Supporting Privacy-Conscious App Update Decisions with User Reviews

Yuan Tian
Carnegie Mellon University
yuan.tian@sv.cmu.edu

Bin Liu
Carnegie Mellon University
bliu1@cs.cmu.edu

Weisi Dai*
Google
weisi@google.com

Blase Ur
Carnegie Mellon University
bur@cmu.edu

Patrick Tague
Carnegie Mellon University
tague@cmu.edu

Lorrie Faith Cranor
Carnegie Mellon University
lorrie@cmu.edu

ABSTRACT

Smartphone app updates are critical to user security and privacy. New versions may fix important security bugs, which is why users should usually update their apps. However, occasionally apps turn malicious or radically change features in a way users dislike. Users should not necessarily always update in those circumstances, but current update processes are largely automatic. Therefore, it is important to understand user behaviors around updating apps and help them to make security-conscious choices. We conducted two related studies in this area. First, to understand users' current update decisions, we conducted an online survey of user attitudes toward updates. Based on the survey results, we then designed a notification scheme integrating user reviews, which we tested in a field study. Participants installed an Android app that simulated update notifications, enabling us to collect users' update decisions and reactions. We compared the effectiveness of our review-based update notifications with the permission-based notifications. Compared to notifications with permission descriptions only, we found our review-based update notification was more effective at alerting users of invasive or malicious app updates, especially for less trustworthy apps.

Categories and Subject Descriptors

H.5.2 [**User Interfaces**]: User-centered design

General Terms

Security and Privacy

Keywords

Security and privacy; Smartphone application

*Work done while at Carnegie Mellon University.

SPSM'15, October 12, 2015, Denver, CO, USA.
 ISBN 978-1-4503-3819-6/15/10...$15.00
DOI: http://dx.doi.org/10.1145/2808117.2808124

1. INTRODUCTION

Many smartphone app markets, such as Google Play, give users the option of either automatic or manual updates for their apps. Mobile operating systems can have potentially more invasive permissions than desktop applications [26] and can greatly complicate users' decision-making process around mobile app updates. A user is typically prompted with an app update notification in two cases: (1) if the user opts for manual app updates or (2) if the user ops for automatic updates but an updated app requires a different set of permissions.

App updates can be very critical to security and privacy. Most of the time, users should update their apps to fix important security bugs, such as when the Facebook app fixed the Heartbleed OpenSSL bug [23]. On the other hand, updates may occasionally create problems for security and privacy, in which case the user probably should not agree to an app update. A rogue application developer may change an application to deviate from previous behavior in a way that violates a user's privacy expectations, possibly as a result of the app being sold or the developer turning malicious. Examples include attackers who buy apps from other developers and insert malicious content [2, 29] and the more common scenario when an otherwise trustworthy developer includes a data-hungry ad library that requires more permissions, both of which will prompt an app update. Although the first example is relatively rare and a truly malicious update would probably be reported quickly, the second would typically not be seen as malicious or violate the Terms of Service even if many users are uncomfortable with the change. As a specific example, many users were uncomforable when Facebook updated their mobile app to include permissions allowing it to access a user's contacts and calendar [25]. In any of these cases, however, the user facing the update notification has no easy way of knowing why the app is being updated, what changes are included, or whether the changes have any security or privacy implications. Hence, the user's important decision of whether to update the app is not supported by useful information.

Currently, app update notifications only include a typical permission screen, which has been found in prior work to have limited effectiveness in helping users make privacy-conscious decisions [16, 11]. We note that Google has recently announced a change to the Android permission model in the next OS version, moving toward in-context permis-

sion prompts instead of one-time permission requests during installation. While this will undoubtedly help users make more privacy-conscious decisions about individual app actions, it does not address the problem of changes to app behavior due to updates. Even with the updated permission model, there are two primary issues. First, the effectiveness of permission prompts depends on how developers implement them – overly frequent prompts could become overwhelming, leading users to ignore them, or prompts without clear contextual meaning could be confusing. Second, users have more reason to trust an installed app that they've been using for a while, so they may be more willing to approve new permission prompts after an update.

Combining the various aspects of permissions and app updates, we see that users need to make a complicated decision about whether to update an app given only limited information about the implications of the update and whether the update will change functionality, permission usage (versus permission request), security features, and sensitive data collection. We thus see that app update management is an important problem in understanding security and privacy concerns in mobile devices. Moreover, we believe that it is essential to provide users with supporting information to help them make meaningful decisions regarding major or minor changes that may have significant security or privacy implications.

As a first important step toward these goals, we note that users' behaviors and opinions around update notifications are not well understood. Although previous work has explored user behavior at the time of app installation [11], we posit that users are much less careful about updates because they trust what they previously installed. Prior to installing an app with a desired functionality, users often search for various options and carefully compare features and potential privacy issues. Contrastingly, app updates are largely or completely automatic. Since the update process is quite different from app installation, we decided to conduct a user study to better understand how users manage and make decisions about app updates. In this study, we ask the following questions about user behavior and perception with respect to Android app updates.

1. Do users manage their app updates automatically or manually?
2. What information do users need or care about when updating an app?
3. Can app update notifications give users more useful information regarding an app's sensitive activities and thereby impact users' decision making?

Our initial survey included 300 participants and asked about users' experiences and understanding of the app update process. One of the outcomes of this survey was the idea that users could significantly benefit from hearing about the experiences of other users. We thus designed a modified update notification mechanism that includes crowd-sourced update reviews from other users, providing a sort of collective intelligence that is independent of any information provided directly by the developer. We then ran a 96-participant field study to test our modified update notifications, which contained highly-rated negative reviews from other users in the app market. Our results show that the additional crowd-sourced information prompted more consideration and reaction from users, especially for less-trusted apps. While our results are based on Android's current permission model, we believe the value of the crowd-sourced review model will extend to Android's recently announced new permission model as well.

