The Privacy and Security Behaviors of Smartphone App Developers

Rebecca Balebako, Abigail Marsh, Jialiu Lin, Jason Hong, Lorrie Faith Cranor
Carnegie Mellon University
{balebako, acmarsh, jialiu, jasonhong, lorrie}@cmu.edu

***Abstract*—Smartphone app developers have to make many privacy-related decisions about what data to collect about end-users, and how that data is used. We explore how app developers make decisions about privacy and security. Additionally, we examine whether any privacy and security behaviors are related to characteristics of the app development companies. We conduct a series of interviews with 13 app developers to obtain rich qualitative information about privacy and security decision-making. We use an online survey of 228 app developers to quantify behaviors and test our hypotheses about the relationship between privacy and security behaviors and company characteristics. We find that smaller companies are less likely to demonstrate positive privacy and security behaviors. Additionally, although third-party tools for ads and analytics are pervasive, developers aren't aware of the data collected by these tools. We suggest tools and opportunities to reduce the barriers for app developers to implement privacy and security best practices.**

I. Introduction

Smartphones such as Android and iPhone offer an array of capabilities to users through a broad and extensive selection of apps. App developers can take advantage of the various sensors on the phones, such as GPS, accelerometer, or camera, to provide entertaining or useful services to the user. Mobile devices are typically always with the user and always on, allowing app developers unprecedented access to information about their users. However, along with these capabilities come great privacy and security risks. While research has looked at smartphone users' perceptions and needs for privacy and security, there has been a dearth of work about the perspectives of app developers

Apps are developed by a broad array of companies and individuals. As the space for innovation is huge, and the barrier to entry is low, many small to medium size app development companies have been able to publish apps. Over 200,000 active developers contribute to the Apple store [1]. There is no training or certification process for app development designed to protect the client. Furthermore, app developers may feel pressure to develop quickly and be the first to market. However, in the race to innovate, privacy and security might not be the top priority for time- and resource-constrained app developers.

In this paper, we examine the ways app developers make decisions and the steps they take to protect security and privacy. Through in-depth interviews with 13 developers, we explored the trade-offs app developers make, how they get information when they need it, and barriers to implementing privacy and security best practices. Informed by the results of these interviews, we formulated several hypothesis about the privacy and security behaviors of app developers. We ran an online survey of 228 app developers to examine factors that predict good privacy and security behaviors, such as encrypting data and providing privacy policies. Our two-step research process is similar to that used in other work examining human subjects' motivations [2].

We first begin by discussing previous work on smartphones and privacy. Then, we describe the interviews and the themes that emerged. In the following section, we describe the online survey and the results of testing specific hypotheses about privacy and security. We find that many developers lack awareness about privacy, and identify a number of barriers to improved privacy and security behaviors. These include the lack of resources in smaller companies and the difficulty of understanding third-party collection of user data. We identify where developers seek privacy and security advice, and point to intervention points and improved tools to help developers.

II. Related Work

We first describe the smartphone app ecosystem, including major platforms and how apps are submitted. We also discuss users' perceptions of smartphone privacy and security. We then describe public policy efforts to guide app developers when making privacy and security decisions and previous efforts to inform app developers about privacy and security.

A. App Development Ecosystem

The two most popular smartphone platforms are Apple's iPhone and Google's Android, with Blackberry and Microsoft holding a smaller market share. Apple and Google both have app markets that allow independent developers to distribute or sell apps, which users can download from their devices. This has allowed many independent developers to sell smartphone software directly to users, and has resulted in a huge variety of apps, with over 800,000 apps on each of the iOS and Android platforms as of October 2013 [3].

Previous work has found a relationship between data collection and advertising as a revenue model. The ad-based

USEC '14, 23 February 2014, San Diego, CA, USA

http://dx.doi.org/10.14722/usec.2014.23 006

revenue model, which often relies on targeted ads, is currently popular [4]. Apps may provide ads through third-party code, such as that provided by Flurry[1] or Google AdSense.[2] Targeted advertising requires collecting information about users, and therefore the targeted advertising revenue model may require more permissions and therefore be more privacy-invasive [5], [6]. Apps may also include third-party code for analytics, whose primary goal is to collect information about the users' interactions with the app.

Some previous work has examined app developer security behaviors, such as that by Egele et al. [7] and Fahl et al. [8], which found the significant portions of apps with security failures or substandard implementations of security code. Throughout our work, we explore app developers' perceptions of their work, including self-reported intentions.

B. User Concerns about Privacy and Security

A wealth of previous work has examined users' perceptions and desires for smartphone privacy and security [9]–[13]. Users are often surprised by what permissions are requested by apps [14], the frequency of data collection, and the data recipients [15]. Furthermore, they often do not understand existing privacy notices, particularly in Android phones [16], [17]. While users are concerned about privacy and security, they are neither informed nor empowered to protect themselves. Therefore, the decisions made by app developers have great impact.

Previous work has examined users' reactions to privacy policies. While privacy policies offer the illusion of notice to users, the reality is that the required time [18] and reading level [19], and vague language [20] pose significant usability barriers. Our work indicates that app developers have similar troubles with privacy policies.

C. Public Policy and Tools

There have been several efforts to educate app developers about privacy and security. We reviewed five privacy guidelines for app developers: three were published by government agencies in Australia [21], Canada [22], and California [23]; one by an industry consortium in Europe [24]; and one by consumer privacy advocacy groups [25]. These guidelines typically offered clear and readable advice and avoided "legalese." While they were lengthy (14-32 pages), some offered privacy and security checklists for developers. These guidelines often suggest that privacy policies can help developers think through their data collection practices in addition to notifying users.

There were five recommendations made by all of the above-cited guidelines, which we paraphrase as follows:

1) Someone must be responsible for privacy.
2) The app should have a clear and easy to find privacy policy.
3) The app should encrypt data during transmission.
4) The app should encrypt data it stores.
5) The app should limit data collection to what is needed.

These are the five main privacy and security behaviors we explored quantitatively in our online survey, and we describe them in greater detail in Section V.

