Introducing Privacy Threats from Ad Libraries to Android Users Through Privacy Granules

Anand Paturi
New Mexico Institute of Mining and Technology
anand@cs.nmt.edu

Patrick Gage Kelley
University of New Mexico
pgk@unm.edu

Subhasish Mazumdar
New Mexico Institute of Mining and Technology
mazumdar@cs.nmt.edu

***Abstract*—Android mobile users are provided with a permissions list before installing an app that displays the list of resources available to that app. Users can review the permissions list and decide to install the app if they trust the app with their information. However, this information is accessible not only to the app provider but may also be available to third party ad libraries included in the app, which users are unaware of. In this paper, we propose a novel icon-based privacy threat representation as an alternative to permissions list that shows privacy threats to users from both app providers and associated ad libraries. Our approach considers users' privacy in terms of three granules: location, identity and query. Our proposed interface aims to educate users about which particular app providers and third parties have access to their privacy granules. We obtained user feedback on our technique in two user surveys (n = 137; 294), one each for testing the icons and the icon-based privacy threat display. We present our findings for ease of use and effectiveness of the novel privacy threat interface and further evaluate its impact on users' installation decision.**

I. Introduction

Android Operating System (OS)-based devices hold a majority share of the smartphone market–81% as of Jan. 2014 [14] along with its official app store – Google Play store hosting over 1.3 million apps. Majority of these apps are developed and uploaded by independent developers, which users download and install to avail different services like location-based services, online games, social networking, entertainment, and many others. In order to provide required services, apps access critical resources on end users' devices like their location, contacts list, phone number, search preferences, Personal Identifiable Information (PII), and more.

The vast majority of Android app providers release their apps for free (around 83% of total available apps [2]), often depending on ads to generate revenue. App providers include different third party ad libraries in their apps that communicate with ad servers to display ads on an end user's device (*via* the host app). In order to provide more customized and targeted ads, ad libraries gather accessible information (like user location, contacts list etc.) from the user's device and send it to their servers along with the device id. Authors in [24] analyzed 13 most popular Android ad libraries and concluded that: (a) Different ad libraries have different practices in place when accessing data from users' devices i.e., some ad libraries might access data that is not disclosed in their documentation while some may not. (b) Third party ad libraries can keep track of users' activity over multiple apps (in which that particular ad library is included) with users' device ids acting as unique identifiers for such data. Although, ad libraries are capable of accessing critical information from the user's device, those details are not revealed to the user during install time through the app's permissions list, since permissions lists only display information accessible to the app provider. Therefore, it would be helpful to provide an interface to end users that communicates the information accessible to both app providers and third party ad libraries, so they can make an informed decision during app installation.

Through our work introduced in this paper, we replace the current permissions list with an icon-based privacy threat representation (also referred as proposed interface). We consider user privacy in terms of three granules: the user's location, information that might reveal his identity (like his device IMEI (International Mobile Station Equipment Identity), email address etc.) and his search queries issued through the app. We derive the threat representation from information accessible to app provider and ad libraries included in that app (also referred as third parties) that might reveal any of above mentioned privacy granules.

We performed two online usability studies to obtain user feedback on our icon-based privacy threat interface. The first usability study was to select three icons, one each for the location, identity, and query granules. The second usability study was to evaluate our novel icon-based privacy threat interface that we created using the icons selected in the first user study. Users were shown the current permissions list interface and our icon-based privacy threat interface for an app and their responses were evaluated to determine ease of use in finding required information in both interfaces, effectiveness of the proposed interface, and its impact on app installation decision making. Our findings indicate that users are open to novel representations as alternatives to current text-based permissions list. Since, our privacy threat interface showed the privacy threat from both app provider and ad libraries included in it, users felt that showing such a privacy threat interface before installation of the app would influence their

USEC '15, 8 February 2015, San Diego, CA, USA

http://dx.doi.org/10.14722/usec.2015.23008

app installation decision.

II. Related Work

Privacy concerns in Android apps have been addressed in two facets: (a) Performing automated application analysis to detect misuse of user's private data and (b) Educating users about privacy threats before application installation. Researchers [8], [3] have proposed methods based on permissions analysis and detected excessive permission usage by Android apps, more than required by their functionality. TaintDroid detects information leakages, in real time by performing dynamic analysis of Android apps [7]. In [10], the authors developed a tool: *ScanDroid*, which analyzes Android apps' source code, along with the *manifest* file included with each app, to produce a data-flow policy specification describing an app's use of information. Some additional security tools have also been developed to protect Android users' privacy [27], [4], [21], [13], where in [21] each user is given the flexibility to choose a more granular level of privacy protections. This mechanism would allow users to grant permissions to applications based on their comfort level and underlying usage context.

Work specific to privacy and security risks posed by third party ad libraries in Android apps has been done by authors in [24], [11] and [5]. In [24], Stevens et. al. performed analysis on 13 Android libraries and found that some ad libraries checked for permissions beyond the required and optional ones listed in their documentation, including sensitive permissions like CAMERA, WRITE CALENDAR, and WRITE CONTACTS. Grace et. al. have created a static analysis tool: *AdRisk* to systematically identify security and privacy risks posed by ad libraries [11]. Authors present their results for 100 ad libraries in terms of data and permissions accessible to each of them. In [5], authors create a tool *Brahmastra* to detect security risks involved in third party ad libraries' code included in Android apps through UI path execution analysis. Authors findings mention that 175 out 220 children's apps display ads that attempt to collect personal information, which is a potential violation of the Children's Online Privacy Protection Act (COPPA). In [6], authors have performed longitudinal analysis of Android ad library behavior by investigating their behavior changes over time by taking a sample of 100,000 apps. They reveal that over last several years more libraries are able to use permissions that pose particular risks to user privacy and security and such behavior of ad libraries is steadily increasing.

Kelley et. al. [18] conducted a user study and identified that users do not pay attention to permissions lists before installing Android applications and are not aware of the security risks posed by apps accessing these permissions. Felt et. al. [9] also conducted studies to understand users' perception regarding permissions list and found the same. In order to enhance the user experience while reading permissions lists and allow them to make an informed decision, Kelley et. al. [19] proposed to display privacy information to users before they decide to install the app and assist them in choosing applications with less permissions. Lin et. al. in [20] designed a variant interface that replaced the regular permissions list. Their variant had opinions of previous users of the app, regarding different permissions the app requested. Authors obtained this opinion through an online survey which evaluated the mental models of Android users regarding the expectation and purpose of permissions in Android apps.