Toward the goal of improving the ability of users to make security- and privacy-conscious decisions in updating mobile apps, we make the following contributions

- We perform the first study of how users make decisions in managing Android app updates, what factors impact users' decisions, and how these decisions could be better supported.
- Based on our finding that most users consider managing updates a non-trivial task and would benefit from additional information, we propose a new update notification mechanism that displays crowd-sourced reviews of the updated app and perform a further user study with this approach.
- Our field study demonstrates that integration of crowd-sourced app reviews helps users make more security- and privacy-conscious app update decisions, especially for less-trusted apps.

The remainder of this paper is organized as follows. We first discuss current app update notifications on Android and present the overview for our survey and field study. We then explain our methodologies, followed by an analysis of the results of our survey and field study. After the analysis, we compare our study with related work. Finally, we discuss the implications of our results and propose guidelines for designing app update notifications.

2. ANDROID APP UPDATE NOTIFICATION SCHEMES

Currently, app update notifications are not standardized. Most update notification interfaces use new permissions to demonstrate that the update will enable the app to collect more information or provide additional functionality.

Google Play provides both manual and automatic app update management for the Android OS. Users have three options: update all apps manually, update automatically anytime (i.e., regardless of type of connectivity), or update automatically only via WiFi. In addition, users can choose manual updates for specific apps of interest, which will override the general setting. If a user chooses to update an app manually, a dialog will show the permissions requested by the app update, as illustrated in Figure 1. The user is asked to decide whether to accept the listed permissions and install the update. If a user chooses to update an app automatically, a dialog will only appear if the updated app requests additional permissions that were not requested by the installed version of the app.

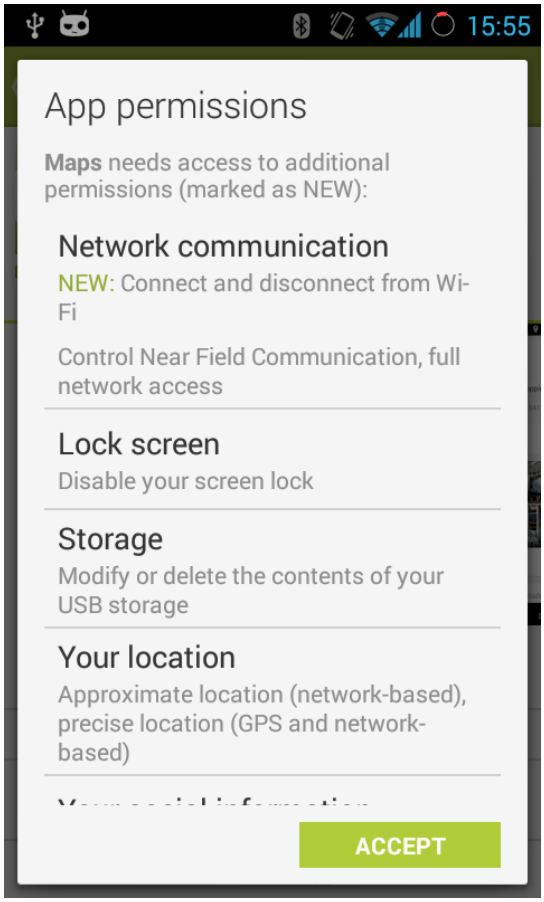


Figure 1: Google Play Update Notification, Version 4.1.1

When the user installs an app, a similar dialog lists all permissions requested by the app. The difference between the update dialog and the installation notification is that additional permissions will be labeled as "NEW" in the update dialog, as illustrated in Figure 1. In June 2015, Google announced a change for future Android OS versions to use in-context permission prompts instead of a one-time permission request at install time [13]. In the new model, the update notification only shows the new features, not the new permission requests. Our user studies were performed before the interface change, but we believe our ideas, experiments, and results apply in both settings for the following reasons. First, the new interface is still based on permissions. Thus, the limitations of permission-based notifications, such as users' lack of attention and understanding, are not mitigated. Second, our proposal of using review-based information can be combined with current in-context prompts to provide users with insights from experienced users.

3. RELATED WORK

Before going into the details of our user studies and proposed app update notification mechanism, we briefly describe related work in several relevant areas.

3.1 Android Permission and Privacy Information

Android platforms use permissions to organize and manage access to resources, such as location, contacts, SMS, and the camera. Apps need to declare permission requests in installation files. Some apps will update their permission requests with the installation of a new version. Previous work demonstrated that apps tend to get more permissions at installation than needed [27, 11]. Requests for unnecessary permissions can also occur during app updates, which is perhaps less transparent to users [17]. Wei et al. conducted a long-term study of the evolution and usage of Android permissions [27]. They found that popular apps tend to be over-privileged and request more permissions over time; additional permission requests are usually related to dangerous permissions. Felt et al. also studied over-privileged apps by analyzing their API calls [11].

Frank et al. studied common patterns of permission requirements for Android applications and found that permission requests are more diverse when compared to Facebook applications [12]. In addition, disreputable apps usually have different request patterns from apps which are more highly regarded. Thus, changes in permission requests might indicate behavioral changes in apps for Android platforms. Additionally, an attack on Android systems called the App Update Attack was studied by Tenenboim-Chekina et al. [19]. Because app updates are a potential way to implant new security vulnerabilities and privacy data leaks, the study of these updates is crucial.

Android permission systems have failed to deliver enough information or to make the information presented understandable or usable [22]. Chin et al. studied user confidence in Android's security and privacy. They found that users reported various concerns, some of which were due to misconceptions or misunderstandings [8]. Though this work examines only the initial app installation, it also suggests that we should search for ways other than traditional Android permissions to notify users about app updates' security and privacy implications.