[1] www.flurry.com

[2] www.google.com/adsense/

Tools have been developed to help developers practice privacy and security behaviors. Many open-source databases, such as mySQL, allow encryption of stored data. Several free or low cost privacy policy generators[3] exist that allow developers to create a policy by answering questions about their app's behaviors. Our interviews examined whether developers were aware of or used these tools.

III. Interview Method

We conducted semi-structured interviews with 13 smartphone app developers in August and September of 2013. Our research goals were to understand what decisions app developers make that they consider privacy and security related, and to better understand what resources they were aware of to help them make those decisions.

Interviewees represented a variety of app types and company sizes, as shown in Table I. We asked "What type of service does your app provide," and the choices were based on a taxonomy developed by Hyrynsalmi et. al [26]. Interviews lasted approximately one hour. The interviews were usually conducted remotely, with only one in-person interview. The audio was recorded for transcription, although participants had the option to refuse audio recording, as some said it made them uncomfortable or unlikely to be forthcoming. Interviewees received $20 as compensation. Our interviewees were overwhelmingly male, which is in-line with evidence that 94% of app developers are male [27].

Participant ID	Company Size	Revenue Model	Service	State
P1	10-30	Advertising, Free trial, Subscription	Digital, Physical, Service, Contents	CA
P2	2-9	Advertising, Free trial, Other	Digital, Service, Contents, Advertisement, Personalized information, Other	CA
P3	2-9	Free trial, Other	Digital, Service	PA
P4	2-9	Pay-per-user	Physical, Service	WA
P5	2-9	Free trial	Digital	WA
P6	100+	Subscription	Other	PA
P7	1	None	Contents	TX
P8	10-30	Subscription	Digital, Service	CA
P9	2-9	Other	Service	CA
P10	1	None	Contents	PA
P11	2-9	Advertising, None	Physical, Personalized information, Other	IL
P12	2-9	None	Personalized information	PA
P13	100+	None	Physical	MI

TABLE I. Interview participant mobile app and company demographics.

Service	Examples
Digital	games, MP3, Ebooks
Physical	selling books
Service	e-mail, banking, ticketing
Stock Information	stock prices
Contents	news, weather, entertainment
Personalized information	location information

TABLE II. Service categories based on classifications by Hyrynsalmi et. al. [26].

We recruited participants for interviews through a number of methods, including in-person recruiting at local meetups

[3] freeprivacypolicy.com, generateprivacypolicy.com, appprivacy.net

for smartphone app developers, online postings on sites such as Craigslist and Backpage, and through our social networks. Recruitment text said, "Participate in an interview to understand and improve smartphone app development." Security and privacy were not mentioned in the recruitment to avoid participant bias. We asked interested parties to first fill out a screening survey to see if they qualified. We included two technical questions to determine whether the applicant had credible knowledge of app development. Valid applicants were invited by email to set up an interview time with one of two researchers. We contacted 20 developers, and 13 completed the interview. Five of the invited developers who did not complete the interview failed to respond to the email invitation, and two invitees were unable to find a suitable interview time.

We did not collect identifying information, such as given name or company name, from participants unless it was volunteered. The interviewed developers ranged from 26 to 58 years old, and were from six states. Most worked in groups of 2-9 developers, but company size ranged from 1 to 100+ employees. Most interviewees were programmers, but one was a product manager. Several interviewees played multiple roles in their company, such as CEO, manager, or quality assurance. Their apps represented a variety of business models and services, and were at various stages of maturity. Some apps were not yet released to the app market, and others had already had several versions on the app market.

Questions included, "What, if any, online resources do you use to help make privacy and security decisions?" and "Have you ever decided not to collect certain information from users due to privacy concerns?" While we generally followed a script, we iterated on the script as each interview informed the next. Participants were asked what subjects we should have addressed, which revealed gaps in our questions and allowed us to improve the interviews.

IV. Interview Results

We describe the themes that emerged from our interviews. We discuss how app developers learn about privacy and security, whether they are aware of regulation and third-party data collection, and where they seek advice and resources for privacy and security decisions. We discuss developers' perceptions of privacy policies and the trade-offs that app developers confront when making privacy and security decisions.

A. Education and Advice about Privacy and Security

Only a few of the developers we interviewed had formal training on privacy and security, typically received through corporate training or certification. Other developers rely on online research to find answers to specific questions. They are not accessing the guidelines published by government agencies, and instead are more likely to rely on their social networks, or specialists within their companies for information.

Many participants did not have formal privacy and security training. This suggests that many developers learn about security and privacy when they are confronted with these issues in the course of their work, at which point they may seek out further education. The lack of education on security and privacy available at the introductory levels was not lost on developers. P3 stated, "Most classes in computer science...there isn't much of a focus on security. That could have a very big impact on how this stuff [implementation of secure code] happens." On the other hand, some participants were confident that they were learning what they needed to know, or had a good background. P13 said "I have no formal training with privacy and security, but I feel that I am a journeyman in privacy knowledge, and pretty expert at security knowledge." Similarly, P10 stated that his privacy and security learning, "is pretty much internal knowledge based on my experience in Web."

Some participants discussed receiving formal training from a variety of sources. Certain businesses have specific training or certification requirements. For example Payment Card Industry (PCI) has security standards for handling credit card information. P11 states, "When you work at E-Commerce, they want you to be what they call PCI compliant." In less regulated areas, participants reported education including certifications, previous work experience, and conferences such as the RSA Conference.

When asked about current and upcoming privacy and security regulations, participants showed little knowledge. While a few app developers brought up issues of the government requesting user data as a concern, none were aware of guidelines such as those discussed in Section II. The exceptions were apps that were marketed to children under 13 or used health information; these developers were aware of the privacy laws specifically related to their cases.

Participants were asked to discuss what resources they used when they needed advice on security- and privacy-related decisions. We received a variety of responses, which could be grouped into a number of common themes, including searching online, consulting friends, and seeking legal or specialist advice.

One of the most common responses was that developers simply searched online when they were looking for advice. As P10 put it, "I would Google it, to be honest, and I would look for articles from developers who have focused on building secure systems and kind of start my research there." Developers consulted Hackernews, TreeHouse, StackExchange, Lynda.com, Google, Facebook's Terms of Use, and various smartphone developer forums to search for advice and examples from other developers.

