		36	
	7		12		23		13		14	
Comfort	14		29		28		23		19	
	24		19		9		10		9	
Low	37		18		13		42		19	
Mocking Yes			more	14	less	15	lower	54	higher	6
	different	58	fewer	25	more	9	higher	10	lower	19
No	true	41	true	60	true	74	true	34	true	73

Table 1: **Detailed mocking survey results.** All values are percentages of the 91 respondents. Levels of knowledge we considered to be unreasonable are marked in bold in the accuracy row.

that mocking would require the cooperation of companies such as Google, Apple, and Microsoft that maintain the dominant three smartphone platforms available today. Unfortunately, we expect that all smartphone platform providers have an interest in perpetuating the exposure of personal information to apps that data mocking is intended to frustrate, making them unwilling to implement this feature.

Another option is to use an open-source platform, which would limit mocking to the 80% of smartphones worldwide running Android [1]. Several previous projects with similar goals including AppFence [8], MockDroid [3], and Android record and replay [7] have shown the feasibility of this approach through platform modifications, although neither implemented mocking as we have described it. However, while utilizing Android would allow platform changes required to implement mocking to be implemented, deploying them to users would still be challenging. We expect that even if they are interested in mocking, few users are willing or able take the steps required to replace the built-in platform software—commonly referred to as "rooting" or "jailbreaking".

There are two potential ways around this roadblock. First, mocking could be integrated into popular alternate Android platform distributions such as CyanogenMod. While this community is small, they may be disproportionately interested in mocking given their willingness to void their warranties and the intentions of smartphone device manufactures. Even a small amount of mocking could lead the smartphone privacy conversation in the right direction. Second, at some point smartphones may be required to provide more configurability at the platform level to support apps, similar to the way that desktop operating systems allow apps to install device drivers, thus potentially opening the door for widespread distribution of data mocking.

	Attribute Count					
	0	**1**	**2**	**3**	**4**	**5**
Unreasonable Fear	48	33	16	2	0	0
Uncomfortable	24	19	18	16	12	11
Interested in changing	18	22	19	21	11	10

Table 2: **Summary of survey results.** Aggregates are shown for the three specific questions addressed in Section 4.2. All values are percentages.

3.2 App Rewriting

A more promising alternative that avoids the need to modify the underlying smartphone platform is to utilize the ability to recover the source of Android APKs through decompilation. Access to the resulting source files would allow rewriting API calls to return mocked data, producing a mockable version of the original app. While decompilation has limitations, we are excited by this technique and actively exploring this approach.

4. SURVEY

To gauge interest in mocking we distributed an IRB-approved survey to students, faculty and staff of the University at Buffalo. No incentives were provided for completing the survey, and all respondents were required to indicate consent before proceeding to the questions. Over four days, we recorded 91 responses.

4.1 Questions

Table 3 summarizes the survey we distributed. It had three parts, each consisting of questions concerning five personal attributes: home location, weight, activity level, income level, and sociability. The goal of the first part was to assess how aware respondents were of the information smartphones apps could collect about them. Respondents were instructed to assume that the hypothetical app had been installed and granted the permissions it requested. The goal of the second part of

the survey was to assess how comfortable respondents were with the data smartphone apps could collect about them. Finally, the third part assessed interest in mocking by determining how interested respondents were in misleading apps about the five attributes.

We intentionally chose a range of attributes. We considered home address and social network as straightforward to determine for an app with the right permissions, in this case the ability to track user's location (home address) or observe who they communicate with (social network). At the time we considered activity level to be more difficult to determine, although now that activity recognition has been integrated into widely-available libraries such as Google Play Services this may be much simpler to measure using the accelerometer. Finally, we chose two attributes that we were not sure could actually be determined by smartphones: income and weight.

Based on the capabilities of current smartphones we classified answers to the accuracy questions as either reasonable or unreasonable. As an example, we considered it unreasonable that an app could know a user's income to within $1 / year or their weight to within 1 lb. However, it is possible that by using the accelerometer and studying a user's gait their weight could be estimated, and the widespread adoption of smartphone-based payment systems such as Google Wallet along with socioeconomic map-matching may make income levels estimable soon. So even the answers we marked as unreasonable may not remain so for long.