Kelley et al. proposed a novel framework that introduced rich privacy information into the process of app installations [15, 16]. In their studies, privacy information could help participants to choose apps with less over-privileging, which is easier to understand than permissions. Instead of directly analyzing the permissions, we propose the inclusion of user reviews to provide additional privacy information.

Amini et al. build a crowd-sourcing platform to evaluate app's privacy and find that the crowd get as accurate result as the experts [3]. Their work is complementary to our proposal of using reviews for update notifications.

3.2 Update Behaviors

Update behavior for Android apps was first studied when Möller et al. analyzed the updates of one app on Google Play quantitatively [20]. They were the first to track how quickly users update one specific app developed by them. However, they did not track whether users set apps to update automatically or explain factors affecting users' decision making.

In our project, we study app updates from a user's perspective. We try to understand users' concerns about updating and use this understanding to design a scheme that assists in making privacy-conscious update decisions. Vaniea et al. also looked into factors which prevent users from updating apps on a Windows platform [26]. They discovered that changes in user interface and loss of functionality discourage app updating. Unlike Windows apps, which have full system privileges, Android apps can only access information if the corresponding permission is granted. Because of this, Windows users do not have to consider whether the app update might gain access to more information. Android users, however, have to be more aware of the impacts an update may have. We examine how Android users might be helped with the complicated update decision-making process.

3.3 Processing User Reviews

Hu and Liu did research on machine mining and summarizing customer reviews [14], while the essential problem of understanding the sentiment expressed in the reviews, was studied by Nasukawa and Yi [21]. These works are complementary to our paper, as these techniques could be helpful in designing effective approaches to choosing representative and trusted negative reviews. We could combine these techniques to find the most effective review automatically.

3.4 Using User Reviews and Community Ratings for Security Decisions

Besmer et al. did a between-subjects study for applying social navigation to access control policy configuration. Their work demonstrated that community reviews do impact user behavior if the visual presentation of the social navigation is strong enough [5]. Ayyavu et al. presented their work about integrating community ratings with heuristic analysis tools for web security [4]. They identify differences between heuristic analysis tools and methods based on community rating and resolve these conflicts. These previous papers presented the impact of user reviews for user security decisions.

Rader et al. found that some users visit forums and blogs to learn about how to make security decisions [24]. Similarly, our survey discovered that some users check reviews before deciding to update. We design our review-based notification on this finding and prove that not only do users care about the reviews, but they are able to use them to make more privacy-aware decisions.

3.5 Security Warnings

Akhawe et al. did a field study on the effectiveness of security warnings in browsers [1], Egelman et al. investigated whether people will click through fishing warnings [10], and Bravo-Lillo et al. designed and tested user interfaces for security decisions [7]. These researchers summarized the interface features which impact security decisions. Though their experiences are helpful for us in designing our user interface, we focus on the content of the notification. We apply our efforts here because notification content may trigger the users to react more securely.

4. STUDY OVERVIEW

Based on the Android app update models with options for users to update manually or automatically and building on the related work described above, we study Android app updates from a user's perspective by conducting two studies: a survey about update behaviors and attitudes and a field study to test different update interfaces.[1] In our first user study, we conducted an online survey to understand how Android users update their apps and what they are aware of or care about during these updates. We asked them about their experiences with app updates to analyze their behavior patterns of updates and the efficacy of current notifications. We also analyzed the survey results to find potentially better designs of update notifications that fit users' mental models. Based on the data from our survey, we propose a new design for app update notifications that alerts users when a potential privacy risk is identified or expected. Because we note that a few users report that they would check reviews explicitly before update to help them make the decision, we include reviews of the update inside the notification. We tested our new design by conducting an online between-group experiment for simulated updates of two popular apps. We discuss design details and analyze the two studies in the following sections.

5. SURVEY OF APP UPDATE BEHAVIORS AND ATTITUDES

Understanding users' concerns and reactions during app updates is important to help them to make the best decision. Toward this goal, we designed and executed a user study. In what follows, we describe our survey methodology and summarize our results.

5.1 Methodology

Our online survey aims to understand how users update their Android apps and investigate what factors impact their decisions.

5.1.1 Design of Survey Questions

In the survey, we asked the participants about their app update behavior. The 11-question survey asked about what information people expect in update notifications as well as the efficacy of Google Play's current notifications.

- We asked users how they usually update apps and why they choose to update this way. Participants were provided with options such as, "Apps on my phone are updated automatically," "Apps on my phone are updated manually," and "I use different strategies in different cases." These questions provided us with insights about users' update behaviors and how users are involved in the app updates.
- We asked participants to report their experiences with updating, particularly in what conditions they would choose not to update or they would regret updating. [Sample questions: "Have you ever chosen not to update an app? (If you dismiss the update notification screen without clicking 'Accept,' your app will not be updated)"; "Please recall some app(s), which you chose

[1]The CMU IRB approved our human subjects experiments.

not to update. (You may give the descriptions or categories of the apps if you do not want to disclose their names)"; "Please tell us about the app and why you chose not to update."] From these detailed user experiences, we are able to assess whether current update notifications are helpful and to assist with productive decision-making.

- We also asked what factors affect the decision to update an app. We provided seven factors (source/author of the app; popularity of the app; new features the app has; trustworthiness of the app; my usage of this app; new permission requests; why the app requested these new permissions) and five agreement degrees. These seven factors were chosen based on previous research on app installation as well as pilot tests. From these results, we gain insight into designing update schemes focused on important factors.
- In addition to these survey questions about update decisions, we gathered demographic information about age, gender, education, and Android experience.