Many developers also consulted their friends and social networks for advice: P7, a professional developer and part-time student, consulted a "Facebook group with... some 300 students," many of whom do mobile development. Others consulted with fellow developers in person, like P5 who said, "I go to a couple meetups, especially if I'm looking for a technical element, or I want to get more into usability." Participants also consulted with contacts who had experience in security or privacy: P10 stated, "I would also talk to my social network, if I knew anyone who has a background in security, about what they would recommend. I fortunately know one or two people."

Lawyers were also consulted when they were available to developers. Some participants worked for companies with dedicated legal staff, such as P13 who stated, "I try to raise [privacy concerns] up to my management level and let them interact with whatever back-end legal that needs to happen.

I try to avoid directly communicating with the lawyers." P12 makes it clear that privacy awareness was the legal division's domain: "Ultimately the legal staff is responsible for making sure that we get the right and accurate information." Generally, the interviews suggest that developers who had access to legal teams seemed to be less personally involved in the understanding of privacy and security regulations.

Some developers relied on terms of service documents provided by the app markets, with P4 stating, "I would expect that those guidelines fall into the realm of what is legally expected in the United States." P8 depended on lawyers to understand regulations that affect app development, leading to less personal knowledge: "The only times we had to change anything, lawyers are on top of it. The reason I didn't bother to know [is that I] depend on a lawyer." As P3 observed, "Unfortunately, I very rarely have time to actually sift through [privacy and security regulations] and try to digest everything that's going on, so I primarily rely on other people to let me know."

B. Security Tools Used More than Privacy Tools

App developers seemed to use and rely on off-the-shelf or third-party tools for security, but did not have as many tools for privacy. The use of third-party tools could also introduce additional privacy concerns, as these tools may collect information that the app developer was unaware of.

Some developers rely on specific tools to help with security. These tools could include encryption built into the database, SSL code built into the platform, or authentication methods such as Facebook authentication. The tools were perceived as being more secure than hand-rolling implementations themselves. For example, P4 discussed the use of Facebook for authentication, "The expectation is that all the crafty security stuff has been handled by them, because I assumed they'd be smart enough to have that locked down, given that they probably hired security people." However, participants noted that tool usage could be a double-edged sword. For example, participants who used Facebook for authentication had access to much of their users' Facebook profile. Developers discussed weighing the advantages of collecting this information in case it might be useful against the privacy concerns of the user.

Very few interviewees used or knew about existing tools specifically for privacy, such as privacy policy generators, or security audits. One interviewee described his experience with a privacy policy generator as being "good enough" for the time, but not able to handle complex cases. Security audits were only considered by one interviewee; he handled health information and was working with businesses that required audits.

Participants also relied on third-party tools for other uses, such as analytics or various other features. Participants seemed generally unaware of the privacy and security practices employed by third-party utilities used in the development of their apps. Many developers had not personally read the terms of service, were unsure if their lawyers or legal departments had done so, and may have even forgotten the names of the ad networks or web traffic analysis companies they had used. P3 described the need for more digestible information, saying, "if either Facebook or Flurry had a privacy policy that was short and concise and condensed into real English rather than legalese, we definitely would have read it."

C. Privacy Policies Are Not Considered Valuable

App developers find creating privacy policies to be a low priority or of low value, believing they only offer legal coverage and may turn off users.

Participants were particularly unconcerned about providing privacy policies. In one interview, P4 said, "I haven't even read [our privacy policy]. I mean, it's just legal stuff that's required, so I just put in there." Both P10 and P11 explicitly stated that they were not concerned, because they worked for small companies, with P10 saying: "I have not heard of any startups or small companies getting into trouble for privacy policies," said one, while P11 noted, "Big companies want to [cover your ass], no one is going to go after a small guy like me. I don't generate enough revenue, so if you do sue me you won't get any money." Other developers stated that they did not collect personally identifiable information, and therefore were less concerned about transparency.

Most participants said that while their privacy policies can be accessed on the app website, they were not directly accessible from within the app. In addition, the type of information collected from users would be difficult to find: P8 admits, "We don't make it very obvious, exactly what data we're collecting. I guess it's kind of in the terms of use or privacy policy or something." Paired with the difficulty of quickly accessing an app's privacy policy, this suggests that users will find it tough to determine how their data is being collected and used by apps [11], [15], [17].

Furthermore, some developers were not convinced that users want privacy policies. P7 said users have "been groomed [into] thinking ... [data] is not private... Because it's all anonymous." They felt that as a result, data collected by their app would not surprise users or cause privacy concerns. P3 described the app developer and user relationship in stark terms: "we have consumers as customers. They either trust us or they don't." Some developers were aware of user concerns, noting, for example, the sensitivity of location data. As P8 put it: "it's definitely important to the user to know that their information is safe with [the app]."

When participants put an effort toward alerting users about information collection, they reported lower user retention. Two interviews reported this concern. "We've gone through pretty great lengths to try to make sure that people know exactly what we're collecting and why we're collecting it," describes P3, "So we end up losing out on some number of users because of warnings....they don't take the time to actually read...so they just sort of see this warning and they're like, oh, it must be something bad."

D. Trade-offs Between Privacy, Security, and Resources

Balancing the need for good security and privacy practices with the cost of actually implementing those practices was a struggle for participants in our interviews. Many discussed privacy and security as being part of the development process but not a top priority, and concerns like monetizing the app or limited resources often trump the desire to follow rigorous privacy and security standards. Some manage to support privacy

and security, like P5, who states: "We are trying to balance where that line [between user concerns and the need to store information] gets drawn. I favor privacy."

P10 tellingly struggles with this trade-off when discussing his company's practice of borrowing from other privacy policies, saying, "I don't see the time it would take to implement that over cutting and pasting someone else's privacy policies.... I don't see the value being such that that's worth it."

When questioned about whether their personal feelings towards privacy affected their development decisions, participants gave mixed responses. Some made strong statements, such as P10 who said, "I personally have very strong feelings about user privacy," and P5 said that as a supporter of privacy rights, he made an effort to collect as little user information as necessary for his app. Even self-described privacy advocates and security experts grappled with implementing privacy and security protection with limited time and resources.