4.2 Results

Table 1 shows detailed results of our mocking survey. When analyzing the results, we were interested in three questions matching the three sections of our survey. First, how reasonable were respondents fears about information that apps might be able to determine? Second, how comfortable were they sharing data with apps? And finally, were respondents interested in modifying the data-driven impressions app might form of them? Table 2 reports aggregate results relevant to these three questions.

First, we found that respondents to be reasonably suspicious of what apps might know about them, with 52% indicating that an app might know at least one personal attribute to a level that we marked as unreasonable today but only 18% indicating that apps might know two attributes of unreasonable levels. The two unreasonable attributes most-frequently reported as knowable by respondents were their income to $100 / year (24%), and the quantity and type of the exercise they engaged in (20%). Overall, users' intuition about what attributes were easy to determine and what were hard matched ours, reflected by the accuracy percentages in the table. No users thought an app would not be able to determine anything about their home address, whereas 39% didn't think an app could determine anything about their weight.

Second, our survey showed that many respondents were uncomfortable with smartphones knowing these aspects of their personal lives. Only 24% were comfortable, defined as a score of 2 or above on the 1–5 scale, with all five attributes, and a majority (57%) were uncomfortable with two or more. Our results match the privacy concerns reported by smartphone users to other surveys [2]. Reported comfort levels on individual attributes were also interesting, with users seeming the least comfortable with smartphones knowing their home address—which is possible—and their income—which, at least today, may not be. Comfort levels regarding knowledge of a users social network were evenly distributed.

Finally, when asked about mocking, of the 91 users that completed the survey, 82% wanted to mock at least one attribute and 60% wanted to mock two, with mocking users requesting an average of 2.6 mocking attributes each. Interest in mocking different attributes was well distributed, with the percentage of mocking responses per attribute varying from a low of 26% for activity level to a high of 66% for income.

Unsurprisingly, users were most interested in mocking attributes that they were uncomfortable with their smartphone knowing, such as home address and income level. Surprisingly, most users seemed to want to appear to make *less* money than they actually do, which is not what we expected. We speculate that this may be because users believe that they will see fewer ads if advertisers believe that they are poor. In any case, it shows that it may be difficult to determine what changes users see as desirable.

4.3 Limitations

Obviously our small survey has many weaknesses. We surveyed a population exclusively drawn from our university and overrepresenting students in the Department of Computer Science and Engineering, groups that may be more aware of and concerned by app data use and more open to experimental approaches. So while these survey results line up with more comprehensive efforts at gauging user privacy concerns, these results may not be valid at a societal scale. However, taken at face value the results hint that smartphone users have reasonable expectations about what smartphone apps might be able to learn about them, are uncomfortable with apps knowing these things, and may be interested in misleading apps.

5. DISCUSSION

Even if feasible and desirable mocking raises a set of unique ethical questions related to the relationship between smartphone users and apps, discussed below.

5.1 Is Mocking Cheating?

A common reaction by many to mocking is to conflate it with cheating, given that it involves misleading apps about a users true nature. This response, however, begs the question: what are the rules of the game? Cheating must be defined with respect to an agreement, and we are not convinced that users have actually agreed to provide an unlimited amount of information about their personal lives to smartphone apps. While installing an app does involve allowing it to access certain information, we believe that it is reasonable for users to expect that apps request information required by the features they provide and use it only to provide those features. Unfortunately, particularly once the information leaves the device, users quickly lose control over their data.