5.1.2 Survey Deployment

In order to deploy our survey, we recruited participants from Amazon Mechanical Turk (MTurk) who satisfied the criteria below:

1. Located in the U.S.
2. MTurk HIT approval rate of 95% or greater
3. 18 years old or above
4. Literate in English
5. Minimum one month Android use with Google Play installation experience

We emphasized the five criteria in the first page of the survey and double-checked in survey questions to properly filter participants. For example, we asked about their personal information again to enforce the first three criteria automatically, and we asked users some open questions about their experiences with installing apps on Android to verify the last two criteria. Users who matched all of our criteira were invited to complete our survey. We paid $0.20 to each survey participant, which took an average of 10 minutes to complete, and obtained 300 valid responses.

5.1.3 Survey Data Analysis

We asked both multiple-choice and open questions in the survey, and we analyzed these two categories of questions differently.

After we collected the results from MTurk, we first analyzed the answers of the multiple-choice questions such as whether they update automatically. We computed the percentages for each options to get insight about users' update behaviors and attitudes toward app update.

Then we coded the results of the open questions such as their experiences about updating to categorize different responses. The coding process was as follows. First, two researchers checked the answers separately and then produced their own code book that categorized responses into different categories. Then, they compared their code books, discussing the differences to decide the codes in the code book. For the two open questions "why refuse to update an app" and "why update apps manually," we developed code books with nine and seven codes respectively. After reaching agreement on the code book, the researchers independently performed two tasks: (a) summarizing of categories of participants' responses and (b) assignment of participants' responses to proposed categories. Next, where the coding agreement was low, another independent round of coding was performed. Finally, after the coding agreement became satisfactory, we analyzed the distributions of the different categories and summarized our findings.

5.1.4 Limitations

We iterated over several rounds to design the survey questions. However, users' data is still self-reported, which is necessarily subjective. To somewhat mitigate this subjectivity, we also performed a follow-up field study to get more objective results. For example, users didn't report "trust" as a major factor when they answered the survey questions, but participants in the field study demonstrated that "trust" is very important for them when making an update decision. The details are discussed in the field study section.

5.2 Survey Results

Among the 300 participants from MTurk, 69.3% of them were male, 30.7% were female, 44.3% had a Bachelor degree or above, and the average age of the participants was 28.3 ($\sigma = 7.8$).

We found that many users are very involved in the app update process. According to their responses, 47.7% of survey participants update their apps automatically, 25.0% update their apps manually, and 25.0% update some apps manually and some apps automatically. 59.3% of the participants have ever chosen *not* to update an app, and 42.6% have regretted updating an app.

We analyzed the reported reasons for choosing an update strategy (agreement rate = 93.5%, Cohen's kappa = 0.920 [9])[2] and found leading reasons for manually updating all or some apps include the following.

- Feel more control ("I like to have control over what happens with my phone.")
- Only update certain apps ("This way I can pick the apps I want to update.")
- Know more about the update ("So I can read about any changes in the app or what has been changed.")
- Address privacy concern ("I don't want to allow apps to take any information they want that I am not comfortable with.")

[2]Agreement rate is computed by comparing the percentage of two researchers put one object into the same category or the same sets of categories. Cohen's kappa is a more robust static to measure the inter-rater agreement since it considers the agreement occurring by chance [28].

Table 1: Reasons for manual app updates reported in the survey.

Reason	# Participants
Feel more control	44
Only update certain apps	38
Phone limitations	34
Want to know more about the update	29
Privacy concern	18
I don't know & others	15
I want to read the reviews	2

- Avoid phone limitation ("I don't always want to bog down how fast my device runs by having apps update automatically.")

A detailed account of these responses is shown in Table 1. Our results demonstrate that users like to engage in decision making for app updating by being aware of updates and making decisions they think are reasonable. Most notably, many users stated that they prefer to wait to see reviews to know whether the update is worthwhile.

We found that the reported reasons for concerns about updating are very diverse. In order to understand the factors that were considered when making decisions about updating an app, we proposed a list of 7 factors that are potentially related. For each factor, participants were asked to provide a 5-point Likert scale rating [18] reflecting their concern about this factor when making the corresponding decision. Participants provided a broad range of ratings. Generally, the variation of their decisions is relatively high: 0.88 to 1.14 on a scale of -2 to 2. Users had quite diverse preferences on these factors. Anticipating this phenomenon, we also provided open questions for participants to recall their experiences regarding not updating an app.

59.0% of participants reported that they had chosen not to update an app, and 42.7% regretted some app updates. A follow-up open question addressing participants' reasons for not updating or regretting updating was provided. We also applied the two-researcher coding procedure to interpret and understand the responses to these other questions. The agreement of the coding results between the two researchers was already 88.7% (Cohen's kappa = 0.866). According to the coding results for reasons of not updating apps, privacy and permission-related concerns were among the most frequently mentioned (see Table 2), while concerns about changes in experience or functionality were related to the quality and service of the apps themselves. This implies that participants do care about privacy when considering whether or not to update an app. Interestingly, a considerable number of responses mentioned considering negative reviews from online or offline sources. This was consistent with our hypothesis that user reviews can be a substantial help in decision making.

When we analyzed reported reasons for regretting updating apps, we found that the most common reasons are worse functionality and user interface. Participants often complained about updates introducing bugs and bad user interfaces (see Table 3). They said things like, "My bank's app stopped working when I updated it. I couldn't use the mobile deposit for a couple months until they fixed it," "Perk App - they took away some of their best features," "I didn't like when they changed the Twitter interface," and "I regretted having updated my Facebook app because I had grown accustomed to the older version and having to get used to the new interface was inconvenient."

Table 2: Top reported reasons of not updating an app in the survey.

Reason	# Participants
Functionality / Experience Changes	30
Privacy Invasiveness / Information Collection	30
Not Using the App Any More or Often	26
Unnecessary Permission Requests	18
Bad Reviews	14
Cost (Data, Payment, Time, Space)	14
Disfavoring the App	7
No Reason / Necessity to Update	6
Annoying Update Notices	3

Table 3: Reasons for regret updating an app in the survey. Over the 128 users who reported that they regretted updating an app, 6 users did not provide a reason.

