Know Your Enemy: The Risk of Unauthorized Access in Smartphones by Insiders

Ildar Muslukhov
The University of British Columbia
Vancouver, BC, Canada
ildarm@ece.ubc.ca

Yazan Boshmaf
The University of British Columbia
Vancouver, BC, Canada
boshmaf@ece.ubc.ca

Cynthia Kuo
Vibrado Technologies
Sunnyvale, CA, USA
cynthia@vibradotech.com

Jonathan Lester
Nokia Research Center
Palo Alto, CA, USA
jonathan.lester@gmail.com

Konstantin Beznosov
The University of British Columbia
Vancouver, BC, Canada
beznosov@ece.ubc.ca

ABSTRACT

Smartphones store large amounts of sensitive data, such as SMS messages, photos, or email. In this paper, we report the results of a study investigating users' concerns about unauthorized data access on their smartphones (22 interviewed and 724 surveyed subjects). We found that users are generally concerned about insiders (e.g., friends) accessing their data on smartphones. Furthermore, we present the first evidence that the insider threat is a real problem impacting smartphone users. In particular, 12% of subjects reported a negative experience with unauthorized access. We also found that younger users are at higher risk of experiencing unauthorized access. Based on our results, we propose a stronger adversarial model that incorporates the insider threat. To better reflect users' concerns and risks, a stronger adversarial model must be considered during the design and evaluation of data protection systems and authentication methods for smartphones.

Author Keywords

Smartphone, Physical Threats, Theft, Loss, Stranger, Insider, User Study

ACM Classification Keywords

K.6.5 Security and Protection: Unauthorized access (e.g., hacking, phreaking)

General Terms

User Study; Smartphone Data Security; Physical Threats; Theft and Loss; Stranger and Insider;

INTRODUCTION

Smartphones are pervasive devices with more than a billion users worldwide [3]. Recent market research shows that the adoption of smartphone continues to rise [17]. Storage capabilities of smartphones have improved significantly in the last decade, allowing users to store greater amounts of data, including photos, videos, SMS messages, and emails. Some owners may consider these data to be sensitive. In this paper, we define data as being *sensitive* if a user would be concerned about someone accessing it without her permission.

High adoption rates of smartphones have made them appealing targets for adversaries. In fact, the annual security threat report by Sophos shows that, in 2012, attackers were focused mainly on exploiting the mobile platforms [30]. Malware in smartphones has attracted significant attention in the research community [5, 7, 8, 10, 15, 16, 24]. However, the research community has focused on one particular type of adversary—one that seeks to perform a remote and highly scalable attack—but ignores an attacker with physical access to the device.

Smartphones are portable; they are more susceptible to theft and loss relative to laptops and desktops [1, 4, 22]. When a smartphone is lost, the person who finds it tries to access sensitive data in 96% of the cases [2]. In a study by Symantec, subjects who found a smartphone accessed sensitive data such as passwords managers and online banking applications. Access to these types of data cannot be justified as necessary for finding the owner of the device.

Shi et al. [28] defined an adversarial model that was used for the evaluation of an implicit authentication system on smartphones. According to that model, unauthorized access can be carried out by strangers (e.g., a person who finds a smartphone or a person who steals the phone from an unfamiliar individual) or by insiders (e.g., a friend, a coworker, an enemy, or a competitor). In general, a **stranger** is defined as *a person who is unfamiliar with the victim*, and an **insider** is defined as *a person who is familiar with the victim*. These two types of adversaries differ in their capabilities and objectives. For instance, we assume an insider may have some knowledge about the legitimate user's behavior, where a stranger does not have any knowledge about the owner.

Most existing research has focused on addressing threats by a stranger (e.g., [6, 20, 2, 27]). Others have assumed that smartphone users consider insiders as a realistic threat [28]. However, there is still no empirical evidence demonstrating the importance of the insider threat. Do users consider insiders to be a serious threat? Does unauthorized access by insiders happen in the real world?

It is the main contribution of this paper to fill this knowledge gap. We conducted two consecutive users studies: interviews with 22 subjects, and an online survey with 724 subjects. First, we found that most users consider the insider threat as important as the stranger threat. Second, we showed that more than 12% of the users have experienced unauthorized access of their data or applications on smartphones. We also identified that some demographic groups are more susceptible to unauthorized access by insiders.

This paper offers the first empirical evidence that smartphone users consider the insider threat to be an important one and that it impacts many smartphone users today. We argue that new proposals that aim to protect sensitive data in lost or stolen smartphones from unauthorized access must consider the insider threat. We also present an adversarial model that describes the capabilities and objectives of strangers and insiders. We base these capabilities and objectives on the literature and on the results of our users studies. Relative to existing adversarial models (e.g. [28]), we add new objectives (e.g., surveil the smartphone owner) and capabilities (e.g., hide attack traces).

This paper offers the following contributions: (1) it provides the first empirical evidence that the insider threat must be considered by designers of data protection systems (DPS), and (2) it presents a stronger adversarial model that can be used during the design and evaluation stages of DPS for smartphones.

RELATED WORK

Several authors have investigated users' concerns with security and privacy of their smartphones. For instance, Chin et al. [9] conducted a user study to understand users' privacy concerns when they use applications on smartphones for sensitive tasks (e.g., online banking). The authors found that users are concerned with sensitive activities on smartphones, and tend to reduce the amount of such activities. Users provided various justification for such behavior, which were rooted in their fears. Interestingly, theft and loss of a smartphone were among users' top five fears. The authors did not further investigate this observation. In comparison, our study focuses on users' concerns with unauthorized access of their data on lost or stolen smartphones.

Muslukhov et al. [23] conducted interviews and identified a list of data types that users store on the smartphones. They also examined why users consider each data type as confidential, sensitive, or valuable. This work provides justification for some objectives and capabilities that an insider may possess, but it does not provide evidence that an insider threat needs to be considered. This paper, in contrast, provides the first empirical evidence that users are concerned with insiders. In addition, the results of the user study suggest that users do experience unauthorized access by insiders today.