In its latest update, Google Play Store has enhanced its permissions list by adding icons along with grouping their permissions to enhance user experience. In the latest permissions list, all available 145 permissions are classified in to different permission groups, where each group is represented by an icon. For example, all location related permissions are listed under the 'Location' group and represented using a location icon. Permissions that cannot be classified in to any group, are part of the 'other' group. However, this new interface does not represent any privacy threat details, communicate to user how the collected information is shared with third parties and further new permissions accessed as part of app updates are not revealed to the end user.

Our work in this paper is first of its kind that extracts privacy threats (posed from both app provider and associated third parties) by performing static and dynamic analysis of an app and further presents the threat information to users in a usable icon-based threat representation interface. We envision our proposed interface would assist Android users in making an informed app installation decision.

III. Background

A. Ad libraries in Android

Ad providers issue SDKs (ad libraries) that developers include in their apps for retrieving ads from the ad provider's servers. The SDK provides APIs that abstract the methodology involved in requesting, parsing responses, and displaying ads in the app. To display an ad, the ad library makes an ad request to an ad server, which responds with the ad shown to the end user. While communicating with the ad server, an ad library may send the user's device's unique id along with other data, such that customized ads can be served to the user. For example, a user's location data could help the ad server send ads pertaining to that location, or even nearby businesses.

Ad libraries execute in the same process space as the apps in which they are included, giving them same privileges as the host app on the device. In addition, ad libraries also have a mandatory set of permissions needed to execute (usually permissions for communicating over network) that must be added to the *manifest.xml* file. In this way, ad libraries will have access to the resources required for their functioning and also any resources accessible to the app. When a user accepts the permissions list displayed to him before app installation, he is directly granting the app provider with those permissions and indirectly granting those same permissions to ad libraries included in that app. At this point, how ad libraries utilize permissions accessible to them for accessing user data and related privacy threats completely depends on their SDK implementation. As mentioned by authors in [24], different SDKs use permissions accessible to them in different ways. For example, *Mobclix* ad library utilized seven undocumented permissions i.e., permissions not specified as either mandatory or optional in its SDK documentation. These included four invasive permissions: READ CALENDAR, WRITE CALENDAR, READ CONTACTS, and WRITE CONTACTS.

B. Granular Privacy for Mobile Users

In this work, we consider privacy requirements of Android users in terms of three granules: Location (L), Identity (I) and Query (Q). L refers to user's geographical location details, I refers to details that would uniquely identify the user; for example, his email address, device IMEI number, social accounts etc. Q refers to a user's search queries or items of interest he looks up using a mobile app, like nearest restaurants, gas stations, bars, etc. Revealing or sharing any of the above granules or data related to them to untrusted parties is considered a privacy threat to that granule. For example, if an app includes an ad library that has access to user's location, then there is a privacy threat to user's L from both the app provider and associated ad library.

Why location, identity and query only as granules? We consider L, I and Q as privacy granules due to their sensitive nature in mobile environment i.e., compromising different combinations of L, I and Q of a mobile user under different usage scenarios would reveal several critical, personal characteristics of the user to untrusted parties. For examples; (a) Knowing a mobile user's location at certain instance of time near a political rally might reveal his political affiliations. (b) Similarly, knowing that same mobile user is interested in nearest gay dance bar, might reveal his personal interests to same untrusted third parties. In addition, user studies have indicated that mobile users are primarily concerned with threat posed to their location privacy from untrusted parties while using location-based services [16], [12] since location compromise could sometimes lead to identifying them and further lead to stalking attacks [22], [25]. Finally, when users' queries (search interests) are linked to their location, identity details, untrusted third parties can easily profile users based on their likes, dislikes during certain time periods in certain locations. Since, majority of Android apps are developed and published by independent developers, firms and they share user data with multiple third party ad libraries, content providers etc., the privacy threat to user's L, I and Q gathered by these apps has never been more serious than in current scenario.

IV. Methodology Overview

We considered access to, or sharing of data pertaining to a privacy granule as threat to that privacy granule. In order to detect threat to users' privacy granules from an app and associated ad libraries, we used a combination of static and dynamic analysis techniques on that app. The aim of our analysis techniques was to detect whether that app or an ad library included in it has access to users' privacy granules. We explain the process below.

A. Identifying Location, Identity and Query Data Access

Determining identity and location information: In Android 4.4, there are 145 permissions that restrict access to device resources (like location sensor, contacts list, device ID etc.) for an app during installation time. We categorized these permissions into 2 categories: L_p and I_p where L_p holds permissions that are required to access user location and I_p holds permissions required to access identity information. For example, to access a user's location, there are two permissions that an app declares in $manifest.xml$ (app's configuration file): ACCESS_COARSE_LOCATION and ACCESS_FINE_LOCATION and we categorized these in to L_p. To access identity related information, permissions like GET_ACCOUNTS, READ_PHONE_STATE etc. that allow an app to access user accounts (like Facebook, Google, Twitter logins etc.) on device and obtain user device's unique identifier respectively are categorized into I_p. All permissions that give access to device's unique IDs, accounts, read contacts, read SMS, calendar, profile information, phone calls on device are categorized in to I_p. Table I shows a complete list of permissions we considered for location and identity information.

Determining query information: Since user queries in apps are accepted through a search functionality, we detected if the target app implements a search functionality. When implementing search functionality, all Android apps are required to define a search configuration file (in XML format) that delivers search queries and provides search suggestions. Existence of this configuration file in an app implies that the app provides search query functionality to the user and we considered all data entered in that search field as query information.

B. Static and Dynamic Analysis

Static analysis to detect data available to app provider: In static analysis, we analyzed the target app in its $.apk$ format (the package file format in which the app is distributed). We extracted and analyzed the $manifest.xml$ from $.apk$ file to determine the permissions requested by the app. We compared this permissions list with L_p and I_p to identify if the app is accessing any location and identity related information from the user's device. The analysis of $manifest.xml$ also gave us ad libraries that are included in the app, since each ad library's package is specified in $manifest.xml$. Finally, to identify if an app accepts query information, we determined the existence of search configuration file in $.apk$ file through static analysis. Thus, we used static analysis to identify (a) App's access to location and identity information, (b) Enumerate included ad libraries and (c) Identify if the app accepts any search query from the user.