Reason	# Participants
Update causes bugs, and worse functions	82
Uncomfortable user interfaces	17
Bad Reviews	14
Cost (Data, Payment, Time, Space)	14
New version accesses more information	11
Disfavoring the App	7
Like the old version better	6
Update took up too much space / was too slow	6

In comparison, seldom do users regret updating apps because of security and privacy. Only 11 participants reported that they realized that the apps were accessing more information or felt regret about updating the apps. As one participant reported, "some apps choose to change the information that they wish to access, sometimes this happens with apps that i *[sic]* have liked but now have no further use for, the perfect example would be Facebook and the way in which they want access to everything, this is no good and thus the app is no longer needed." This result suggests that most users are not be able to observe privacy problems with a new update before they install the update. Therefore, update notifications are the best opportunity for users to realize that the privacy feature of the update is not what they want.

When analyzing survey results, we found that although many users reported that they were concerned about privacy they may not be able to make informed choices. A few users even explicitly mentioned that they had to read reviews to help them understand the update.

6. FIELD STUDY OF APP UPDATE INTERFACES

Our survey results suggest that reviews may help people decide whether or not to accept an app update. We would like to compare whether reviews, instead of automatically extracted permission requests, can be a better way to help

Table 4: Conditions for the decision making study.

CCS, Perms	Candy Crush Saga, permission requests
CCS, Reviews	Candy Crush Saga, negative reviews
Maps, Perms	Google Maps, permission requests
Maps, Reviews	Google Maps, negative reviews

users make informed choices. We designed and implemented an Android app that simulates two different designs of update notification for Google Maps and Candy Crush in order to compare the effectiveness of them using a crowd-powered method.

6.1 Methodology

We carried out a field study in which participants were shown different designs of update notifications. We collected participants' behaviors, including the amount of time they spent on each notification screen and their decision about whether or not to update the app.

6.1.1 Design of field study

We ran a screening survey on MTurk in which we selected only users of Android devices that ran Google Play. We also required that the participants should have installed Candy Crush Saga and Google Maps on their phones. Because our app will pop up update notifications for Candy Crush Saga or Google Maps. The reasons for choosing Candy Crush and Google Maps are two-fold: they are very popular apps and many users have both of them on their phones; they are at different trust levels so that we can test whether users' trust to apps can affect the update decisions. If a user matched our criteria, we sent out a link to our app and asked the participant to install it on their devices.

The participants were told that we were doing a study on the power usage of Candy Crush Saga when they downloaded the app. The app showed the current battery level, which made it looked like an authentic battery usage monitoring app. In fact, the app was used to show update notifications. The reasons why we pitched our app as a battery management app are that we don't want the users to change their update behaviors for the experiment and we needed the users to keep the app running on their phones. After about 12 hours, the app started to push simulated notifications of app updates to the notification area (Figure 2(a)), which looked like the notifications from Google Play. In this way we were able to simulate app update notifications on real devices and collect real response to the design.

After the user clicked on the notification, a dialog containing simulated update messages appeared. We designed this experiment as a between-group study with 4 conditions (Table 4) which is the combination of 2 different apps (Google Maps, Candy Crush Saga) and 2 different messages (permission requests, negative reviews). Depending on the condition a participant was randomly assigned to, they saw different messages. For example, a participant in the "CCS, Reviews" condition saw a negative review of the update for Candy Crush Saga (Figure 2(b)), while a participant in the "Maps, Perms" condition saw messages about Google Maps requesting a new permission (Figure 2(c)). Note that the permission-based notification and the review-based notification were describing the same level of privacy invasion. For example, permission-based notification displayed that the update asked for "camera" permission and the review-based notification explained that the update can take pictures.

In the dialog they had 3 choices. They could either accept the update ("Yes"), decline the update ("No"), or choose to make a decision later ("Not now"). If they clicked on "Not now," in a few hours the notification would appear again in the notification area. If they chose either "Yes" or "No," they reached their last step in the study, in which we first showed them the debrief message about the update message being simulated, and then ask them why they made such a decision. Their final choice ("Yes" or "No"), with their explanation of the choice, was then encrypted and reported to our server. We also collected the number of times they clicked on "Not now," and the duration of their time staying on the dialog to understand how long they spent on reading the message and making the decision.

The "Not now" choice was only available in the first 2 appearances of the dialog. In the 3rd, the participants could only choose between "Yes" and "No."

6.1.2 Data Analysis of the Field Study

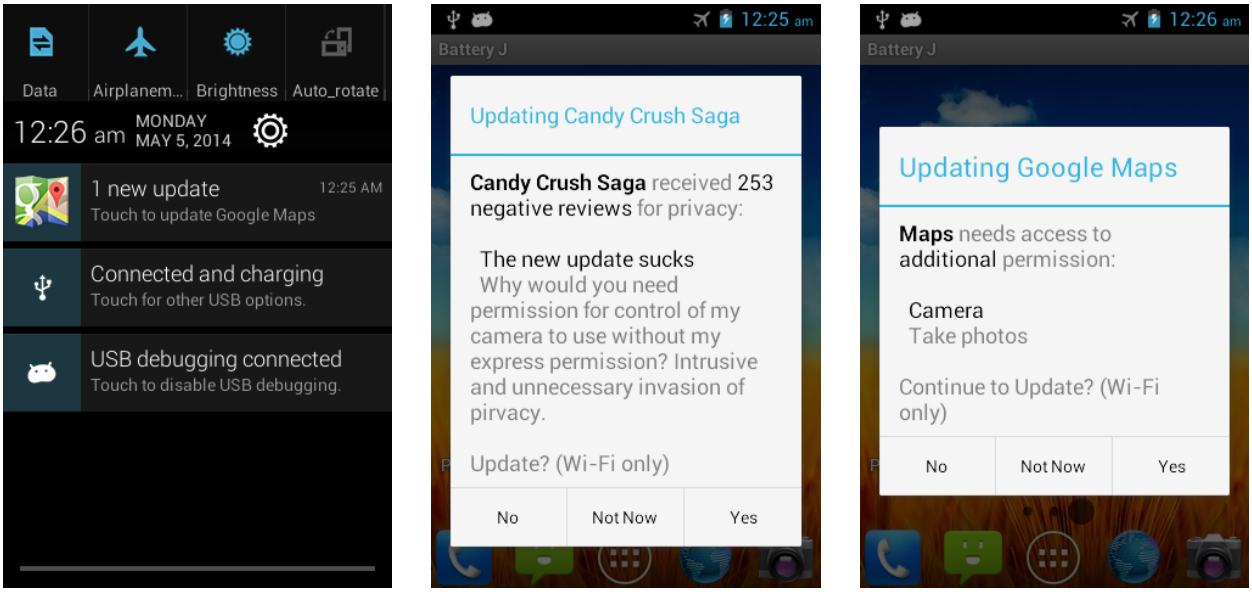
We correlated the decision results with different update notifications to check which notification triggered greater user reaction. We also collected and coded participants' decision-making reasons and recorded behavioral data such as how long users stayed on the notification.