Dorflinger et al. [13] investigated users' attitudes on gradual security levels and novel authentication methods. The main contribution of that work is the analysis of users' concerns with various authentication methods and users' perception of the security that each method provides. The authors did not investigate user behavior or the sensitivity of various data types on smartphones. We compare users' concerns with sensitive data in the presence of strangers and insiders. In addition, we report users' experiences with unauthorized access of their smartphone data.

Similarly, Ben-Asher et al. [6] focused on studying user attitudes toward alternative authentication methods and how sensitive some data and smartphone functionalities are. The authors considered a limited set of data types (7 data types) and did not differentiate between data sensitivity with insiders and strangers. Furthermore, this paper did not report users' experiences with unauthorized access.

Finally, Shi et al. [28] present an adversarial model that is used for implicit authentication evaluation. In comparison with our adversarial model, the model proposed in this work is limited. In particular, our model considers that an attacker may want to spy on the smartphone owner without revealing that. In addition, we assume that an insider is not able to capture authentication secrets and hide the attack traces.

RESEARCH QUESTIONS

We set out to answer the following research questions: **RQ1** - Are users concerned about unauthorized access of their data or smartphone functionality by an insider?, and **RQ2** - Have users experienced unauthorized access of sensitive data, either as victims and/or as adversaries? The answers to these research questions allow us to understand whether DPS should be designed with an insider type of adversary in mind. In addition, understanding why users are acting as insiders themselves gives us a better understanding of insiders' incentives, objectives, and capabilities; this is important for a valid evaluation of DPS in smartphones.

METHODOLOGY

To answer our research questions we conducted two users studies: a set of semi-structured interviews (study 1) and an online survey (study 2).

Study 1 - Interviews

We conduced 22 semi-structured interviews in Vancouver, Canada, between September and November of 2011. Participants received 25 dollars for participating. We did not inform participants about the real nature of the study. Instead, we advertised that the interview was about users' experiences with smartphones and smartphone applications. To compare subjects' concerns with strangers and insiders, we asked them to think aloud for the following scenarios: (a) "Assume you just lost your phone on a bus (it might be stolen), what would be your reaction and would you have any concerns with such a loss", (b) "Assume that you are at a party and someone took your phone, what would be your reaction and would you have

any concerns? Does it matter if that person knows you?". During these sessions we asked users to be specific about their concerns by providing examples or naming specific applications.

Study 2 - Online Survey

The result of Study 1 provided us with an insight on why users treat the two adversaries differently. They informed the follow-up survey questions in terms of the types of previous experiences we should cover and the data types that users store on their smartphones.

In Study 2, we conducted an online survey, which allowed us to recruit a larger and more diverse participant pool. We conducted four pilot studies (between January and May 2012) with 60 subjects in total, to insure the clarity of the questions and correctness of data collection process. We did not combine data from pilot studies into the final results.

The online survey consisted of four parts. In the first part, general questions were asked on smartphone use. We asked subjects whether they used a phone locking system and if they also used a code (either PIN, Draw-a-Secret, or a password) to unlock. We then asked them to visit a web page through their smartphones, in order to record their smartphone *User-Agent*[1] string and eliminate 942 subjects who did not use a smartphone.

The second part of the survey included questions about respondents' previous experience with their smartphones, e.g., loss or damage. We also asked subjects whether they had previously accessed someone's smartphone without the owner's permission, and whether they had found someone accessed their smartphone, without their permission.

In the third part of the survey, we asked subjects about the types of data they stored on the phone. For this part of the survey, we gave them a pre-populated list of data types (compiled based on the results of Study 1) and asked them to select those that they stored. We allowed them to add new data types if necessary. Furthermore, we asked subjects separately for personal and work-related data types.

In the final part of the survey, we asked subjects to rate their agreement with the following statement, "I would not have any concerns if *Personal/Work Data Type* could be viewed by such a thief" on a 5-point Likert scale for each data type. The following options were provided: Strongly Agree, Agree, Neutral, Disagree, Strongly Disagree. The rating task was performed twice, once for a stranger scenario and once for an insider scenario. The stranger scenario was presented as, "Assume your smartphone just have been stolen by a person who does not know you [sic]," and the insider scenario as, "Assume your smartphone just have been stolen by a person who does know you [sic]."

[1]A *UserAgent* is a string that every browser sends to the web server. For instance the following string is sent from an HTC Sensation 4G that runs Android 2.3.4 - *"Mozilla/5.0 (Linux; U; Android 2.3.4; en-us; HTC Sensation 4G Build/GRJ22) AppleWebKit/533.1 (KHTML, like Gecko) Version/4.0 Mobile Safari/533.1"*

In addition, we asked each subject to rank the importance of each data type they stored on their smartphones to corroborate the results of the rating task. Ranking was performed twice, once for the stranger scenario and once for the insider scenario. In each ranking task, subjects were asked to rank the data types by their level of concern, with most concerned at the top and least concerned at the bottom.

We instrumented the survey website with tools that allowed us to track: how much time each subject spent on each question; IP addresses of the PC and smartphone used for the survey; and UserAgent strings for the PC and smartphone. Later, these data were used to remove subjects that either skimmed through the survey (23 subjects), or did not use a smartphone (942 subjects). The UserAgent string was also used to measure the representativeness of our subjects in terms of mobile platforms and OS versions.

In our data analysis, we used the Fisher Exact Test (FET) or Chi-Squared Test (CHI) for tests on contingency tables. To analyze the differences between sensitivity rates for strangers and insiders, we used the U-test (Wilcoxon rank sum test). To analyze the differences between sensitivity ranks for strangers and insiders, we used the Wilcoxon signed-rank test (WSRT).

Study 2 was conducted between May 16 and June 23, 2012. The survey was available in the US, UK, Australia, New Zealand, and Canada on Amazon's Mechanical Turk (MTurk); through other advertisement services, such as Kijiji and Craigslist; and through word of mouth. We received ethics board approvals for both studies.