Others, while voicing personal concern about privacy, discussed the need to work with clients' wishes. In reference to the privacy of user data in apps developed for his clients, P11 says, "What they want is what they want." Another developer was very invested in privacy protection, but expressed concern that with the threat of his app being copy-catted, advertising was a safer bet for earning revenue than pay-to-download.

This suggests that developers have to weigh their personal desire to respect privacy against the ability to monetize or sell their app, and in particular, developers who work as part of a larger company or who work on commission may be less free to implement good privacy practices than self-employed developers and those who work for small companies. Furthermore, developers consistently discussed the constraints such as time, effort, and money it would take to implement best privacy and security practices.

The cost of collecting and storing data is perceived as minimal. At the same time, interviewees indicated that the cost of developing the code or policies to delete old data or accounts is not prioritized. This is not a question of tools; many of the same tools that allow users to encrypt data also allow them to delete data. Instead, this is a pervasive belief that data may become useful in the future and is therefore worth the resources required to collect and store.

V. Survey Method

Based on the interview results, we formed two hypotheses about privacy and security behaviors in app development. We hypothesized that company size would be related to privacy and security behaviors and that revenue models would also be related to privacy and security. In order to test these hypothesis quantitatively, we performed an online survey of 228 United States app developers and product managers. The survey gathered relevant demographics about the developers and their companies, and examined how developers make decisions about privacy and security.

Our survey was designed to take less than 30 minutes, and participants were compensated with a $5 Amazon gift card. Participants were recruited though several online forums, such as reddit subgroups, technical Facebook pages, and through six United States cities on backpage.com. To avoid biasing participation, it was not advertised as a security or privacy survey.

We included four knowledge and attention check questions in our surveys to help us eliminate non-developers and invalid responses. Due to our stringent requirements, we discarded 232 results that either did not have valid responses or were outside the United States. We were left with 228 valid responses from within the United States.

The privacy and security behaviors we examined are those that were recommended by all five of the privacy and security guidelines for app developers that are discussed in Section II. We describe the questions used to measure the privacy and security behaviors.

Security Behaviors

- SSL usage: By encrypting data going over the network, app developers can protect users from data snooping on insecure connections. We measured SSL usage with the question, "Do you use SSL when transmitting data?"
- Encrypting collected data: Encrypting data stored by the app, either in a database or on the phone, protects the user in the case of data breaches. We considered two variables: whether data was encrypted either in the database or when stored on the users' phones.

Privacy Behaviors

- Having a Chief Privacy Officer or equivalent: The existence of a CPO or equivalent indicates that the company is paying attention to privacy and has a specialist who is accountable for privacy. We measured this with the question, "Does your company have a Chief Privacy Officer (or equivalent)?"
- Providing a privacy policy: Privacy policies may indicate that the app company has considered their practices and is being transparent to the user. We measured this with the question, "How does your app inform users about what information it collects?" and the response "Privacy policy on website."

We recognize that there are concerns with self-reported data [28]. We present the results as app developers' own conceptions of their work, not as ground truth. Our findings may differ than those of previous research based on scans of the app stores. For example, our questions are on a per-developer basis, and developers may have created more than one app. Our results are for all platforms, and both free and paid apps. Furthermore, our survey was done in August 2013 and may be more recent than published papers' results.

Behavior	percent
Use SSL	83.8%
Encrypt data on phone	59.6%
Encrypt data in database	53.1%
Encrypt everything (all data collected)	57.0%
Revenue from advertising	48.2%
Have CPO or equivalent	78.1%
Privacy Policy on website	57.9%

TABLE III. Percentage of respondents who reported various privacy and security-related behaviors. Participants could select multiple options.

VI. Survey Results

We first present the demographics of our survey participants, including their training in privacy and security, and where they look for advice when making privacy and security decisions. We then discuss the app companies they work for, including size, revenue model, and use of third-party ad and analytics tools. We also present some exploratory work on data collected by app developers, including data types that have not been measured in previous work. We then describe our hypotheses about security and privacy behaviors; that they are correlated to each other, correlated to company size and correlated to revenue, and report our results.

Role	Participants
Programmer or Software Engineer	58%
Product or Project Manager	31%
Tester or Quality Assurance	20%
Manager	12%
CEO or President or Owner	8%
Marketing	4%
Student	4%
User Support	3%
Not currently working/Currently unemployed	1%

TABLE IV. Percentages of participants in different roles. Participants could select multiple options.

A. Participant Demographics

Most of our respondents were programmers, product managers, or quality assurance testers. The average age was 30 years old (range: 18-50 SD = 5.6). We did not collect additional personal demographics such as gender. Participants selected their professional role from a multi-select list. Our recruitment stated specifically that we were looking for app developers or product managers, so it was not surprising that 78% of participants were programmers or software engineers, product managers, or both. Other participants were testers, managers, and CEOs. The role breakdowns are shown in Table IV.

We asked participants to describe their formal privacy and security training. Our results directly contradicted our interviews, in which few people claimed to have formal privacy or security courses. However, most interview participants worked for small companies. In the survey, only 7.3% claimed to have no formal privacy or security training. 62.9% of respondents claimed to have taken a privacy or security training course. Many also stated that they had received corporate training on privacy or security (62.5%) or attended a professional development seminar or workshop (43.5%). App developers in companies of size 31-100 were the mostly likely to receive corporate training, and companies with only one employee were the least likely to receive training.