We used statistical analysis to assess behavioral differences between notification conditions. First, we ran a logistic regression to build the model for the impact of review-based or permission-based notifications. We use the app name (Google Maps or Candy Crush Saga) and notification format (review or permission) as independent variables and the update decision (update or not update) as the dependent variable for the logistic regression. Then, we ran a t-test about stay time with "Yes" and "No" conditions for the update decision. We confirmed the relationship of longer time on notification screen and privacy-aware decisions. Another t-test was run on Candy Crush Saga looking at time differences between negative reviews and new permissions. The result further reinforced our analysis regarding time differences.

6.1.3 Limitations

We made the assumption that the reviews we chose were trusted and expressed their privacy concerns clearly. Though the notification might be misleading if the review is not correct, this type of assessment was outside the scope of this project. Finding representative negative reviews is another interesting problem. Natural language processing and crowd sourcing could be helpful in identifying review meaning and reputation when choosing representative reviews in the update notification.

A pervasive limitation of review-based schemes, ours included, is their inability to function without prior production of reviews. Researchers could work on automatically evaluating updates and producing initial reviews for the users.



(a) Android notification item (1 new update)

(b) Condition 2: CCS, Reviews

(c) Condition 3: Maps, Perms

Figure 2: The simulated notification and message dialogs

The update notifications we displayed were not coming from the operating system, but we went to great lengths to copy the visual design of the standard Android interface. In fact, only one participant mentioned at any point in the study that our notification was not from Google Play.

Users might have fatigue if they see the update notifications frequently. This problem could be mitigated by adopting attractors such as animation and swipe that are proposed by Bravo-Lillo et al. [6]

6.2 Results for the Update Decision Making Study

From March to May of 2014, we recruited participants on MTurk to install our app that simulates update notifications and collected update decisions from them. We got 736 valid responses from the MTurk screening survey, 411 of which opted to install our app and join the study. 194 participants actually download our app; 96 finished all the steps and provided us reasons of their update decisions. We had between 22 and 25 participants in each condition, and each of the 4 conditions had similar demography distributions.

Of the 96 participants who completed the study, the average age was 29.6 ($\sigma = 6.9$), with 36 females and 60 males. Education and occupation were very diverse, ranging from unemployed to engineers, high school diplomas to graduate degrees.

After analyzing the results, we had the following observations:

- Negative reviews were better at informing users about privacy violations in app updates.

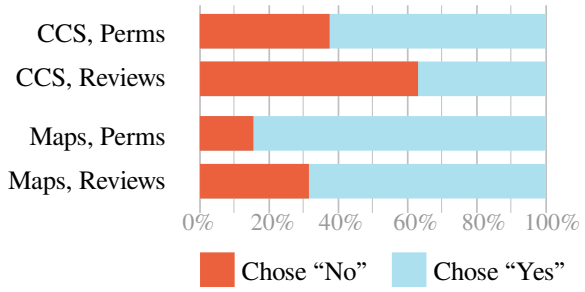


Figure 3: Results from the decision-making field study. “Not now” is not an option for the final decision so it’s not presented here. Participants in the “Reviews” conditions tended to refuse updates more.

- The trustworthiness of an app was very important to users when making update decisions, which was different from self-reporting about decision factors in the online survey.
- People who spent more time reading update notifications generally made better decisions from a privacy standpoint.
- Participants spent more time when prompted with our negative review-base notification than with the existing notification that only indicates the new permissions to be included.

In the following sections, we discuss and analyze results for the between-subject study of app updates on Android phones. We first examine how decisions are affected by different update notifications and the reasons for users' responses. Then we talk about our analysis of update behaviors, such as time and decision patterning in the update decision process.

6.2.1 Update Decisions: Results and Reasons

Different update decisions under different notifications show that users react better to reviews than permissions. Users who saw negative reviews (47.3%) tended to refuse app updates more than those who saw the new permission update notification (26.5%), (logistic regression, $p < 0.03$) used by Google Play and other marketplaces. Also, users generally updated Google Maps (77.7%) more than Candy Crush Saga (49.0%),(logistic regression, $p < 0.01$) because of their trust of Google Maps (indicated in reported reasons).

As illustrated in Figure 3, 63% of people refused to update Candy Crush Saga when they were shown negative reviews, while only 37.5% of the users who were notified by new permissions refused the same update. Similarly, though users tended to update Google Maps more, the percentage that rejected updates when shown negative reviews (31.6%) was still around twice the percentage who rejected updates when shown the new permissions (15.4%).

Our analysis of users' update reasoning also supported the conclusion that people were more alarmed by negative reviews about privacy than by new permissions to access sensitive data.