Others point out that mocking likely violates the terms of service (TOS) that app users are frequently required to agree to during the installation process. We have not yet performed the detailed examination of common smartphone app TOS required to determine whether this is true, but would be surprised if TOS agreements could be written in ways preventing users from misleading apps. The reason is that there is considerable overlap between mocking behaviors and things that users might legitimately do in order to alter how smartphone apps perceive them. For example, can an app TOS prevent a user from leaving their smartphone home during a night out? Require that users keep their smartphones on their person all the time? Prevent a user from loaning their device to a friend for a period of time? Or require that a user perform all of their smartphone interaction with the same device? The natural answer to these questions seems to be no, but this begs the question of whether there is a meaningful difference between actually leaving the device at home or bringing it but mocking apps into thinking it is at home.

Another common objection is that the current smartphone app ecosystem depends on collecting user information in order to subsidize app development, most commonly by using personal information to embed targeted advertisements in apps, and that users benefit from lower app prices as a result. There are two problems with this argument. First, as mentioned earlier smartphone users have indicated on surveys that they are not comfortable with this model of subsidizing apps through personal data collection. Second, believing that this model is really a good deal for smartphone users requires them to trust the same companies that are actively trying to monetize this information. A sense that they are not receiving adequate compensation for their information may drive user discomfort with this business model. In any case, because smartphone users have different privacy concerns and expectations, not every user should be required to trade data for service.

5.2 Is Mocking Safe?

A more serious concern with mocking concerns the effect it might have on apps, particularly ones that are health-related. If mocking confuses an app designed to remind a user to take a pill, it could have serious health consequences. Here it is important to distinguish between the possible side-effects of mocking on legitimate apps and the intentional effect of mocking on apps that the user is intending to mislead. For example, if a doctor asks a patient who wants to get fit to install a pedometer to help increase their activity level and the patient chooses to mislead this app with mocked activities, then the main problem is not really the mocking feature. If the user wants the app to help them become healthier, they will cooperate; if not, they can always refuse to be monitored altogether.

We believe that remaining safety concerns can be addressed through careful system design. Mocking systems should allow users to configure which apps to mock to avoid mocking safety-critical apps. It may also be helpful to allow apps to issue explicit requests to not be mocked which users could approve or ignore as a reminder to adjust mocking settings on a per-app basis.

5.3 What Effect Would Mocking Have?

Finally, we consider the effect that data mocking would have on smartphone data collection and privacy if deployed on a significant number of devices. First, we would expect to see an interest among app developers in deploying countermeasures to detect or eliminate mocked data. While single-app attacks may be defeated by carefully engineering the mocking system to provide consistent false data, a more difficult or impossible set of attacks are launched by colluding with other devices or with surrounding infrastructure. Interdevice collusion is more feasible, since it could be launched by cooperating instances of the same app. Foiling these attacks might require mocking cross-device interaction to fool the local app, or simply disabling interfaces such as Bluetooth allowing device-to-device communication.

Collusion with the infrastructure would represent a more serious challenge to data mocking. As an example, if mobile data networks began reporting smartphone user's location directly to app providers apps could use this information to pierce the mocking context by comparing the location being reported to them by the smartphone to infrastructure-reported location. While this type of collusion is the most effective way to shut down our mocking approach, it would also represent an unprecedented level of cooperation between network providers and the companies selling apps and services. We anticipate that these types of agreements would be highly-unattractive to smartphone users already concerned about their privacy.

What can data collected by your smartphone reveal about you?
The questions below assess how much you think your phone is able to determine about you without asking. All the questions assume that you have installed the application and granted it the permissions it requested.

- Without asking, what could an application determine about your yearly income?
  - An application could predict my income exactly, to within $1 / year.
  - An application could predict my income to within $100 / year.
  - An application could predict my income to within $10,000 / year.
  - An application could not determine anything about my income level.
- ... your weight?
- ... your social network?
- ... your activity level?
- ... where you live?

What are you comfortable with applications knowing about you?
The questions below assess how comfortable you are with smartphone application being able to determine the same things about you we asked about in the first section. **Please indicate your comfort level between not comfortable at all (1) and completely comfortable (5).**

- How comfortable are you with smartphone applications knowing your yearly income?
- ... your weight?
- ... about your social network?
- ... your activity level?
- ... where you live?