In order to determine how app developers were making privacy and security decisions, we asked participants from whom they sought advice about privacy and security. This is useful for two reasons: first, it provides some insight into the level of expertise available to developers, and second it may allow better framing of educational campaigns for app developers about privacy and security. Figure 1 shows from whom participants sought advice, based on their company size. The company size significantly affected whether participants sought advice from their social network, security or privacy

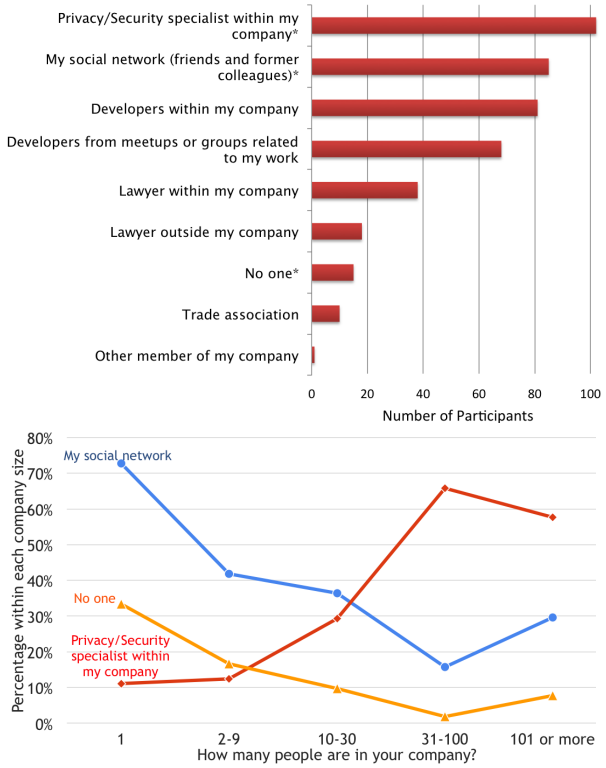

Fig. 1. Survey participants' response to the question "Who, if anyone, do you turn to when you have questions about consumer privacy and security?" Responses significantly different based on size of company are marked with *. The bottom figure shows the 3 significant selections by company size. Developers at small companies rely on their social networks or no one, while developers at larger companies rely on specialists within the company.

exports in their company, or no one (Kruskall-Willis test, $p<.001$). Participants from companies with under 9 employees were more likely to get advice from their social network, or to ask no one. Developers in larger companies (31-100 employees or 100+ employees) were more likely to ask a privacy or security specialist within their company.

B. App Company Characteristics

We discuss the categories of app companies represented by the survey participants. We do not claim that this is a proportionate sample of app development companies in the United States. Instead, we discuss the characteristics to put our other findings into context.

Equal numbers of participants were building or planning to build iOS (142) and Android (142) apps, with much smaller numbers for other platforms (38 for Windows, 10 for Blackberry and 6 for Palm, and 1 other). Over one quarter of participants (63) said they were developing for both Android and iOS. The survey participants represented different size companies and development groups. The percentages of developers in companies sized 1, 2-9, 10-30, 31-100, and 101 or more were 4.9%, 14.8%, 19.7%, 48.4%, and 12.1% respectively. However, the size of app development groups (employees working directly on the app) were typically between 2-30 people.

Participants were asked to categorize their app, using a list that was a combination of Apple iTunes store and Android

Revenue Model	Total	Only Source
Advertising	48.0%	9.4%
Paid Download	44.8%	10.3%
Free trial, upgrade to premium	37.7%	8.1%
Subscription	25.1%	8.5%
Pay-per-use	21.5%	2.2%
In-app purchases	19.7%	4.0%
All of above except Subscription		7.6%
Advertising and Paid Download		8.5%
Hosting	3.6%	.9%
Other	3.6%	3.1%
None	2.2%	2.2%

TABLE V. REVENUE MODELS OF RESPONDENTS. RESPONDENTS MAY HAVE CHOSEN MULTIPLE RESPONSES; THE MIDDLE COLUMN REPRESENTS ALL PARTICIPANTS WHO SELECTED THAT MODEL, AND THE RIGHT COLUMN SHOWS HOW MANY PARTICIPANTS SELECTED ONLY THAT MODEL. THE APP REVENUE MODELS USED ARE BASED ON LEEM ET. AL [29].

Play store categories. All categories were represented, with Games (17.7%), Entertainment (12.5%), and Finance (10.8%) appearing most frequently.

Ad or Analytics company	Survey
Google analytics	82.1%
Google ads	64.1%
Flurry analytics	16.6%
Medialets	11.2%
AdWhirl	8.1%
AdMob	13.9%
AirPush	14.8%
Amazon ads	43.9%
No ads	13.5%
No analytics	12.6%

TABLE VI. PERCENTAGE OF RESPONDENTS WHO REPORTED USING VARIOUS ANALYTICS COMPANIES. PARTICIPANTS COULD SELECT MULTIPLE OPTIONS. ONLY LIBRARIES WITH 10% OR MORE OF RESPONDENTS ARE SHOWN.

Apps may collect data for their own use, but interviewees also indicated data is collected for secondary uses such as advertising or analytics. Our interviews indicated that app developers were not always aware of the data collection of the third-party API's or toolkits they were using. Table VII shows respondent's knowledge of third-party data collection practices. Just over one-third of app developers claimed that they knew exactly what data is collected by third-party tools. These responses may represent more of the developers' self-perception than reality. For example, of the developers who claimed they did not use third-party tools, the majority answered separate questions about using third-party tools differently: 70% said they used at least one ad company, and 87% used an analytics company. This suggests confusion either about the question (different definitions of "third-party tools") or their own apps' behaviors.

Seventy percent of respondents stated that their apps were already earning revenue. As seen in Table V, a variety of revenue models and types of apps were also represented, with advertising and paid downloads being the two most popular options. 52% of participants reported relying on more than one revenue model; the two most popular combinations were advertising combined with paid download and all revenue models except subscription and hosting. The more revenue models selected, the more likely they were already earning revenue (χ^2 tests $p<0.001$). The distribution of revenue models was similar for both iOS and Android. Due to the nature of our questions, these results could represent the revenue model used by an app developer across multiple apps that they have developed, as opposed to the models used within one app.

Response	percent
My app doesn't use third-party tools	41.7%
I know exactly what kinds of data the third-party tools are collecting	35.9%
I have some ideas about third-party data collection but don't know for sure	22.0%
I don't know	.04%

TABLE VII. RESPONSES TO "HOW FAMILIAR ARE YOU WITH THE TYPES OF DATA COLLECTED BY THIRD-PARTY TOOLS?"