We coded 96 participant responses regarding update decisions, finding that the leading reason to *not* update was having read negative reviews for privacy about the update (19 times), while other factors such as sensitive permissions (9 times) and phone limits (8 times) were referenced much less frequently.

For details about user responses when choosing not to update, we list a few representative quotations from them about the usefulness of the negative privacy reviews. Seven users reported that they did not want to try things that other users had negative thoughts about. For example, p_1 said "The pop up showed a negative review where someone said the app update asked for permission for use their camera. I didnt want that either, so I said no" and p_2 mentioned "Too many negative reviews; if that many people had negative thoughts and feelings about it, its probably not something I want to try." Twelve users claimed that reviews helped them to identify the privacy problem with the update. For example, p_3 reported that "The negative comment about it wanting to access my camera for it to update was suspicious." p_4 told us that "lots of bad reviews. the review you showed says it needs to use my camera I don't see any need for that." Finally, p_5 claimed that "Privacy is very important to me and being able to enter to my personal data I will not except. the negative reviews helped take me the decision."

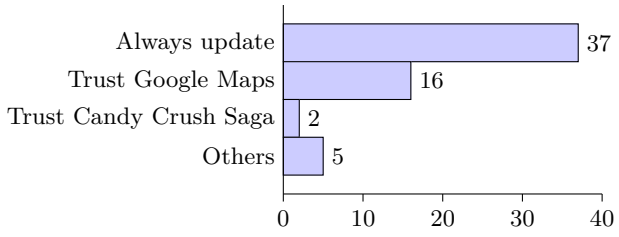


Figure 4: Reasons for updating the app in the decision making study

On the other hand, the major reasons people chose to update were their trust of the application or their desire to keep things up to date (reported by 37 participants) (Figure 4). The data also indicates that users' greater willingness to update Google Maps over Candy Crush Saga is primarily based on the trust of the application. Although trust was not reported as a major factor for update decisions in the online survey, it played a very important role in update decisions. For example, 16 participants reported their trust in Google Maps and 2 reported their trust in Candy Crush Saga when they decided to update the app.

Typically, some users always keep their apps up to date. For example, p_6 reported "I chose yes because I like to keep all of my apps updated." This phenomenon suggests that many users are not conscientious about the potential privacy risks in app update.

There are also many users who update apps because they trust the app. For example, p_7 said that "I trust Google products, especially maps, plus they run my Android software," and p_8 mentioned "updating is what I would expect to do. I said because I trusted Google maps."

Interestingly, some users have misunderstandings about permissions. For example, p_9 was confused that the update actually required more permissions: "It looked like they were removing camera permissions. So I would like that."

6.2.2 Update Behavior Patterns

We also collected users' behavior patterns during the update process to better understand their attitudes and behaviors. These reported behaviors also demonstrate that participants pay more attention to review-based notification than permission-based notification.

We recorded the time they stayed on the notification dialog and noted the times they select "not now" before making

a final decision. However, we revised the app and ran the experiment iteratively and only included the record of time users stayed on the notification dialog in later versions of our app. After removing the outliers, we had 27 valid user records.

Using staying time as a representation of people's attention, we could record the time people spent reading update notifications and compare different levels of attention in different cases. Specifically, the app recorded the elapsed time between the appearance of the dialog box and the choice of "Yes," "No," or "Not now."

After analyzing behavior patterns according to different conditions and decisions, we found some interesting correlations in the relationships between attention, notifications, and decisions. Firstly, when people made the decision not to update the app, they tended to stay longer at the notification interface (mean = 25.2 s, median = 13.0 s), compared to when they decided to update (mean = 6.5 s, median = 5.0 s). Also, participants who saw negative reviews spent more time reviewing the notice than participants who saw the permissions screen (t-test, $p = 0.0286$). Participants who saw Candy Crush Saga reviews spent a median of 13 s (mean = 20 s, σ = 19.50 s), where the permissions screen garnered a median of only 5 s (mean = 5.90 s, $\sigma = 3.69$ s).

Using Candy Crush Saga, negative review notifications produced a significantly longer read time than the permissions notification. However, the time difference between the review notification and permissions screen in Google Maps was much less pronounced.

7. IMPLICATIONS AND DESIGN RECOMMENDATIONS

Our study demonstrates that though many users do care about privacy, they fail to make informed update decisions with traditional permission-based notifications. This may be because many users misunderstand or ignore permission-based notifications. One user, for example, told us that he thought the camera permissions were being removed when he read the notification in its traditional format. Such misunderstandings also imply that app updating is a complicated process involving both the addition and removal of various features. Users often struggle to decipher, and therefore tend to ignore, traditional notifications. When most users viewed traditional notifications, they either did not notice permissions or they misunderstood the permission model. We believe that Google's recently proposed in-context permission system does not resolve these complications for two primary reasons. First, it is up to developers to present permission requests at the best time, and, second, many users will blindly approve permission requests with overly technical descriptions. We further believe that our proposal of including crowd-sourced reviews into the update notification still provides useful information under this new permission model.

Many users reported that they always updated all their apps. This might not always be a good practice for security and privacy considering the potential invasiveness of malicious app updates. This phenomenon implies that the scheme used for updating should not be the same as the scheme used for installation because users appear to pay less attention during updates. The new interface design should pressure users to pay more attention to updates by displaying reviews or, at the very least, by adding one more click before the update.

We also notice that many users update apps because of their trust in the apps. This finding could guide the future design of update notifications in two ways. First, the notification screen must clearly and accurately indicate the identity of the app developer. Second, the notification screen must warn the user that even apps you trust could also cause privacy problems when they ask for more permissions.