Dynamic analysis to detect data available to third parties:

In dynamic analysis, we captured the network traffic of the app while running it and analyzed the captured traffic data to determine what information was being sent to different third party ad library server IPs and isolated information of interest pertaining to our granules. We primarily used TaintDroid and Robotium [1] for our dynamic analysis. TaintDroid is a tool that determines which sensitive information is triggered and where it is sent for an app. Robotium is a Android test automation framework that we used to write test cases for auto app navigation by executing graphical elements on each screen. We captured logs of network communication during the auto navigation process and analyzed them to find items of interest like location information, identity information, and search query terms.

Data set of Apps: We collected the top 50 free Android apps (as of February 2014) from Google Play Store and implemented above mentioned static and dynamic analysis methodology on them. Thus, we could identify threats to

TABLE I. PERMISSIONS CONSIDERED FOR LOCATION AND IDENTITY PRIVACY GRANULES

Location		
ACCESS_FINE_LOCATION	ACCESS_COARSE_LOCATION	INSTALL_LOCATION_PROVIDER
Identity		
ACCOUNT_MANAGER	READ_CALL_LOG	SEND_RESPOND_VIA_MESSAGE
AUTHENTICATE_ACCOUNTS	READ_CONTACTS	SEND_SMS
BODY_SENSORS	READ_EXTERNAL_STORAGE	SUBSCRIBED_FEEDS_READ
CALL_PHONE	READ_HISTORY_BOOKMARKS	USE_CREDENTIALS
CALL_PRIVILEGED	READ_PHONE_STATE	WRITE_CALENDAR
CAMERA	READ_PROFILE	WRITE_CALL_LOG
GET_ACCOUNTS	READ_SMS	WRITE_CONTACTS
MANAGE_ACCOUNTS	READ_SOCIAL_STREAM	WRITE_PROFILE
PROCESS_OUTGOING_CALLS	RECEIVE_MMS	WRITE_SMS
READ_CALENDAR	RECEIVE_SMS	WRITE_SOCIAL_STREAM

privacy granules posed by each of the top 50 apps and ad libraries included in them.

V. GRANULAR PRIVACY THREAT REPRESENTATION

The primary requirement for the new icon-based privacy threat interface was: It should present privacy threats posed to users' privacy granules from all parties associated with an app and it should be intuitive and usable on the screen size of a mobile device (averaging 3.5 inches diagonally).

A. Design Rationale

Research suggests that the text heavy privacy policies are very difficult for end users to comprehend [17], [23]. On Android, the permissions list is the only source currently available to users that communicates to them what information is accessed from their device. However, as mentioned in [9], [19], users find it difficult to understand the permissions list. This is primarily due to the technical nature of the permissions list, lack of risk notation in the permissions list, and finally due to the list's inability to capture users' attention. Considering the above limitations, we propose a novel icon-based interface for privacy threat visualization. Our reasoning behind selecting an icon-based interface is as follows: (a) Since, we are going to represent privacy threat in terms of only three granules, it would not overload users to represent those three attributes as simple icons. This will make use of the small mobile screen space efficiently and assist in gaining user attention. (b) When icons are combined with appropriate text explanation, they could cover time, speed up users' decision process and improve their comprehension.

B. Design Process

Our design process included three steps:

Step 1 - Designing icons for location, query, and identity granules: We created three icon variants for location, query, and identity. For the location icon, we used prevailing icons that are used in mapping and navigation software. For the query icon, we designed three variants that represent searching for information over Internet. For the identity icon, we designed three variants that represent a person or his personal details. All icons were presented in gray scale to remove biases to colors from our users.

Step 2- Adding threat and safe symbols: To represent possible threat to privacy granules, we used the standard danger symbol, an exclamation mark surrounded by a red triangle. To represent safety, we used a standard check mark surrounded by a green circle. In cases where we were unsure if access was possible, we displayed a question mark surrounded by an yellow circle. The above symbols are displayed at bottom right of the privacy granule icons based on the analysis results for a particular app.

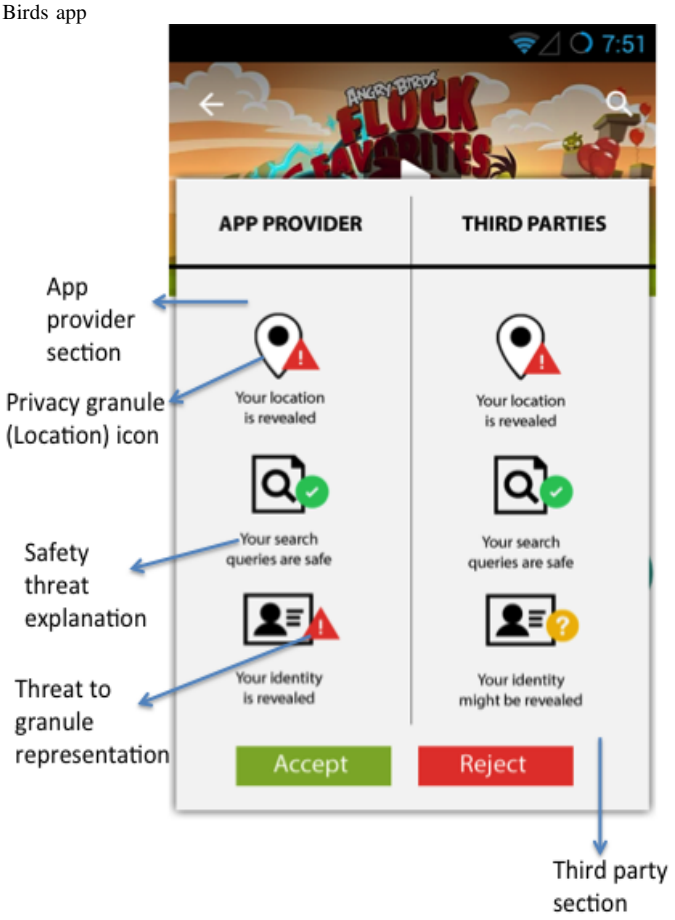
Step 3 - Putting the final interface together: Our final interface design that communicates privacy threat to the user is a replacement to current permissions list. The interface has two sections: 'App Provider' and 'Third Parties'. The 'App Provider' section shows threat(s) posed to user's privacy granules based on data accessible to the app provider. The 'Third Parties' section shows threat(s) posed to user's privacy granules through the data accessible or shared with third party ad libraries.