Overall, most app developers (87.4%) used at least one analytics company, with one in five using two or more analytics companies. Table VI shows which companies were used by app developers. Most apps also used an advertising company: 86.5% selected one or more advertising companies in use, using on average 1.78 ad companies (SD=1.33). Interestingly, app developers were likely to use an ad company regardless of whether they relied on advertising for revenue. Of app developers who did not select advertising as a revenue source, 82% still resorting to using at least one advertising company. We speculate that app developers may be including advertising API's without earning money from ads; however this merits further exploration.

In our survey, 41.7% of developers self-report that they do not use a third-party tool. It is important to understand app developers' self-perception, as it will likely influence their need to consider third-party tools' data collection when creating privacy policies or handling data. If developers are not aware of or fail to consider some libraries, they will not report on their behavior when making privacy decisions. Our 2012 scan of free Android apps indicates that 50.2% of free Android apps did not use ads, analytics, social networks, and payment APIs, which is higher than our survey findings suggest [6]. We find that 36.3% of developers reported using exactly one ad library.

C. Collection of Sensitive Data

As we did not discuss the collection of sensitive data in our interviews, we did not formulate specific hypotheses to test. Therefore, we show the results of some exploratory analysis. Table VIII shows which data the app collected or stored. Due to an error with the survey, 5 participants did not answer this question. They were removed from the analysis of this question.

We asked about data that may be privacy or security sensitive. Several data items corresponded to Android or iOS permissions and warnings (such as location), but other data can be collected without warning the user. This includes which apps are installed, or sensor data from accelerometers. The user would only know about this data collection if it were included in a complete privacy policy. Other data that don't trigger permission notifications are credit card information or password; these are input by the user but require that the app developer handle them securely.

An average of 5.5 out of the 10 sensitive variables we asked about were collected. Based on our interviews, we

were not surprised that most apps did not collect or store users' passwords or credit card information. Instead, apps that need this information may often rely on third-parties such as Facebook to do authentication or to handle credit card information. Unsurprisingly, apps collected information pertinent to their app, such as level attained in a game. It is startling that three quarters of app developers collected which other apps are installed on the user's device. Apps may do this to explicitly collaborate with other apps or services, such as a todo list app accessing a calendar app. However, information about installed apps can have privacy implications, such as family or health status if related apps are installed.

Data Type	Collect or Store (%)
Parameters specific to my app	83.9%
Which apps are installed	73.9%
Location	71.6%
Advertising ID	70.6%
Sensor information not location-related	63.0%
Phone ID	54.5%
Contacts	54.0%
Phone Number	44.1%
Password	35.5%
Credit card information	30.3%

TABLE VIII. PERCENTAGES OF RESPONDENTS WHO COLLECTED OR STORED SELECTED DATA.

The category of app also significantly affected the amount of data collected (ANOVA p=.007). Of the categories with 10 or more responses, finance used the most sensitive variables on average (μ=6.36) while entertainment collected the least (μ=4.73). Only 20% of respondents with a finance app had advertising revenue, while 57% of entertainment apps had advertising revenue.

Leontiadis et al. found that free apps required more permissions than pay-to-download apps [4]. Our findings support this. We find statistical differences in the amount of data collected by revenue (ANOVA, p<0.001), and find that the amount of data used by developers with paid-download revenue models only (μ=3.78, SD 3.03) is significantly different from the amount of data collected by advertising-only revenue models (μ=6.48, SD 2.40) (ANOVA multiple comparison, p=0.013 with Bonferroni correction).

D. Hypothesis Testing and Results

In this section, we describe our hypotheses about privacy and security behaviors and the results of testing each hypothesis. Table III summarizes the percentages of respondents who claimed to engage in the each privacy and security behavior. Table VIII summarizes the number of respondents who collected or stored the data types we examined.

1) Hypotheses 1: Behaviors are correlated: First, we hypothesized that security and privacy behaviors would be positively correlated, and that there would be developers who were generally privacy and security concerned and demonstrated all or most behaviors, while others would not display any such behaviors. Our hypothesis is mostly supported; all behaviors are significantly and positively correlated at the p=.05 level except Privacy Policy and SSL, as shown in Table IX.

H1: *Security and privacy protective behaviors are correlated.*

	CPO		Encrypt Everything		Privacy Policy	
	ϕ	p	ϕ	p	ϕ	p
Encrypt Everything	·272∗	<·001				
Privacy Policy	·159∗	·018	·228∗	·001		
SSL	·257∗	·001	·217∗	·005	·157	·063

TABLE IX. CORRELATIONS BETWEEN THE SECURITY AND PRIVACY BEHAVIORS. THE PHI COEFFICIENTS (ϕ) INDICATE THAT THE BEHAVIORS ARE GENERALLY POSITIVELY BUT WEAKLY CORRELATED. * INDICATES SIGNIFICANT CORRELATION AT THE P=.05 LEVEL.

For hypothesis H2 and H3 we ran eight χ^2 tests separately. We conservatively correct the standard p-value of .05 with Bonferroni correction, and use a significance level of 0.006 (0.05 divided by the number of tests).

2) Hypotheses 2: Company size: We are aware that startups or app development companies with small teams and little investment may not have the resources, in terms of time or money, to invest in privacy and security. Therefore, we suspected that small companies may be less likely to engage in the privacy and security behaviors that require additional employees (a CPO), additional time (creating a privacy policy), or additional resources. For example, encryption may require more equipment or software. Using SSL may require additional developer time or experience.

H2a: *Company size correlates to having a CPO.*
H2b: *Company size correlates to having a privacy policy.*
H2c: *Company size correlates with encrypting everything.*
H2d: *Company size correlates with using SSL.*

We found that the size of a company does help determine whether they have a CPO (χ^2 test p<0.001), whether they have a privacy policy (χ^2 tests p=.002), and whether they encrypt everything (χ^2 tests p<0.001). However, the company size was not correlated with SSL using the conservative corrected significance level (χ^2 tests p=.009). As one respondent wrote in an open-text field, "We are a small, two-person shop. Although we don't have CxO positions, we do understand the need to protect the privacy of our users. Our app embeds a privacy statement in an easily identifiable location."