RESULTS

Study 1 Demographics

We recruited 22 subjects for the interview study. Ten were male. The majority (12) were between 19 and 30 years of age. Ten held a Bachelors degree from a university or college. Half of the participants were recruited on the university campus, and the rest were recruited in a downtown area. The participants had 19 different occupations, including one unemployed participant. Nine participants used their smartphones for work-related activities, but only three of them received their smartphones from their employers.

Study 2 Demographics

For the online survey, we recruited 2,092 subjects. Only 1,725 respondents completed the survey. Further investigation revealed that only 783 of the subjects used their smartphones as required. Also, we removed participants who finished the survey in less than 10 minutes (23 participants) to exclude those who skimmed through the survey. The minimum time required to go through the questionnaire was identified during early pilot studies. Finally, we excluded 36 non-MTurk subjects to avoid an unbalanced participant pool.

The remaining 724 participants completed the survey in 25 minutes on average (std. dev., s=12.5). The majority of the participants were from the US (634); the rest were divided between Canada (50), the UK (29), Australia (9), and New Zealand (2). The majority of subjects used Android OS (391/51%) and iOS (278/37%). We did not find a statistically

significant difference for our sample platform distributions and the distributions reported by Google and Kunzler [21, 18] (FET,p>0.08). Three hundred seventy of the subjects were male (51%). The average age for the subjects was 25.6 years (s=5.98). The average annual income was $43k (s=$19k).

The participants had diverse occupations, including more than 500 different titles in 16 various industry fields, such as agriculture, business, construction, education, etc.

We compared the demographics of our subjects with the results reported by Smith [29]. To the best of our knowledge, Smith's study is the only study that provides statistics on a representative sample of the US population of smartphone users (n=2,253), and the majority of our subjects were from the US. For this part only we excluded all subjects that were not from the US (90). For the rest of the analysis, we used all (724) subjects. The analysis of differences between our subjects from the US and the ones reported by Smith's study [29] did not reveal a statistically significant difference in gender distribution. However, there was a statistically significant difference in age, income, and education. In particular, our participants appeared to be younger (29.6, $\sigma = 9.69, \chi = 361.6676, df = 3, p < 0.001$). This, however, is not surprising, as it was previously shown that MTurk subjects tend to be younger [25]. Although the difference in education and income distributions were statistically significant (FET, $p < 0.001$), we consider them practically insignificant due to small relative values. The average income in Smith's study was higher by 6% ($46k, sd = $20k), and the difference in education levels revealed that our sample had 9% more subjects with high school diploma and 9% fewer subjects with college or higher degree.

Our data analysis indicates we recruited a diverse and a representative sample, at least for the US, with a slight bias towards younger smartphone users.

Research Question 1 – "Are users concerned about unauthorized access of their data or smartphone functionality by an insider?"

In the interview study (study 1), we found an even split between two groups of subjects: (a) those who cared about the privacy of their sensitive data and functionality on their smartphones (11); and (b) those who did not care (11). Seven out of 11 subjects in the first caring group used a locking system, and the remaining four had used locking systems in the past but had stopped. These four subjects provided the following reasons why they stopped using a locking system: (a) inability to disable a locking system temporarily for a short period of time (e.g., 10-20 min), without the necessity to setup it again afterwards; (b) social discomfort of using a lock in front of their friends, where they believed trust was implied; and (c) too frequent authentication prompts.

The majority of the subjects stated that they would have higher concerns if their friends' and associates' contact details were revealed to a stranger. Subjects thought that they were implicitly required to protect the confidentiality of other people's contact details. Disclosure of contact details was seen as a negative impact on their reputation. On the other hand, subjects had higher concerns with insiders in regard to personal messages and photos. Two subjects turned a smartphone lock on only when they were at home, due to past experiences. This study provided us with interesting qualitative data, but it did not allow us to justify that the insider threat is comparable to the stranger threat.

We confirm these results in the online survey study (study 2). In particular, we found that half of the survey participants (379, 52%) used a locking system. We refer to these participants as the lock-using group (LOCK). More than 64% (243) of subjects in LOCK group did so to avoid unauthorized data access by others, and 73% (278) of them did so to avoid unauthorized access to the functionality of the phone.

The remaining (345) participants did not use a locking system. We refer to this group as OPEN. Further investigation revealed that–similar to the interview study results–the OPEN group included 155 subjects that had something sensitive on their smartphone and had used a locking system before, but had stopped due to various usability problems (too frequent authentication prompts, necessity to authenticate even if non-sensitive data are accessed). The other 190 subjects in the OPEN group did not have any sensitive data on smartphones.

Interestingly, most of the subjects in the LOCK group used either a PIN-code (206) or a Draw-a-Secret (DAS) (168) authentication method, whereas only 52 used alpha-numeric passwords. Note that subjects were able to select multiple types of authentication methods if they owned several smartphones, thus $\sum n \neq 379$. The participants in the interview study explained the choice of PIN or DAS by ease of use, in comparison to full-fledged passwords.

In order to compare users' concerns with a stranger and insider, we first asked subjects about the types of data they stored on the smartphones and which one were used for personal or work purposes. The top 15 most used data types are provided in Table 1. Note, that this table is based on users' responses, and, thus, is not supposed to be precise or complete.

Data Type (Label)	%
1 - Photos and videos (phv)	94
2 - SMS/MMS messages (sms)	93
3 - Call history (cah)	90
4 - Emails (eml)	87
5 - Contacts details (cod)	87
6 - Music (mus)	81
7 - Browser search history (bsh)	74
8 - Browsing history (bwh)	73
9 - Events in calendar (evt)	73
10 - Notes and memos (n&m)	72
11 - Data in social networking applications (osn)	68
12 - Progress in games (gam)	68
13 - Documents (doc)	64
14 - Voice recordings (voc)	42
15 - Passwords saved in applications or passwords managers) (pwd)	37

Table 1. The top 15 data types used by the subjects for personal use. No work-related data types were listed.