Figure 1 shows the new proposed interface. Each icon is described by a one line sentence that explains the threat or safety aspect for that respective granule. In addition to the 'Accept' button, the new interface also provides a 'Reject' button that enables the end user to opt out of app installation in cases where might feel uncomfortable with the app's (and included third parties) data access capabilities. With the proposed interface, we envision to present privacy threats from both app provider and third parties in a consolidated display to the end user.

VI. USER STUDY

We performed user study of new icon-based privacy threat interface to obtain user feedback on our approach. Our user study was divided into two phases: (a) In phase I, we conducted a user study for obtaining feedback on icon variants for location, query and identity. The aim of phase I user study was to select three icons, one each for location, query, and identity from variants we created. We wanted to use the icons that would be the most recognizable and understandable to users. (b) In phase II, we created a user study to evaluate the new icon-based privacy threat interface. In this phase, users were shown the current permissions list interface and new icon-based privacy threat design interface for same app and user feedback was obtained.

Fig. 1. Our proposed novel privacy threat design as shown to users for Angry Birds app



A. Phase I User Study

Study Design: We conducted an online survey for obtaining user feedback on all variants of icons we created for location, query and identity. We recruited participants using Amazon's Mechanical Turk (AMT). We designed each Human Intelligence Task (HIT) as a set of explanations and questions. We explained to users what query, identity, and location information mean in mobile environments and showed them our icon variants. For each icon variant we obtained their opinion using a 5-point Likert scale ranging from Highly Likely (5) to Highly Unlikely (1). For example, we showed the users all variants of an identity icon and asked them "How likely are you to select image 1 to represent a person's identity?" All HITs were completed in 6 days, we received 160 responses out of which 23 were discarded, since they were partial. Hence, we obtained 137 valid responses for phase I. On an average it took 5 minutes 20 seconds for participants to complete the HIT and each participant was paid $0.20 per HIT.

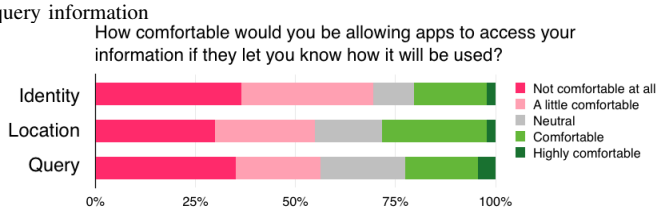
User Validation: We used the following validation test to ensure that target users of our survey were Android users. All participants were asked to provide Android OS version of their device before continuing to take the survey. We provided instructions for users to find the OS version on their device. 52.94% of our participants used Jelly Bean, 16.17% used Gingerbread and 13.23% used Kit-Kat. Finally, we limited our participants only to United States and required them to have a lifetime approval rate higher than 90% i.e., the rate of successfully completing previous tasks.

B. Results of Phase I User Study

The aim of our phase I questionnaire was to pick best icon for identity, location and query based on user feedback and comments. We obtained Likert scale ratings (1-5) for each icon from users. This allowed us to gather data reflecting users' opinion on each icon separately, rather than obtaining their opinion in terms of relative comparison. We show all user responses for each icon in figure 3. We present our analysis results in table II. It shows the mean (μ), standard deviation (σ) and Z-score to percentile rank for each icon. We used the Z-score technique, since it converts the raw score into normal score. We calculated the Z-score where observation value is a reasonable benchmark to which the mean is compared with. We set the observation value to 75% for a 5 point Likert scale i.e., (5*.75=3.75).

In addition to asking users for feedback on icons, we asked users two more questions to know their interest and comfortability about third party data sharing. The questions were: (i)"Are you interested in knowing how data collected from your smartphone by mobile apps is being used and shared by the app provider?". We allowed users to answer with a Yes/No option to answer. 114 participants out of a total 137 i.e., approx. 83% answered Yes, which shows users' interest in having access to this information and better understanding about the data sharing policies of apps.
(ii)"How comfortable would you be allowing apps to access your information if they let you know how it will be used?" i.e., their comfortability level when apps let them know how their location, identity and query information is going to be used by app providers. We collected user responses on a Likert scale ranging from "Not comfortable at all" to "Highly comfortable". We show user responses for this question in figure 2. It can be seen that only a small portion of users were highly comfortable while majority of users were (highly) uncomfortable if information related to their identity was revealed.

Fig. 2. User comfort with applications using their identity, location, and query information



C. Phase II User Study

The aim of our phase II user study was to determine the effectiveness of our solution; specifically determining the effectiveness of including privacy threat from third parties associated with the app, compare our privacy threat representation with existing solution i.e., with current Android permissions list and finally obtain user feedback on our granular approach to privacy threat representation.

Fig. 3. User selected icons

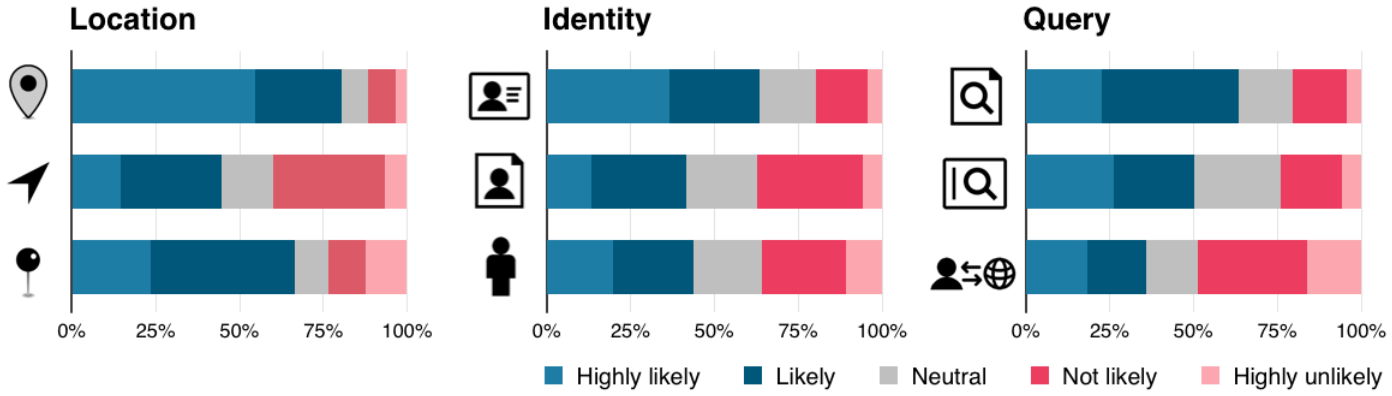


TABLE II. PHASE I SURVEY ANALYSIS RESULTS

Location				Identity				Query			
Icons	μ	σ	Z-score to Percentile rank	Icons	μ	σ	Z-score to Percentile rank	Icons	μ	σ	Z-score to Percentile rank
	4.22	**1.07**	**76%**		**3.96**	**1.21**	**56%**		**4.09**	**1.10**	**62%**
	3.11	1.22	30%		3.13	1.17	29%		3.47	1.23	41%
	3.55	1.30	43%		3.19	1.32	33%		2.91	1.36	26%