The percentages of companies engaging in the above privacy and security behaviors grows as the company size grows, up to the 31-100 employee companies. For example, all of the respondents with company sizes of 1 said they did not have a CPO or equivalent, while only 58.8% of respondents in companies from 2-9 had someone responsible for privacy, compared to 89.6% and 92.6% of companies size 10-30 and 31-100 respectively. This is shown visually in Figure 2, and is similar for the other privacy and security behaviors. However, this trend of improved privacy and security practices does not hold for company sizes greater than 100. We speculate that app developers in larger companies may not be as aware of all their company's practices.

3) Hypotheses 3: Revenue model: We were curious about the impact of the revenue model on privacy and security behaviors, and hypothesized that certain revenue models, such as advertising, were less likely to show privacy and security behaviors

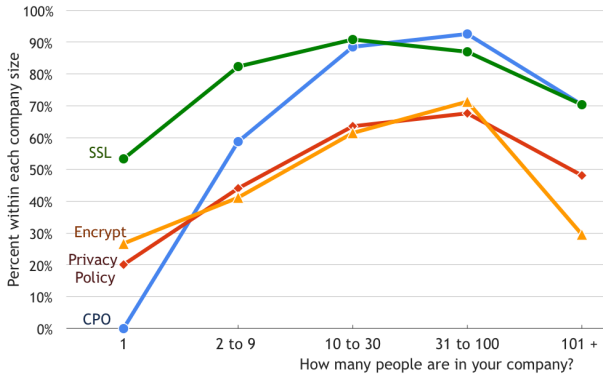

Fig. 2. The size of the company is related to whether or not the company has privacy and security behaviors. Companies with 31-100 employees are the most likely to engage in these behaviors.

H3a: *Revenue model is correlated to having a CPO.*
H3b: *Revenue model is correlated to having a privacy policy.*
H3c: *Revenue model is correlated with encrypting everything.*
H3d: *Revenue model is correlated with using SSL.*

Since 47 unique combinations of revenue models were reported, we examine the most common models and combinations, which are shown in Table V. All other combinations (with fewer than 10 responses) were combined into an "other" category. At our conservatively corrected p-value, none of the results were significant (CPO p=.035, encrypt p=.029, SSL p=.037, privacy policy p=.019). However, we note a few interesting cases. An advertising revenue model indicates low adoption of privacy policy, but is average on the other measures. It is disconcerting that in-app purchase, which might be transmitting payment information, have the lowest adoption of SSL. However, we note that all 17 of the developers who used every model except subscription also claimed to implement all the privacy and security sensitive behaviors. The only common feature we found across all 17 of these developers is that they all received corporate privacy and security training as well as college classes.

VII. Discussion

Our results indicate that many developers lack awareness of privacy measures, and make decisions in ad hoc manner. While most developers claimed to be using SSL and to have a CPO or equivalent, only slightly over half of our survey participants claimed to employ the other recommended privacy and security measures such as encrypting everything or having a privacy policy on their website. Our interview respondents discussed encrypting some, but not all of their data, and having little belief that privacy policies were useful. The survey respondents indicated a high level of data collection. Roughly three-quarters of developers collected information about the other apps installed on the users device. Some interviewees discussed collecting data that they didn't need, but thought might be useful in the future

While several government agencies, non-profit groups, and industry groups have developed guidelines for app developers on suggested privacy and security practices, the app developers we interviewed were not aware of and had not read these documents. This suggests that public policy around privacy and security is not reaching developers. In this section, we discuss hurdles to better privacy and security behaviors, and provide recommendations to encourage privacy-sensitive behaviors.

A. Third-Party Tools Should be More Transparent about Data Collection

Most app developers in our survey used third-party advertising or analytics services. Previous work shows that these libraries have permission to collect sensitive data [5], [6]. The developers we interviewed discussed their difficulties reading the policies and terms of use for the third-party APIs or services that they integrated into their apps. Popular ad and analytics companies should provide information about their data collection to app developers in an easy-to-read format. They should explain both what they collect and the purpose of that collection. This information could be provided in two places: as part of a quick-start guide, so developers can review before integrating the code, and after the developers has configured the third-party settings, so they can review how their choices impact their users' privacy and write their privacy policies.

Unfortunately, third-party tools may collect information about the smartphone user while having little or no relationship with the user, and thus have little incentive to protect user privacy. This may indicate a need for legislation to incentivize third-parties to provide clear information about their data collection to app developers and the end users.

In addition, survey participants demonstrated some confusion about whether they were using third-party tools, providing contradictory responses in different questions. This indicates that tools that automatically detect and describe third-party data collection may be helpful for developers.

B. With a Little Help From my Friends

App developers often mentioned searching for resources about security and privacy on the web. In addition, app developers in small companies rely on their friends and social networks for advice about privacy and security, while developers in larger companies may have experts within their company or legal counsel to turn to. Security and privacy advocates may find traction by intervening at a social level, such as by meeting with developers to discuss and improve their practices.

C. Legalese Hinders Reading and Writing of Privacy Policies

Less than half of small companies (fewer than 10 employees) informed their users about data collection through privacy policies on their websites. Several of our app developer interviewees had never read their own policy, and many others did not view it as a tool to communicate with users. Privacy policies were perceived as a tool that might protect them against lawsuits, but that small companies would not be targeted for lawsuits. This suggests that there is a need to emphasize that privacy polices need not be legalese, and can be an opportunity to communicate with their users. Furthermore, some interviewees expressed concern that full disclosure scares users away. This suggests that required,

standardized privacy notices might be a benefit for privacy-protective apps. Efforts of the government to develop such notices may provide guidance [30]. If all apps are required to provide notices, those who have good practices would not be punished for transparency.

Several interviewees believed that complying with the app stores' policies would provide sufficient legal protection, or that the app store would be monitoring them for compliance. This suggests that platform developers and market controllers are well-placed to encourage privacy and security behaviors. Platforms can highlight best practice notices and checklists, making them clear and accessible to app developers.

D. Small Companies Need Privacy and Security Tools

The smaller companies were the least likely to engage in privacy and security behaviors. Companies with fewer resources are less able to devote time or money to privacy and security issues. Therefore, small companies may need additional help or resources so they can overcome the hurdles to developing privacy policies and encrypting data. We suggest that privacy and security tools should be specifically targeted at small development companies with few resources. OS developers or open-source developers could focus on providing free tools to developers. These tools should be usable and not require legal expertise. In addition, companies of all sizes could be nudged to minimize data collection with tools that help developers decide what data to collect and when to delete it.