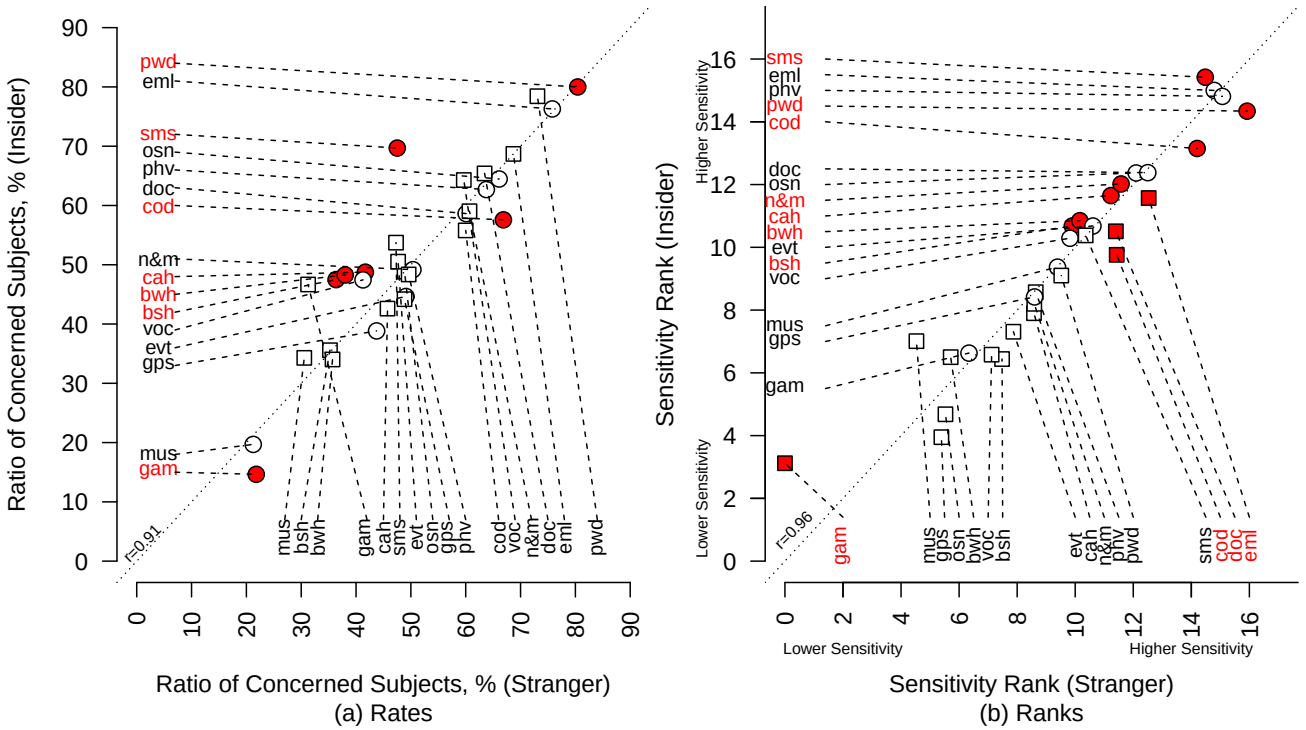

Figure 1. The proportion of concerned users with sensitivity in the presence of a stranger (horizontal axis) and in the presence of an insider (vertical axis). Data labels across the vertical axis and circles in the plots represent data types for personal use; data labels across the horizontal axis and squares in the plots represent data types for work related use. Filled shapes and red-colored data labels represent statistically significant differences between subjects' concerns with respect to a stranger and an insider (U-test for rates, WSRT for ranks, $p < 0.05$). The meanings for the abbreviated data type labels are in Table 1.

In order to answer RQ1 we tested if the difference between users' concerns with both types of the adversaries is statistically significant. We acknowledge that a larger number of subjects will reveal more statistically significant results. We, however, argue that according to the Central Limit Theorem, such results, in terms of absolute values, should not be substantially different from what is presented in this work.

First, we analyzed Likert scale ratings with U-test, since the data are ordinal, and parametric tests (t-test, ANOVA) are not applicable. The results of U-test revealed that subjects rated their concerns differently for only six data types (out of 32) . In particular, subjects expressed more concerns about an insider threat for SMS messages, call history, browsing history, and search history in the browser. And subjects were more concerned about strangers for contact details and progress in games. These results agreed with the results from the interview study.

Figure 1a shows the proportions of subjects that were concerned with strangers (x axis) and insiders (y axis) for every data type. The proportion of concerned subjects for a data type **A** was estimated as a fraction of the number of subjects that were either concerned or highly concerned with unauthorized access to the total number of subjects that stored the data type **A** on their smartphones. Note, that even though personal passwords also showed a statistically significant difference, we ignored it due to small absolute difference. This plot shows that users' concerns with regards to both adversaries are highly correlated (r=0.91), which suggest that both types of adversaries are worth considering. In other words, users are concerned about insiders, and it is comparable to their concerns about strangers.

In order to validate these results, we also asked the subjects to rank each data type for each type of the adversary.

Statistical analysis revealed 11 statistically significant differences (WSRT, $p < 0.05$). These differences, however, had small absolute values, thus, could be ignored. For the purpose of presentation, we averaged the ranks and plotted them on Figure 1b. Similar to the ratings, the correlation between ranks of users concerns for both adversary types was high (r=0.96).

From these results, we conclude that users are concerned about insiders gaining unauthorized access to their data or applications. Furthermore, the level of their concerns about insider access is comparable to level of their concerns about stranger access. This implies that both system and usability practitioners should evaluate their proposals for DPS against insiders as well as strangers.

Research Question 2 – " Have users experienced unauthorized access of sensitive data, either as victims and/or as adversaries?"