TABLE III. PHASE II SURVEY DETAILS. APP USED IN EACH SURVEY AND CRITERIA FOR SELECTING THAT APP

App name	Criteria for selection
Angry Birds	Sends location to third party ad libraries, compromising the location granule. Top 50 app.
Fruit Ninja Free	Sends location to third party ad libraries, compromising the location granule. Top 50 app.
Deer Hunter	Sends users' IMEIs to third party ad libraries, compromising the identity granule.
Yelp	One of the most used location-based services apps, shares user data with third parties and content distributors. Might be sharing user search queries with third parties thus compromising their search query details.

Study Design: Based on our findings from analysis of top 50 Android apps, we identified data accessible to app providers and ad libraries included them. We created 4 surveys based on our findings. Each of the 4 surveys was focused around a single app and privacy threats it might pose to end users. The primary reason for selecting 4 different apps was to cover different possible threat scenarios posed to users. For example, for *Angry Birds* app, we detected that it is sending location information (through traffic analysis) to third party ad library whereas for $Yelp$ app we could not determine such data sharing through traffic analysis. However, $Yelp$'s privacy policy states that it shares data with content distributors and their contractors have access to user data [26]. Hence, it was important to present these different threat scenarios using our new icon-based privacy threat interface to users. Table III presents details about apps used in each survey and rationale behind selecting for them.

In all 4 surveys, users were asked same set of questions pertaining to different apps. In each survey, we first explained our privacy granule definitions, showed them the current permissions list for a particular app, and asked questions related to easiness of finding threat to their privacy granules from the app provider. Next users were shown our proposed interface (customized with findings for each app) for the same app and we asked questions related to easiness of finding threat to their privacy granules from app providers and third parties. Next, we asked users to evaluate the effectiveness, impact of the proposed interface during install time. Finally, we asked users their opinion about our granular approach and their perception regarding the identity granule.

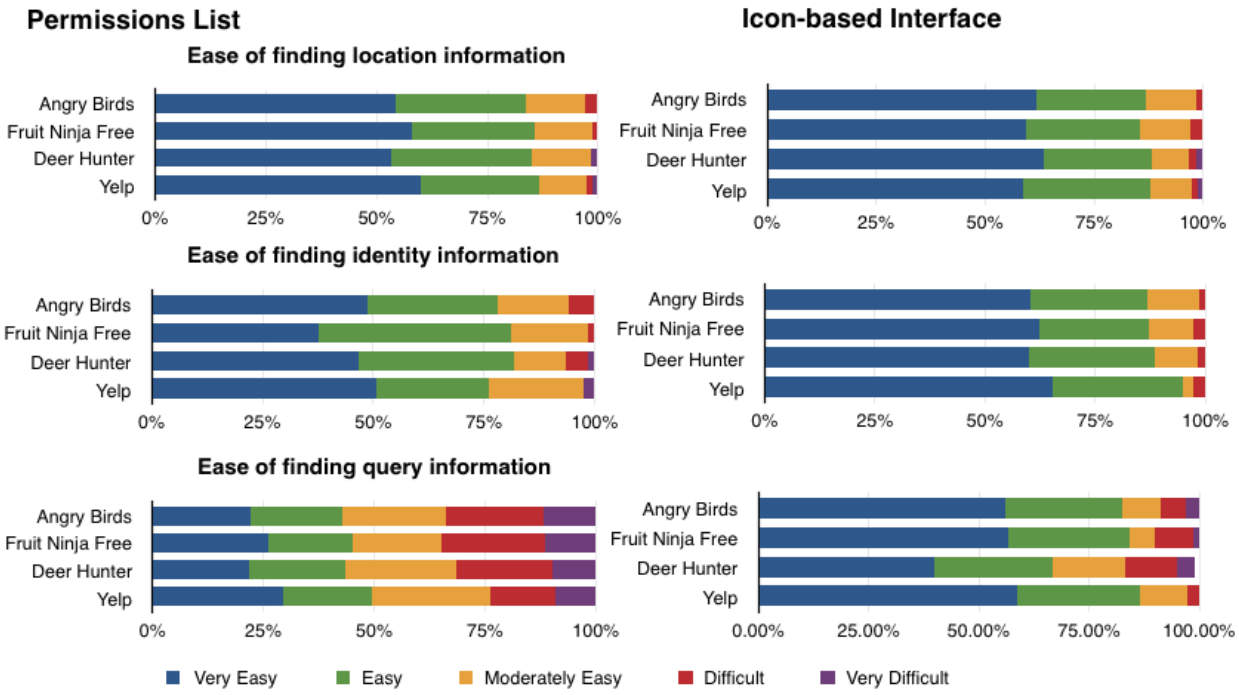
User Validation: In addition to using same methodology as phase I for user validation, we also used another validation question in this phase. After presenting the users with the proposed interface, we asked them which privacy granules had threat from third parties associated with the app. Users were given radio button-based answer options with each option being different combinations of privacy granules. Correct response to this question would ensure that participants have actually studied the new interface before answering questions related to the proposed interface. In our analysis, we only considered responses of participants who got it correct for the validation question.

We recruited participants for all 4 surveys using AMT and designed each HIT as a set of explanations and questions. All HITS were completed in 7 days, we received 324 responses out of which 52 were discarded, since they were either partial or did not meet our validation criteria. Finally, we obtained 272 valid responses for all 4 surveys. On an average, it took 9 minutes 21 seconds for participants to complete the HIT and each participant was paid $0.30 per HIT.