Our experiences studying update behaviors demonstrate the power of reviews for app update management. We learned that assisting users with the complicated task of app updates requires a notification that is clear and easy to understand. Currently, we only tested the effect of negative reviews for privacy, and the initial results were very promising. In the future, researchers could experiments with mixed reviews (positive and negative) or extend to reviews about other features such functionality and interfaces.

8. CONCLUSION

Through our online survey of users' attitudes towards app updates and field study of users' decisions about app updates, we acquired a better understanding of their attitudes and behaviors. The results show that privacy invasion is one of the most frequently cited reasons for not updating an app. When compared to the traditional permission-based notifications, our proposed review-based notifications were significantly more effective at catching the attention of users, leading them to make more privacy-protective decisions. These results provide insight for the future design of app update notifications by including peer reviews to help users identify potentially invasive app updates.

9. ACKNOWLEDGEMENT

We would like to thank Jaeyeon Jung (Microsoft Research), Lujo Bauer (Carnegie Mellon University), and Cici Zhao (Duke University; now at Amazon.com) for their suggestions and help.

10. REFERENCES

[1] D. Akhawe and A. P. Felt. Alice in warningland: A large-scale field study of browser security warning effectiveness. In *Proc. USENIX Security*, 2013.

[2] R. Amadeo. Adware vendors buy Chrome Extensions to send ad- and malware-filled updates. *Ars Technica*, January 2014.

[3] S. Amini, J. Lin, J. I. Hong, J. Lindqvist, and J. Zhang. Mobile application evaluation using automation and crowdsourcing. 2013.

[4] P. Ayyavu and C. Jensen. Integrating user feedback with heuristic security and privacy management systems. In *Proc. CHI*, 2011.

[5] A. Besmer, J. Watson, and H. R. Lipford. The impact of social navigation on privacy policy configuration. In *Proc. SOUPS*, 2010.

[6] C. Bravo-Lillo, L. Cranor, S. Komanduri, S. Schechter, and M. Sleeper. Harder to ignore? Revisiting pop-up

fatigue and approaches to prevent it. In *Proc. SOUPS*, 2014.

[7] C. Bravo-Lillo, S. Komanduri, L. F. Cranor, R. W. Reeder, M. Sleeper, J. Downs, and S. Schechter. Your attention please: designing security-decision UIs to make genuine risks harder to ignore. In *Proc. SOUPS*, 2013.

[8] E. Chin, A. P. Felt, V. Sekar, and D. Wagner. Measuring user confidence in smartphone security and privacy. In *Proc. SOUPS*, 2012.

[9] J. Cohen. Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. *Psychological bulletin*, 70(4):213, 1968.

[10] S. Egelman, L. F. Cranor, and J. Hong. You've been warned: an empirical study of the effectiveness of web browser phishing warnings. In *Proc. CHI*, 2008.

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

[12] M. Frank, B. Dong, A. Felt, and D. Song. Mining permission request patterns from Android and Facebook applications. In *Proc. ICDM*, 2012.

[13] Google. Android M Developer Preview & Tools, 2014. `http://android-developers.blogspot.com/2015/05/android-m-developer-preview-tools.html`.

[14] M. Hu and B. Liu. Mining and summarizing customer reviews. In *Proc. KDD*, 2004.

[15] P. G. Kelley, S. Consolvo, L. F. Cranor, J. Jung, N. Sadeh, and D. Wetherall. A conundrum of permissions: Installing applications on an Android smartphone. In *Proc. FC*, 2012.

[16] P. G. Kelley, L. F. Cranor, and N. Sadeh. Privacy as part of the app decision-making process. In *Proc. CHI*, 2013.

[17] S. M. Kywe, C. Landis, Y. Pei, J. Satterfield, Y. Tian, and P. Tague. Privatedroid: Private browsing mode for android. In *IEEE International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*, Sep 2014.

[18] R. Likert. A technique for the measurement of attitudes. *Archives of psychology*, 1932.

[19] L.Tenenboim-Chekina, O. Barad, A. Shabtai, D. Mimran, L.Rokach, B. Shapira, and Y. Elovici. Detecting application update attack on mobile devices through network features. In *Proc. INFOCOM*, 2013.

[20] A. Möller, F. Michahelles, S. Diewald, L. Roalter, and M. Kranz. Update behavior in app markets and security implications: A case study in Google Play. In *Proc. Workshop on Research in the Large*, 2012.

[21] T. Nasukawa and J. Yi. Sentiment analysis: Capturing favorability using natural language processing. In *Proc. K-CAP*, 2003.

[22] L. T. Nguyen, Y. Tian, S. Cho, W. Kwak, S. Parab, Y. S. Kim, P. Tague, and J. Zhang. Unlocin: Unauthorized location inference on smartphones without being caught. In *International Conference on Security and Privacy in Mobile Information and Communication Systems (PRISMS)*, Jun 2013.

[23] W. Pelegrin. These Android, iOS, and WP8 apps are affected by the Heartbleed Bug. `http://www.digitaltrends.com/mobile/heartbleed-bug-apps-affected-list/`, 2014.

[24] E. Rader, R. Wash, and B. Brooks. Stories as informal lessons about security. In *Proc. SOUPS*, 2012.

[25] Thomas Ella. Why does the android app now want permission to read my texts?, 2014. `http://on.fb.me/1CdGVzQ`.

[26] K. E. Vaniea, E. Rader, and R. Wash. Betrayed by updates: how negative experiences affect future security. In *Proc. CHI*, 2014.

[27] X. Wei, L. Gomez, I. Neamtiu, and M. Faloutsos. Permission evolution in the Android ecosystem. In *Proc. ACSAC*, 2012.

[28] J. M. Wood. Understanding and computing cohen's kappa: A tutorial. *WebPsychEmpiricist. Web Journal at http://wpe. info/*, 2007.

[29] Y. Zhou and X. Jiang. Dissecting Android malware: Characterization and evolution. In *Proc. IEEE S&P*, 2012.