In study 2, we asked the participants to select the kinds of previous "negative" experiences they had undergone. A summary is provided in Table 2. We found that half of subjects

Description of the experience	n/%
E1 - I have left my mobile phone at some place, but recovered it later (e.g., at my friends' place, in a restaurant, at parents house, at school, etc.)	363/50
E2 - I have broken my mobile phone before, so that it was not usable	335/46
E3 - I have lost my mobile phone before and did not find it	165/23
E4 - Someone used my mobile phone without my permission with intention to use its functionality (phone call, browsing the Internet, etc.)	100/14
E5 - I used someone's mobile phone without owner's permission for some functions (phone call, browsing the Internet, etc.)	102/14
E6 - Someone used my mobile phone without my permission with intention to look at some of my data	89/12
E7 - I used someone's mobile phone without owner's permission to look into his/her data	66/9

Table 2. The distribution of the previous "negative" experience of the participants $(N = 724)$**.**

had left their phones in some place. In such cases, their smartphones became an easy target, since an insider would have had plenty of time to go through data. 12% of the subjects had found that someone was accessing sensitive data on their phones without their permission. Furthermore, 9% had sneaked into someone else's phone, to access data without the owner's permission. These results provide empirical evidence that unauthorized access to smartphone data and functionality happen in the daily life of smartphone users.

Similar results were collected during the interviews. In particular ten (out of 22) subjects had negative experiences with their smartphones, such as complete damage (to an unusable state) or unrecoverable loss. Additionally, they encountered unauthorized access while at home or at work, or committed such access themselves. For example, one participant (female, 19-24, student) used to live in a shared accommodation with other students, and she found that while she was sleeping, her roommates used her phone to look at pictures and make expensive phone calls. Shortly after the discovery, she decided to lock the phone with a PIN. Another subject (female, 19-24, student) stated that she always locked her phone at home. She justified this by necessity to hide SMS messages from her parents and brother. Finally, yet another participant (female 19-24, student) who had found someone's phone, stated that she looked through all the pictures stored on the device. When asked why she did this, she replied, "Wouldn't you do the same?" Interestingly, subjects who committed unauthorized access stated that they tried to hide the traces of intrusion, by taking phone while the owner does not see and returning it back. This supports our definition of an insider in our adversarial model, i.e., his ability to hide the traces of an attack.

We performed a logistic regression analysis in order to identify groups of smartphone users that had higher chances to be a victim of an authorized access. Logistic regression is best suited for models with binomial independent variables– in this case, those who have or do not have an experience. In this analysis, we only analyzed the experience related to an unauthorized access (i.e., E4-E7). We built a model for each experience separately, four models in total. If a subject had such an experience, then we coded it as 1, otherwise 0. As independent variables, we considered the following values: *A -Age*, *G - Gender*, and *L - Lock Use*. For binomial independent variables (Gender, Lock Use), we used a bipolar representation (-1,1). Equation 1 shows the form of the model we investigated, where E_x stands for one the experiences from E4-E7.

Experience	a_0	a_1	p	RD	AIC	R^2
E4	-2.95	-0.53	< 0.001	546	550	0.09
E5	-2.90	-0.51	< 0.001	554	558	0.08
E6	-2.70	-0.36	< 0.001	521	525	0.05
E7	-3.13	-0.52	< 0.001	425	429	0.05

Table 3. Parameters of logistic regression models, where a_0 is intercept, a_1 is the coefficient in front of Age variable, p is the biggest p-value for both a_0 and a_1, RD is the residual deviance, AIC is Akaike Information Criterion, and R^2 is Nagelkerke R-squared.

$$E_x = \frac{e^{a_0+a_1G+a_2L+a_3A+a_4GL+a_5GA+a_6LA+a_7GLA}}{1+e^{a_0+a_1G+a_2L+a_3A+a_4GL+a_5GA+a_6LA+a_7GLA}} \quad (1)$$

Logistic regression analysis revealed that, for all four models, all interaction effects were not statistically significant ($p > 0.174$), thus could be removed from the model. Furthermore, Gender and Lock Use also showed statistically insignificant prediction power on the experience ($p > 0.185$). That is why we simplified our models to the form shown in Equation 2. The parameters of the models are shown in Table 3.

$$E_x = \frac{e^{a_0+a_1A}}{1+e^{a_0+a_1A}} \quad (2)$$

First, the logistic regression analysis revealed that our models did not have strong predictive power since R^2 values were low. However, the coefficients of intercept and age showed a statistically significant difference from zero. Negative values of the intercept and the coefficient for age showed that the younger subjects have higher chances of experiencing an unauthorized access. This is also depicted in Figure 2, where a larger ratio of younger subjects had experienced E4-E7. This might be attributed to various factors. For instance, social norms might not be strong in this age group; younger smartphone users might tend to share their devices more frequently; or young students often share accommodation with others while attending college or university.

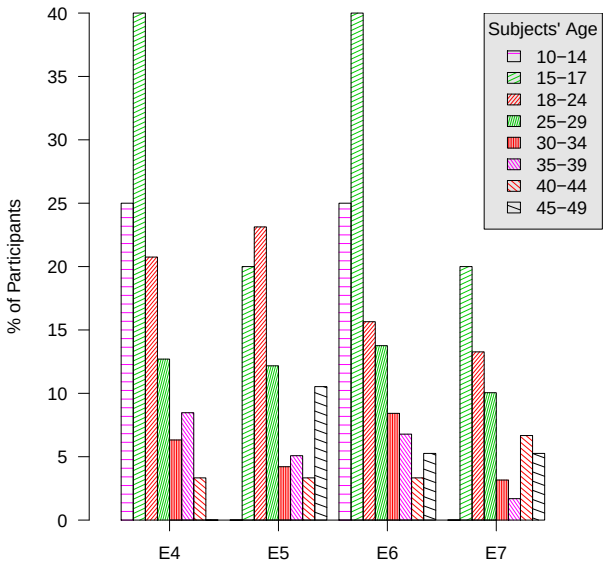

Figure 2. Distribution of the experiences E4-E7 (meaning for these labels are provided in table 2) over participants' age groups. We removed all the subjects that were younger than 10 and those that were 50 or older for clarity purposes.

From the descriptive statistic results and the logistic regression analysis we conclude that users do experience unauthorized access to the data or functionality on their smartphones, and unauthorized access is more common among younger smartphone users. These results also suggest that further research is needed in order to understand which factors increase the chances of attacks, and how these factors impact security and usability of DPS.