D. Results of Phase II User Study

Ease of Use of the New Icon-based Interface

Fig. 4. Detailed Likert responses for ease of use



We measured the ease of use of the proposed interface using two criteria:

(a) ***Criteria 1***: **Easiness in identifying privacy threats from app provider**: While using this criteria as measurement, we asked users how easy it was to find if there is threat to their privacy granules by reading the permissions list. The question for determining ease of use to find threat to location granule was: "How easy it is to find if your location information is being accessed by reading the permissions list?". Users were given Likert scale options to answer ranging from "Very Difficult (1)" to "Very Easy (5)". The same question was again asked to users after showing them the proposed interface and their responses were collected on a similar Likert scale. Figure 4 shows Likert scale answers for each survey. Analysis of Likert responses are shown in table IV. It can be observed that ease of finding threat to identity and query information in our proposed interface were statistically significant (p <.05, p <.005 and p <.0001) in pair-wise t-tests for all apps. These results suggest that users felt identifying threat to identity and query is more easy in the proposed interface than the permissions list. At the same time, we observed that finding threat to location information was easier in the proposed interface than in the permissions list. However the difference was not statistically significant.

(b) ***Criteria 2***: **Easiness in identifying privacy threats from third parties**: We did not compare our proposed interface with permissions list for this criteria, since the permissions list does not present threats posed from third parties. Users were asked "How easy it is to find if your identity information is shared with third parties?" to determine if they could detect threat to their identity granule from third parties. Similar questions were asked for detecting threats to location and query granules. Users were given Likert scale options to answer ranging from "Very Difficult (1)" to "Very Easy (5)" to respond. Table V shows analysis of user responses for all surveys. We calculated Z-score of Likert responses with the observation value set to 70% (3.5) for a 5 point Likert scale and determined the significance of user responses. It can be observed that all user responses were statistically significant (at p <0.05) establishing that users were easily able to detect threat to their privacy granules from third parties through our proposed interface.

TABLE V. ANALYSIS OF LIKERT RESPONSES RECEIVED FOR EASINESS OF FINDING THREATS FROM THIRD PARTIES

*p<0.05	Location		Identity		Query	
	Z-score	p	Z-score	p	Z-score	p
Angry Birds	2.11	*	2.01	*	2.15	*
Fruit Ninja Free	2.04	*	2.07	*	1.99	*
Deer Hunter	2.04	*	2.29	*	2.06	*
Yelp	1.96	*	1.98	*	2.00	*

Effectiveness of including threats from third parties

We asked users to rate the effectiveness of the proposed interface in order to understand the impact of showing privacy threats posed from third party ad libraries before they installed the app. Users were given Likert scale options to answer ranging from "Strongly Disagree(1)" to "Strongly Disagree(5)" and presented following questions.

How much do you agree with following statements:
S1.This interface would help you to understand privacy risks posed to your location, identity and search query from both the app and other third parties.
S2.This interface would help you to decide whether or not to install the app because third parties accessing data is important to know.

Table VI shows analyzed results for each question. We

TABLE IV. COMPARISON OF LIKERT RESPONSES RECEIVED FOR EASINESS OF FINDING THREATS FROM APP PROVIDERS. DEGREES OF FREEDOM FOR T-TEST ARE SHOWN IN PARENTHESIS BESIDE THE APP NAME. FOR PERMISSIONS LIST AND PROPOSED INTERFACE, WE SHOW MEAN AND STANDARD DEVIATION IN PARENTHESIS

$*\ p < 0.05$ $**\ p < 0.005$ $***\ p < 0.0001$	Location			Identity				Query			
App Name	Permissions List	Proposed Interface	T	Permissions List	Proposed Interface	T	p	Permissions List	Proposed Interface	T	p
Angry Birds(67)	4.35(0.82)	4.47(0.76)	1.21	4.21(0.92)	4.46(0.76)	2.07	*	3.19(1.33)	4.26(1.05)	5.61	***
Fruit Ninja Free(68)	4.42(0.77)	4.42(0.81)	0.00	4.17(0.77)	4.46(0.80)	2.05	*	3.25(1.38)	4.29(1.02)	5.20	***
Deer Hunter(59)	4.35(0.84)	4.47(0.85)	0.85	4.2(0.95)	4.47(0.75)	2.16	*	3.23(1.29)	3.85(1.22)	3.01	**
Yelp (74)	4.43(0.85)	4.43(0.82)	0.00	4.21(0.96)	4.57(0.68)	2.90	*	3.45(1.31)	4.43(0.79)	5.88	***

calculated Z-score with observed value at 70% (3.5). For all apps, users' responses were statistically significant (p <0.05) for both questions. This suggests that comprehensive presentation of privacy threats from all entities associated with an app before installing the app would help users make a better installation decision.

TABLE VI. ANALYSIS OF USER RESPONSES FOR QUESTIONS ON EFFECTIVENESS OF THE PROPOSED INTERFACE

* p < 0.05	S1			S2		
	μ & σ	Z-score	p	μ & σ	Z-score	p
Angry Birds	4.63(0.54)	2.08	*	6.42(0.55)	2.04	*
Fruit Ninja Free	4.52(0.50)	2.02	*	4.56(0.50)	2.10	*
Deer Hunter	4.55(0.50)	2.09	*	4.6(0.49)	2.23	*
Yelp	4.49(0.50)	1.97	*	4.55(0.50)	2.09	*

Self Reported Installation Decision

We wanted to know if the proposed interface design would influence users' app installation decision. To this end, we asked all users (at beginning of the survey) if they had used the app before with "Yes/No" options. Later, at the end of the survey we asked users how likely were they to install the app after viewing the proposed interface. We presented them with Likert options to answer ranging from "Not likely at all (1)" to "Very likely (5)". The question asked was: "How likely are you to install this app after seeing the above interface?". Figure 5 shows the Likert responses for each app. For each app, more than 60% of users responded either with "Not Likely" or "Not Likely at all" to install the app after viewing the proposed interface. We present our findings for app install decision change in Table VII. The 'Yes → 'Not likely/Not likely at all' represents number of users who have used the app and have changed their responses to 'Not likely' or 'Not likely at all' to install after looking at the proposed interface. It can be observed that majority of users who have used the Angry Birds and Fruit Ninja apps tended towards not to install them. A possible reason for this could be unauthorized access of their location information by third parties in these apps.

How do users perceive their identity?