VIII. Conclusion

While there is general awareness of need for security measures, such as encrypting information or using SSL, there was a lack of understanding around privacy best-practices. Small companies rely on social networks and search engines for privacy and security advice. Privacy and security tools for developers must be quick, simple, and cheap, so that they can be used by time- and resource-constrained small companies. Platforms should make sure that it is easy to implement good security practices. App stores should provide privacy and security checklists, as they are uniquely positioned to reach developers. Third-party tools should make their data collection clear to developers and end users. More work is needed to make developing clear privacy policies a simple and routine part of app development.

Acknowledgements

This research was funded in part by Google, NQ, John and Claire Bertucci Fellowship, and NSF grants DGE0903659, CNS1012763, and CNS1228813.

References

[1] "App store stats summary," http://148apps.biz/app-store-metrics/, Sept 2012.

[2] A. N. Joinson, "Looking at, looking up or keeping up with people?: motives and use of facebook," in *Proc. of CHI '08*, 2008.

[3] H. McCracken, "Who's Winning, iOS or Android? All the Numbers, All in One Place," http://techland.time.com/2013/04/16/ios-vs-android/, April 16 2013.

[4] I. Leontiadis, C. Efstratiou, M. Picone, and C. Mascolo, "Don't kill my ads!: balancing privacy in an ad-supported mobile application market," in *Proc. of HotMobile '12*. ACM, 2012, pp. 2:1–2:6.

[5] T. Book, A. Pridgen, and D. S. Wallach, "Longitudinal analysis of android ad library permissions," in *Proc. of MoST 2013*, 2013.

[6] J. Lin, "Understanding and capturing people's mobile app privacy preferences," Tech. Rep. Ph.D Thesis CMU-CS-13-127.

[7] M. Egele, D. Brumley, Y. Fratantonio, and C. Kruegel, "An Empirical Study of Cryptographic Misuse in Android Applications," in *Proc. of CCS '13*.

[8] S. Fahl, M. Harbach, T. Muders, L. Baumgärtner, B. Freisleben, and M. Smith, "Why Eve and Mallory Love Android: An Analysis of Android SSL (in)Security," in *Proc. of CCS '12*.

[9] Z. Benenson, O. Kroll-Peters, and M. Krupp, "Attitudes to IT Security when Using a Smartphone," *Proc. of the FedCSIS*, pp. 1179–1183, 2012.

[10] E. Chin, A. Felt, V. Sekar, and D. Wagner, in *Proc. of SOUPS*.

[11] A. Felt, S. Egelman, and D. Wagner, "I've got 99 problems, but vibration ain't one: A survey of smartphone users' concerns," in *Proc. SPSM*, 2012.

[12] A. Mylonas, A. Kastania, and D. Gritzalis, "Delegate the smartphone user? security awareness in smartphone platforms," *Computers & Security*, 2012.

[13] E. Fife and J. Orjuela, "The privacy calculus: Mobile apps and user perceptions of privacy and security," *International Journal of Engineering Business Management*, vol. 5, no. 6, p. 7, 2012.

[14] J. Lin, S. Amini, J. Hong, N. Sadeh, J. Lindqvist, and J. Zhang, "Expectation and purpose: Understanding users' mental models of mobile app privacy through crowdsourcing," *UbiComp 2012*, 2012.

[15] R. Balebako, J. Jung, W. Lu, L. Cranor, and C. Nguyen, ""A Lot of Little Brothers:" Measuring User Confidence in Smartphone Security and Privacy," in *Proc. SOUPS*, 2013.

[16] 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 *Financial Cryptography and Data Security*, 2012.

[17] A. Felt, E. Ha, S. Egelman, A. Haney, E. Chin, and D. Wagner, "Android permissions: User attention, comprehension, and behavior," *Proc. of SOUPS*, 2012.

[18] A. M. McDonald and L. F. Cranor, "The Cost of Reading Privacy Policies," *I/S: A Journal of Law and Policy for the Information Society*, vol. 4, no. 3, pp. 540–565, 2008.

[19] C. Jensen and C. Potts, "Privacy policies as decision-making tools: an evaluation of online privacy notices," in *Proc. of CHI*, 2004.

[20] I. Pollach, "What's wrong with online privacy policies?" *Commun. ACM*, vol. 50, no. 9, pp. 103–108, Sep. 2007.

[21] "Mobile privacy mobile privacy: A better practice guide for mobile app developers," Office of Australian Information Commissioner, Sept. 2013.

[22] "Seizing opportunity: Good privacy practices for developing mobile apps," Office of the Privacy Commissioner of Canada, Information and Privacy Commissioner of (Alberta, British Columbia), Oct. 2012.

[23] K. D. Harris, "Privacy on the go, recommendations for the mobile ecosystem," Attorney General California Department of Justice, 2013.

[24] "Privacy design guidelines for mobile application development," GSMA Mobile Privacy, February 2012.

[25] "Best practices for mobile application developers: App privacy guidelines," Future of Privacy Forum and the Center for Democracy And Technology, July 12 2012.

[26] S. Hyrynsalmi, A. Suominen, T. Mäkilä, A. Järvi, and T. Knuutila, "Revenue models of application developers in android market ecosystem," in *Software Business*. Springer, 2012, pp. 209–222.

[27] A. Cravens, "A demographic and business model analysis of today's app developer," Tech. Rep.

[28] C. Jensen, C. Potts, and C. Jensen, "Privacy practices of internet users: Self-reports versus observed behavior," *International Journal of Human-Computer Studies*, vol. 63, pp. 203 – 227, 2005.

[29] C. S. Leem, H. S. Suh, and D. S. Kim, "A classification of mobile business models and its applications," *Industrial Management & Data Systems*, vol. 104, no. 1, pp. 78–87, 2004.

[30] "Privacy multistakeholder process: Mobile application transparency," http://www.ntia.doc.gov/category/privacy/u, Jul. 2013.