LIMITATIONS

The design of this study has several limitations. First, all data reported in this study are subjective, and participants' biases might be present in the results. We tried to reduce any bias by avoiding security terms and jargon in the questionnaire. We also conducted several pilot studies to improve the clarity of the questions.

Second, our study is limited to smartphone owners. The results in Table 2 should be considered as a lower bound. In particular, users might be hesitant to self-report that they accessed someone's phone without permission, or users might not know that their smartphones were accessed by someone else.

Finally, the participants of this study were recruited through the MTurk platform. We tried alternative recruiting methods; unfortunately, these proved to be less effective. We do not consider this as a major threat to the validity of our results, since we found that the demographics of the participants were similar to a representative sample of the US smartphone users.

THREAT MODEL

In this section, we present our threat model. We base this model on proposal by Shi et al. [28], and we add new capabilities and objectives of an adversary to it. These new capabilities and objectives enable data protection systems designers to carry out a better evaluation of theirs proposals against a stronger adversary.

Unauthorized access to data can be gained by **a stranger**, or **an insider**. While acknowledging that there are various types of insiders, we consider all insiders to possess the same capabilities. Strangers and insiders differ in capabilities, resources, and objectives. In what follows, we discuss our threat model from the following perspectives: *assets*, *risks*, *risk agents*, *agents' capabilities*, and *objectives*. Unauthorized access can be performed by strangers or/and insiders, who are considered as *risk agents*. Risk agents might consider gaining access to assets on the smartphone as their objective.

Assets on a smartphone could include sensitive data items and/or sensitive functionality. Note, that not all sensitive data are necessary confidential. We consider the definition of the term "sensitive" to be broader than confidential, i.e., confidential data is a subset of sensitive data. We refer our readers to Figure 1 for a sensitivity comparison of various data types. If an adversary accesses a smartphone, there is a *risk* of sensitive data disclosure. Similarly, sensitive functionality might be *misused* (e.g., for identity theft). For instance, an adversary might try to pass herself as the owner of the device.

The following list includes possible *objectives* of an adversary:

- **O1** - Resale/ransom value of the smartphone - sell the device for profit or return it for ransom;
- **O2** - Keep the smartphone for personal use - get the device for free and use it for herself,
- **O3** - Read sensitive data - get an unauthorized access to sensitive data,
- **O4.1** - Use sensitive data/functionality - get an unauthorized access to sensitive data/functionality and use it for profit (e.g., identity theft),
- **O4.2** - Use sensitive data/functionality - get an unauthorized access to sensitive data/functionality and use it for non-profit reasons (e.g., pranks),
- **O5** - Hide traces of unauthorized activities - hide all traces of the unauthorized access.

There are two possible scenarios for a stranger to get hold of a smartphone. In the first scenario, an *active stranger* can gain possession of the device by stealing it. In this case, we assume that such an adversary would be interested in either objectives **O1** or **O2**. In addition, he might be also interested in objectives **O3** and **O4**. Such an adversary would not be interested in **O5**, if he has no intention of giving the device back to the user. In the second scenario, a passive stranger gets the device by accident, i.e., finds it. We assume that such an adversary would want to get the device back to the owner, thus would be only interested in **O3** and **O5**. We ignore such cases when a passive adversary has similar objectives as an active one, since the active adversary is stronger than passive one; thus, consideration of the active adversary is sufficient.

There are two possible scenarios with an insider. First, an insider might be *conservative* such that he seeks only access to data and/or functionality, but wants to avoid detection and possible social complications that might follow. Such an adversary is highly interested in **O3**, **O4.2**, and **O5**. He would not be interested in **O1**, **O2**, or **O4.1**. Second, an insider might be *extreme*, i.e., they are also interested in **O1** (thus no **O5**) and **O4.1**. He would not be interested in **O2**, since it would be difficult to use a stolen device from a peer without a great risk of being discovered.

There are several capabilities an adversary needs in order to accomplish the aforementioned objectives:

- **C1** - Gain physical access to the device (either permanent, i.e., by stealing it, or temporary, by taking it without permission with the intent to put it back),
- **C2** - Hide traces - by placing the device back where it was and not leaving traces in the system of recent activities (e.g., by marking read email as unread, by deleting sent SMS messages, etc.),
- **C3** - Observe the victim using the smartphone multiple times,
- **C4** - Be in close proximity to a victim,
- **C5** - Observe the authentication secret.

Both an insider and a stranger have **C1, C2**, however, for a stranger, sometimes it is harder to accomplish **C2**, e.g., to get the phone back to the owner. An insider is able to observe the owner and be in close proximity (**C3** and **C4**), and the stranger is not. Based on this, it is much harder, if not impossible, for a stranger to have multiple observation for shoulder surfing attacks on (PIN or Draw-a-Secret) authentication methods, while an insider is able to carry out multiple observations (**C5**).

Our assumption that an insider is capable of capturing the authentication secret for PIN and Draw-a-Secret (DAS) methods is based on the related work. In particular, Raguram et al. [26] evaluated iSpy system, which is able to reconstruct users' input on smartphones from recorded videos, which would enable an adversary to capture authentication secrets. Zakaria et al. [32] and Dunphy et al. [14] investigated the possibility of increasing eavesdropping resistance for authentication method. The authors found that even more complex implementations of the DAS method, with a higher entropy than the version deployed today, are not resistant to shoulder surfing. Similarly, Dunphy et al. report that, on average, it took seven observations for an attacker to capture the authentication secret for a picture-based authentication method. Finally, De Luca et al. [11] extended the DAS method with a modality that considered how users apply pressure during authentication process. Ideally, such a system would resist adversaries, who cannot easily observe the way a secret is entered. Meng et al. [31], however, showed that users are capable of learning keystroke dynamics, thus, the question of whether it is safe to assume that an adversary cannot learn and observe how to apply pressure is still open. The capabilities and objectives of different adversaries types are summarized in Table 4.