To evaluate the feasibility of our approach, we wanted to understand how users perceive threat to their identity granule. To this end, we asked users questions related to their identity information.

At the end of the survey, we asked users if they wished to see the identity granule divided in to more sub-groups than what it currently represented such that we could further enhance our granular approach in future work. The question was: "In the proposed interface, identity has information related to your email address, phone number, IMEI number, phone call details, contacts details, SMS, calendar details, microphone, camera and photos. Do you think that this grouping should be split?". Users were given radio button answer options that were: "I think this group is good", "I think this group is good but a couple of those things should get their own icons", "I need all this information split up into different groups" and "I want to see every permission separately". Out of total 272 responses received, 62 responded with "I think this group is good" and 150 responded with "I think this group is good but a couple of those things should get their own icons" suggesting that users are interested to see more granular approach to their identity granule. We presented a follow up question: "Which information that might reveal your identity would you like to see in a separate group represented by an icon for itself ?". Users were given check-box based answer options: Camera, Contacts, Calendar. Photos and SMS. We show top 5 combinations based on number of user responses in figure 6. It can be noted that majority of users consider threat to their contacts and SMS as threat to their identity.

Fig. 5. Users' likelihood of installing apps after viewing the proposed interface

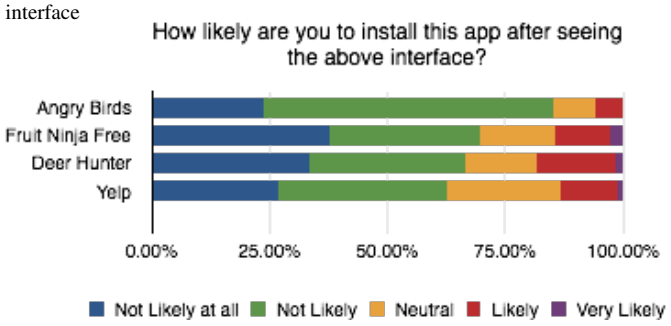
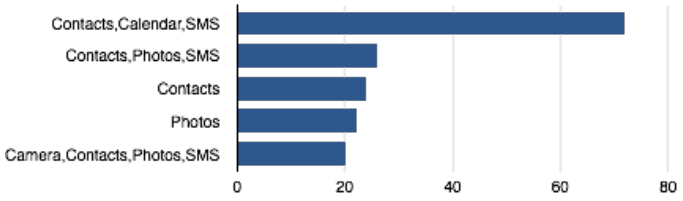


TABLE VII. USERS' SELF REPORTED APP INSTALLATIONS

	Before viewing icon interface		After viewing icon interface
App name	Yes	No	Yes → Not likely/Not likely at all
Angry Birds	38	30	31
Fruit Ninja Free	24	45	14
Yelp	33	42	14
Deer Hunter	4	56	1

Fig. 6. Identity granules users are most interested in



VII. Discussion

The goal of our work was to create a usable alternative solution to current Android permissions list that presents privacy threats from all parties associated with an app to end users. We wanted to find if users would comprehend our new icon-based interface and make an informed decision during installation of that app. Overall, majority of the users appreciated our proposed interface and provided feedback that it helps them better understand which entities are able to access their data and in some cases such interface would influence their decision to install the app.

A. Prevalence of ad libraries and users' concerns

Our analysis of top 50 free Android apps revealed that 29 of them included 22 distinct ad libraries among them. Popular apps like *My Talking Tom*, *Super-Bright LED Flashlight*, *Fruit Ninja Free* and *Angry Birds* have access to user's information on device and include 11, 9, 7 and 5 ad libraries respectively. This shows the prevalence of ad libraries among most used apps and wide scope for user information dissemination among untrusted third parties. Findings pertaining to sharing of sensitive user information with multiple third parties by *Angry Birds* [15] also supports this claim.

Majority participants stated that they would feel least comfortable with such information sharing practices. One of the user comments in this regard was: "it makes me uncomfortable to know how much information is revealed", helps us understand users' concern with undisclosed information sharing and data access procedures in place. Therefore, we believe our interface is a huge step towards educating users by providing a consolidated threat interface that includes privacy threats from app provider and associated third parties.

B. Usability of the proposed interface

Icons for location and identity granules selected from our phase I survey are same as those used in current Android permissions list for those granules, which further establishes their wide acceptance. It should be noted that our Phase 1 survey was conducted in February'14, before icons were introduced in Android permissions list.

Our user studies suggest that using icons with a simple design and appropriate threat representation captures users' attention better than regular text-heavy descriptions. Majority of users responded extremely positive to our icon-based privacy threat interface. For instance, out of 89 comments received for *Angry Birds* app, 69 were positive, 10 were neutral, and 10 were negative towards the new interface. Some of the positive comments were: "I like to know who has access to my information so that I can decide whether to share it.Doing so backhandedly annoys me as I'm not in control of something personal.", "I think its useful to know what kind of information is being shared when deciding to purchase/download an app.", "Its very useful to see, There is no downside to it, its very informative and helps you distinguish between good apps and bad apps." and "knowing what information is being shared with third parties will affect my judgement on weather or not to install an app."

Improving the proposed interface: While majority of users appreciated the proposed interface, there was a set of users who provided feedback to further enhance the design and usability of the proposed interface. Primarily, users wanted to see more precise details about information sharing i.e., they wanted to know exactly what information was being shared with which third parties and why. Especially in cases where third parties could access identity information and there was no confirmation regarding this. Some comments related to this from users were: "I liked seeing what exactly is expected to be shared and not shared, I didn't like the possible description. I think it would be best to have it be a yes or no, plain and simple.", "It doesn't tell how it will be used." and "it's useful to see, but still unknown *why* they want access/permission at all." Considering more privacy granules, especially representing identity as more granules might allow users to see more precise information regarding what exactly is being shared that might reveal their identity. As specified above, users are more interested to see threats to their contacts and SMS details. However, screen sizes of smart phones might pose a challenge to represent high number of privacy granules in the proposed interface. An avenue we plan to pursue in this regard is to implement a scrollable privacy threat interface and subject that to more user studies and feedback.