Would you like to alter what your smartphone knows about you?
These questions assess your interest in altering what your smartphone knows about you. Assume that a system exists that would allow you to change the qualities as indicated by the questions.

- If a smartphone application could accurately determine my income level
  - I would like to appear to have a lower yearly income than I actually do.
  - I would like to appear to have a higher yearly income than I actually do.
  - I am comfortable revealing my true income level to the application.
- ... my weight
- ... my social network
- ... my activity level
- ... where I live

Table 3: **Mocking survey questions.** Respondents were asked three groups of questions about five aspects of their personal lives their smartphone could observe. For each group one sample question and answers is shown.

Our ultimate hope in discussing smartphone mocking is to initiate a conversation about what our personal data is worth. Today, because smartphone users lack effective tools to control the data they provide to apps, they are effectively surrendering their personal information without receiving anything in return. So while this data is clearly worth something, as evidenced by advertisers scrambling to develop novel location-based analytics, as long as we give it away for free we will never know how much. If mocking causes apps to begin to be suspicious of the personal data they can collect, this may help make legitimate information about users even more valuable.

6. CONCLUSION

In this paper, we have introduced the notion of data mocking to help users protect their digital personas on their mobile devices. We have argued that mocking is feasible, shown survey data indicating that it is desirable, and discussed the ethical implications of the approach. We are completing a prototype of a system implementing data mocking and looking forward to testing it on the PhoneLab smartphone testbed [9].

7. REFERENCES

[1] Global Smartphone Market Share Platform. http://www.businessinsider.com/iphone-v-android-market-share-2014-5.

[2] Mobile Privacy: A User's Perspective. http://www.truste.com/why_TRUSTe_privacy_services/harris-mobile-survey/.

[3] A. R. Beresford, A. Rice, N. Skehin, and R. Sohan. Mockdroid: trading privacy for application functionality on smartphones. In Proceedings of the 12th Workshop on Mobile Computing Systems and Applications, HotMobile '11, pages 49–54, New York, NY, USA, 2011. ACM.

[4] W. Enck, P. Gilbert, 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. In Proceedings of the 9th USENIX conference on Operating systems design and implementation, OSDI'10, pages 1–6, Berkeley, CA, USA, 2010. USENIX Association.

[5] A. P. Felt, E. Chin, S. Hanna, D. Song, and D. Wagner. Android permissions demystified. In Proceedings of the 18th ACM conference on Computer and communications security, CCS '11, pages 627–638, New York, NY, USA, 2011. ACM.

[6] A. P. Felt, E. Ha, S. Egelman, A. Haney, E. Chin, and D. Wagner. Android permissions: user attention, comprehension, and behavior. In Proceedings of the Eighth Symposium on Usable Privacy and Security, SOUPS '12, pages 3:1–3:14, New York, NY, USA, 2012. ACM.

[7] J. Flinn and Z. M. Mao. Can deterministic replay be an enabling tool for mobile computing? In Proceedings of the 12th Workshop on Mobile Computing Systems and Applications, HotMobile '11, pages 84–89, New York, NY, USA, 2011. ACM.

[8] P. Hornyack, S. Han, J. Jung, S. Schechter, and D. Wetherall. These aren't the droids you're looking for: retrofitting android to protect data from imperious applications. In Proceedings of the 18th ACM conference on Computer and communications security, CCS '11, pages 639–652, New York, NY, USA, 2011. ACM.

[9] A. Nandugudi, A. Maiti, T. Ki, F. Bulut, M. Demirbas, T. Kosar, C. Qiao, S. Y. Ko, and G. Challen. Phonelab: A large programmable smartphone testbed. In 1st International Workshop on Sensing and Big Data Mining (SenseMine 2013), November 2013.

[10] M. Nauman, S. Khan, and X. Zhang. Apex: extending android permission model and enforcement with user-defined runtime constraints. In Proceedings of the 5th ACM Symposium on Information, Computer and Communications Security, ASIACCS '10, pages 328–332, New York, NY, USA, 2010. ACM.