Adversary Type	**Objectives**	**Capabilities**	**Return**
AS - Active stranger	O1-O4	C1-C2	No
PS - Passive stranger	O3, O4.2, O5	C1-C2	Yes
CI - Conservative insider	O3, O4.2, O5	C1-C5	Yes
EI - Extreme insider	O1, O3-O4	C1-C5	No

Table 4. The summary of objectives and capabilities for different types of adversaries. Column *Objectives* contains the types of objectives an adversary might be interested in. Column *Capabilities* shows the capabilities an adversary has. Column *Return* shows whether an adversary is interested in returning the smartphone to the owner.

For the HCI community, capabilities C3 and C4 are crucial for a proper evaluation of novel authentication methods. That is why we argue that novel authentication methods for smartphones should be evaluated against our adversarial model, with emphasis on shoulder surfing attacks. In addition, capability C5 opens a new direction for research in usable audit systems for smartphones. Since smartphone are used by all kinds of users, with various backgrounds and capabilities, the usability of such audit system is of high importance.

DISCUSSION

We now summarize the findings of this study and discuss their implications for the field of data security on smartphones in the presence of an insider.

First, the results of our work show that users are concerned about insiders, and smartphone users experience unauthorized access of their data by insiders. These results strongly suggest that research that aims to improve the security of data in lost or stolen smartphones have to evaluate their proposals against the insider threat. This requires an understanding of insiders' capabilities, which we define in our adversarial model.

Second, some researchers have made an assumption that some locations are safer than the others (e.g., Riva et al. or Hayashi et al. [20, 27]). In particular, they treat home, work, and school as safe environments. In these locations, the authors usually propose to either disable authentication or to make it simpler to increase usability. The results of our work suggest the opposite: these locations might be safe if you only consider strangers, but they are not safe once you consider insiders.

Third, the results of this work show that younger demographic groups have higher risk of experiencing unauthorized access. It is, however, still not clear which factors increase or decrease the probability of unauthorized access. Further investigation of these factors and their impact on security and usability of DPS is needed.

Fourth, we found that 95% of those users who lock their smartphone used weak authentication methods that are not resistant to eavesdropping attacks. It is obvious that further research is needed in secure, yet usable, authentication methods

for smartphones. New proposals, however, have to consider a stronger adversarial model, which we present in this work. In particular, new proposals have to be evaluated against attacks where an adversary either learns the authentication secret or the way in which the secret is entered, i.e., capabilities C3 and C4.

Finally, the reduction of an insider's ability to hide his traces is also a possible direction for future research. Modern smartphones do not provide users with a way to identify if the smartphone has been accessed by someone. If an insider unlocks the phone, he can view all the pictures without leaving an audit trail. Even though it seems that it is different with messages (e.g., SMS or email), one can always revert the state of such messages from "read" to "unread" (either through standard functionality or through third party applications). We argue that the HCI community should push forward research in usable audit systems for smartphones. Such systems can benefit significantly from various sensors that are available today on modern smartphones.

CONCLUSION

The results of our study suggest that insiders are an important threat, and they impact smartphone users today. In particular, we found that users are concerned with insiders, and more than 12% of them experienced unauthorized access of their data or functionality on their smartphone. Furthermore, more than 9% of our subjects stated that they have accessed someone's smartphone without the owner's permission.

We highlight direction for future research of usable DPS, such as audit system. These systems might not only improve accountability in smartphones, but also serve as a deterrence factor for an adversary. Additional studies, however, are needed in order to understand how effective can these system be, given highly constrained user interface of modern smartphones.

The results of the user studies revealed that almost all subjects (95%) who locked their smartphones, used PIN or DAS authentication methods. According to the recent research [14, 19], these methods are not resistant to eavesdropping, especially when users are distracted by many other factors [12]. Thus, the adequacy of existing DPS against shoulder surfing attacks mounted by insider threats is questionable.

Finally, in this paper we presented an adversarial model, which defines the objectives and the capabilities of strangers and insiders. We argue that such stronger adversarial model should be used during the design and evaluation stages of novel DPS and authentication methods for smartphones.

REFERENCES

1. Lost and found: The challenges of finding your lost or stolen phone. **http://www.mylookout.com/**. last accessed August 18, 2011.
2. Webview. **http://www.symantec.com/about/news/resources/press_kits/detail.jsp?pkid=symantec-smartphone-honey-stick-project**, 2012.
3. Number of Smartphones Around the World Top 1 Billion - Projected to Double by 2015. **http://finance.yahoo.com/news/number-smartphones-around-world-top-122000896.html**, 2013. Accessed March 12, 2013.
4. Banks, L. Mobile devices pose security dilemma for CIOs. **http://www.cio.com.au/article/346474/mobile_devices_pose_security_dilemma_cios/**, 2010.
5. Barr, K., Bungale, P., Deasy, S., Gyuris, V., Hung, P., Newell, C., Tuch, H., and Zoppis, B. The vmware mobile virtualization platform: is that a hypervisor in your pocket? *SIGOPS Oper. Syst. Rev. 44* (December 2010), 124–135.
6. Ben-Asher, N., Kirschnick, N., Sieger, H., Meyer, J., Ben-Oved, A., and Möller, S. On the need for different security methods on mobile phones. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*, MobileHCI '11, ACM (New York, NY, USA, 2011), 465–473.
7. Bugiel, S., Davi, L., Dmitrienko, A., Heuser, S., Sadeghi, A.-R., and Shastry, B. Practical and lightweight domain isolation on android. In *Proceedings of the 1st ACM workshop on Security and privacy in smartphones and mobile devices*, SPSM '11, ACM (New York, NY, USA, 2011), 51–62.
8. Chaugule, A., Xu, Z., and Zhu, S. A specification based intrusion detection framework for mobile phones. In *Proceedings of the 9th international conference on Applied cryptography and network security*, ACNS'11, Springer-Verlag (Berlin, Heidelberg, 2011), 19–37.
9. Chin, E., Felt, A. P., Sekar, V., and Wagner, D. Measuring user confidence in smartphone security and privacy. In *Proceedings of the Eighth Symposium on Usable Privacy and Security*, SOUPS '12, ACM (New York, NY, USA, 2012), 1:1–1:16.
10. Conti, M., Nguyen, V. T. N., and Crispo, B. Crepe: context-related policy enforcement for android. In *Proceedings of the 13th international conference on Information security*, ISC'10, Springer-Verlag (Berlin, Heidelberg, 2011), 331–345.
11. De Luca, A., Hang, A., Brudy, F., Lindner, C., and Hussmann, H. Touch me once and i know it's you!: implicit authentication based on touch screen patterns. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems*, CHI '12, ACM (New York, NY, USA, 2012), 987–996.
12. De Luca, A., Langheinrich, M., and Hussmann, H. Towards understanding atm security: a field study of real world atm use. In *Proceedings of the Sixth Symposium on Usable Privacy and Security*, SOUPS '10, ACM (New York, NY, USA, 2010), 16:1–16:10.
13. Dorflinger, T., Voth, A., Kramer, J., and Fromm, R. "My Smartphone is a Safe!" - The User's Point of View Regarding Novel Authentication Methods and Gradual