C. Limitations

Both of our studies were performed through Amazon's Mechanical Turk, which has a number of known biases. We attempted to mitigate these by working to capture only actual Android users as participants and by asking a mix of free-response and multiple choice questions. As our designs were shown in studies, our participants were not actually downloading apps on to their own devices, which may cause any number of changes in the way they behave in the real world compared to these results. Especially, users' decision to install the app after looking at the proposed interface might be different in real world depending upon the underlying usage context. Further testing of this and other Android permission replacement interfaces demands real world study. Finally, we only tested a small number of apps and are in the process of expanding this work to a much wider range of app types and possible privacy threats.

VIII. Conclusion and Future Work

Owing to lack of clarity in current Android permissions list regarding threat posed from ad libraries included in apps, we proposed an alternative icon-based privacy threat display (using granular privacy approach) for Android users that presented privacy threats from both app providers and third parties. Through rigorous user studies, we have uncovered that mobile users are interested to adopt more such usable and informative interfaces that might have an impact on their app installation decision.

Our long-term vision is to use our methodology for profiling ad libraries in terms of what information they can gather about users such that they could track them across multiple apps. We are currently analyzing top 100 apps in all categories of Google Play Store and plan to represent our findings for each app (through a publicly accessible web interface and mobile app) using the proposed interface and assist users in making an informed decision regarding installation of an app of their interest.

REFERENCES

[1] https://code.google.com/p/robotium/.

[2] AppBrain, "Free vs. paid Android apps," http://www.appbrain.com/stats/free-and-paid-android-applications.

[3] D. Barrera, H. G. Kayacik, P. C. van Oorschot, and A. Somayaji, "A methodology for empirical analysis of permission-based security models and its application to android," in *Proceedings of the 17th ACM Conference on Computer and Communications Security*, ser. CCS '10. New York, NY, USA: ACM, 2010, pp. 73–84.

[4] 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*, ser. HotMobile '11. New York, NY, USA: ACM, 2011, pp. 49–54.

[5] R. Bhoraskar, S. Han, J. Jeon, T. Azim, S. Chen, J. Jung, S. Nath, R. Wang, and D. Wetherall, "Brahmastra: Driving apps to test the security of third-party components," in *Proceedings of the 23rd USENIX Conference on Security Symposium*, ser. SEC'14. Berkeley, CA, USA: USENIX Association, 2014, pp. 1021–1036.

[6] T. Book, A. Pridgen, and D. S. Wallach, "Longitudinal analysis of android ad library permissions," *CoRR*, vol. abs/1303.0857, 2013.

[7] 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*, ser. OSDI'10. Berkeley, CA, USA: USENIX Association, 2010, pp. 1–6.

[8] 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*, ser. CCS '11. New York, NY, USA: ACM, 2011, pp. 627–638.

[9] 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*, ser. SOUPS '12. New York, NY, USA: ACM, 2012, pp. 3:1–3:14.

[10] A. P. Fuchs, A. Chaudhuri, and J. S. Foster, "Scandroid: Automated security certification of android applications," 2009.

[11] M. C. Grace, W. Zhou, X. Jiang, and A.-R. Sadeghi, "Unsafe exposure analysis of mobile in-app advertisements," in *Proceedings of the Fifth ACM Conference on Security and Privacy in Wireless and Mobile Networks*, ser. WISEC '12. New York, NY, USA: ACM, 2012, pp. 101–112.

[12] U. Hengartner and P. Steenkiste, "Access control to information in pervasive computing environments," 2003.

[13] 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*, ser. CCS '11. New York, NY, USA: ACM, 2011, pp. 639–652.

[14] IDC, "Smartphone OS Market Share, Q1 2014," http://www.idc.com/prodserv/smartphone-os-market-share.jsp.

[15] T. W. Jimmy Su and J. Zhai, "Man Accused of Stalking Ex-Girlfriend With GPS," https://www.fireeye.com/blog/threat-research/2014/03/a-little-bird-told-me-personal-information-sharing-in-angry-birds-and-its-ad-libraries.html.

[16] E. Kaasinen, "User needs for location-aware mobile services," *Personal Ubiquitous Comput.*, vol. 7, no. 1, pp. 70–79, May 2003.

[17] P. G. Kelley, J. Bresee, L. F. Cranor, and R. W. Reeder, "A "nutrition label" for privacy," in *Proceedings of the 5th Symposium on Usable Privacy and Security*, ser. SOUPS '09. New York, NY, USA: ACM, 2009, pp. 4:1–4:12.

[18] 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 *Proceedings of the 16th International Conference on Financial Cryptography and Data Security*, ser. FC'12. Berlin, Heidelberg: Springer-Verlag, 2012, pp. 68–79.

[19] P. G. Kelley, L. F. Cranor, and N. Sadeh, "Privacy as part of the app decision-making process," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '13. New York, NY, USA: ACM, 2013, pp. 3393–3402.

[20] J. Lin, S. Amini, J. I. Hong, N. Sadeh, J. Lindqvist, and J. Zhang, "Expectation and purpose: Understanding users' mental models of mobile app privacy through crowdsourcing," in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, ser. UbiComp '12. New York, NY, USA: ACM, 2012, pp. 501–510.

[21] 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*, ser. ASIACCS '10. New York, NY, USA: ACM, 2010, pp. 328–332.

[22] F. News, "Man Accused of Stalking Ex-Girlfriend With GPS," http://www.foxnews.com/story/2004/09/04/man-accused-stalking-ex-girlfriend-with-gps/.

[23] R. W. Reeder, P. G. Kelley, A. M. McDonald, and L. F. Cranor, "A user study of the expandable grid applied to p3p privacy policy visualization," in *Proceedings of the 7th ACM Workshop on Privacy in the Electronic Society*, ser. WPES '08. New York, NY, USA: ACM, 2008, pp. 45–54.

[24] R. Stevens, C. Gibler, J. Crussell, J. Erickson, and H. Chen, "Investigating user privacy in android ad libraries," in *IEEE Mobile Security Technologies (MoST)*, San Francisco, CA, 2012.

[25] U. Today, "Authorities: GPS system used to stalk woman," http://usatoday30.usatoday.com/tech/news/2002-12-30-gps-stalker_x.htm.

[26] Yelp, "Privacy Policy," http://www.yelp.com/tos/privacy_en_us_20130910.

[27] Y. Zhou, X. Zhang, X. Jiang, and V. W. Freeh, "Taming information-stealing smartphone applications (on android)," in *Proceedings of the 4th International Conference on Trust and Trustworthy Computing*, ser. TRUST'11. Berlin, Heidelberg: Springer-Verlag, 2011, pp. 93–107.