Security Levels on Smartphones. In *SECRYPT 2010 - Proceedings of the International Conference on Security and Cryptography, Athens, Greece, July 26-28, 2010, SECRYPT is part of ICETE - The International Joint Conference on e-Business and Telecommunications*, SciTePress (2010), 155–164.

14. Dunphy, P., Heiner, A. P., and Asokan, N. A closer look at recognition-based graphical passwords on mobile devices. In *Proceedings of the Sixth Symposium on Usable Privacy and Security*, SOUPS '10, ACM (New York, NY, USA, 2010), 3:1–3:12.

15. Enck, W., Gilbert, P., Chun, B.-G., Cox, L. P., Jung, J., McDaniel, P., and Sheth, A. N. Taintdroid: an information-flow tracking system for realtime privacy monitoring on smartphones. In *Proceedings of the 9th USENIX conference on Operating systems design and implementation*, OSDI'10, USENIX Association (Berkeley, CA, USA, 2010), 1–6.

16. Enck, W., Ongtang, M., and McDaniel, P. On lightweight mobile phone application certification. In *Proceedings of the 16th ACM conference on Computer and communications security*, CCS '09, ACM (New York, NY, USA, 2009), 235–245.

17. Farago, P. iOS and Android Adoption Explodes Internationally. **http://blog.flurry.com/bid/88867/iOS-and-Android-Adoption-Explodes-Internationally**. Accessed January 15, 2013.

18. Glen, K. iOS 5.1 Reaches 61% Adoption in Just 15 Days. **http://www.mactrast.com/2012/03/ios-5-1-reaches-61-adoption-in-just-15-days/**, 2012. Accessed July 18, 2012.

19. Hayashi, E., Hong, J., and Christin, N. Security through a different kind of obscurity: evaluating distortion in graphical authentication schemes. In *Proceedings of the 2011 annual conference on Human factors in computing systems*, CHI '11, ACM (New York, NY, USA, 2011), 2055–2064.

20. Hayashi, E., Riva, O., Strauss, K., Brush, A. J. B., and Schechter, S. Goldilocks and the two mobile devices: going beyond all-or-nothing access to a device's applications. In *Proceedings of the Eighth Symposium on Usable Privacy and Security*, SOUPS '12, ACM (New York, NY, USA, 2012), 2:1–2:11.

21. Inc., G. Dashboards — android developers. **http://developer.android.com/about/dashboards/index.html**, 2012. Accessed July 18, 2012.

22. Landman, M. Managing smart phone security risks. In *2010 Information Security Curriculum Development Conference*, InfoSecCD '10, ACM (New York, NY, USA, 2010), 145–155.

23. Muslukhov, I., Boshmaf, Y., Kuo, C., Lester, J., and Beznosov, K. Understanding users' requirements for data protection in smartphones. In *Workshop on Secure Data Management on Smartphones and Mobiles* (2012).

24. Ongtang, M., McLaughlin, S., Enck, W., and McDaniel, P. Semantically rich application-centric security in android. In *Proceedings of the 2009 Annual Computer Security Applications Conference*, ACSAC '09, IEEE Computer Society (Washington, DC, USA, 2009), 340–349.

25. Paolacci, G., Chandler, J., and Ipeirotis, P. G. Running experiments on amazon mechanical turk. *Judgment and Decision Making 5*, 5 (2010), 411–419.

26. Raguram, R., White, A. M., Goswami, D., Monrose, F., and Frahm, J.-M. iSpy: automatic reconstruction of typed input from compromising reflections. In *Proceedings of the 18th ACM conference on Computer and communications security*, CCS '11, ACM (New York, NY, USA, 2011), 527–536.

27. Riva, O., Qin, C., Strauss, K., and Lymberopoulos, D. Progressive authentication: deciding when to authenticate on mobile phones. In *Proceedings of the 21st USENIX Security Symposium*, Usenix Security '12, USENIX Association (Berkeley, CA, USA, 2012), 301–316.

28. Shi, E., Niu, Y., Jakobsson, M., and Chow, R. Implicit authentication through learning user behavior. In *Information Security*, M. Burmester, G. Tsudik, S. Magliveras, and I. Ilic, Eds., vol. 6531 of *Lecture Notes in Computer Science*. Springer Berlin / Heidelberg, 2011, 99–113.

29. Smith, A. Nearly half of american adults are smartphone owners. **http://pewinternet.org/Reports/2012/Smartphone-Update-2012.aspx**. Accessed March 5, 2012.

30. Security Threat Report 2013. **http://www.sophos.com/en-us/security-news-trends/reports/security-threat-report.aspx**. Accessed January 15, 2013.

31. Tey Chee Meng, P. G., and Gao, D. I can be you: Questioning the use of keystroke dynamics as biometrics. In *Proceedings of the 20th Annual Network and Distributed System Security Symposium*, NDSS Symposium'13 (San Diego, CA, USA, 2013).

32. Zakaria, N. H., Griffiths, D., Brostoff, S., and Yan, J. Shoulder surfing defence for recall-based graphical passwords. In *Proceedings of the Seventh Symposium on Usable Privacy and Security*, SOUPS '11, ACM (New York, NY, USA, 2011), 6:1–6:12.