Location Privacy: User Behavior in the Field

Andrew Fisher

Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2012-247
http://www.eecs.berkeley.edu/Pubs/TechRpts/2012/EECS-2012-247.html

December 14, 2012

Acknowledgement

I thank my collaborator Leah Dorner and my advisor David Wagner for their efforts, advice, and tireless discussions.
Additionally, I thank Nathan Good, Serge Egelman, and Bjoern Hartmann for
helpful comments and feedback on this work.
This research was supported by Intel through the ISTC for Secure Computing
and by a gift from Google.

Abstract

Location Privacy: User Behavior in the Field

Drew Fisher

Current smartphone platforms provide ways for users to control access to information about their location. For instance, on the iPhone, when an application requests access to location information, the operating system asks the user whether to grant location access to this application. In this paper, we study how users are using these controls. Do iPhone users allow applications to access their location? Do their decisions differ from application to application? Can we predict how a user will respond for a particular application, given their past responses for other applications?

We gather data from iPhone users that sheds new light on these questions. Our results indicate that there are different classes of users: some deny all applications access to their location, some allow all applications access to their location, and some selectively permit a fraction of their applications to access their location. We also find that apps can be separated into different classes by what fraction of users trust the app with their location data. Finally, we investigate using machine learning techniques to predict users' location-sharing decisions; we find that we are sometimes able to predict the user's actual choice, though there is considerable room for improvement. If it is possible to improve the accuracy rate further, this information could be used to relieve users of the cognitive burden of individually assigning location permissions for each application, allowing users to focus their attention on more critical matters.

Introduction

Mobile phones are a fast-growing platform with rich data. Permission systems are a (mostly) effective way to limit what third-party applications can do [12]. Phone owners can control the sharing of their current location through mobile platforms' permission systems.

On Apple's iOS, the platform provides a subsystem called Location Services which provides access to the user's current location [15]. Apple allows users to limit applications' use of Location Services on a per-app basis. In particular, when an iOS application first attempts to access Location Services, the operating system displays a prompt to the user and asks the user whether to grant or deny access to location information for this application. The operating system then remembers the user's decision and applies it to all future requests from this application. Google's Android provides a similar subsystem, but in contrast, Android only allows users to toggle access to location on a global level. That is, if an Android user would like to disable location access privileges for a single app, their only option is to turn off the GPS for the entire phone.

These features provide a way for users to control how information about their location is accessed.

Our work is situated in the context of a great deal of prior research on location privacy. However, much of the past research on location privacy took place before smartphones and location-based applications were widely used. The world has changed significantly since this research, in several ways.

People have had several years to become more accustomed to using smartphones. In addition, smartphones have become much more capable devices and see much wider deployment. The smartphone userbase has changed.

The public is becoming increasingly aware of potential negative consequences of sharing information (particularly their location) on smartphones. There have been press reports about mobile phone operating systems tracking and storing phone location over extended periods of time, and Congress has heard testimony from Apple and Google regarding this tracking [4, 21, 13]. Application markets have seen instances of malware and greyware, and even curated markets like the iPhone App Store are not perfect. Recently, there was a public outcry against a company called Path which uploaded iPhone users' entire contact books to Path's servers [20].

There has been a significant amount of research on smartphone privacy, and much of this research has been focused on location privacy. However, there is little research on how users are using the privacy controls available in deployed smartphone

platforms. This suggests it would be interesting to understand user behavior in the wild.

Considering that most people say they are willing to share their location using phones [5], we think it would be useful to understand actual user behavior. We'd also like to consider a broader demographic than was used for past work, which tended to focus on university undergraduates aged 18–22.

Given studies indicating that most people are willing to share their location [5], given the past literature on runtime warnings indicating that users learn to reflexively click "OK" and that users are unlikely to immediately recognize potential negative effects of poor location privacy decisions [23, 10, 6], one might naturally expect that most iPhone users will ignore the platform's runtime location access prompts and simply grant access to any app that requests such access. **Surprisingly, we find that this is not the case.**

In this work, we study several questions about iPhone users' use of location privacy controls:

- Do users grant applications they have installed access to location information? Do they ever deny applications access to location information, and if so, how frequently does this occur? Does this rate vary from user to user? Can users can be separated into different classes based upon their sensitivity regarding location privacy?

- Do users treat different apps differently? Do they grant some apps access to location information at a higher rate than other apps?

- Given a user's past decisions about location access, can we infer how they will respond to a location prompt for a new application? Does it help to know how other users of that application have responded? This is useful, because it may point ways to improving runtime permission systems. One way we can avoid inundating the user with runtime prompts and stave off habituation is to ask the user fewer questions. If we can work out what the user most likely wants, we have an opportunity to take appropriate action without prompting the user.

Background

To understand how our study is designed, it helps to first have some background on how the iPhone platform and interface behave.

On iOS, the platform provides a subsystem called Location Services [15]. Installed apps that wish to use the user's location can request the phone's current location and notification of location changes from Location Services. The first time each app requests access to location services, the platform will show a warning prompt, as in Figure 1.1. The user selects either "OK" or "Don't Allow," and the platform respects that choice. The platform stores that decision and will apply that choice without prompting the user again if the app requests location access in the future.

Since users may change their mind about these privacy decisions, the platform collects these decisions in an editable list displayed in a Settings page for Location Services, as shown in Figure 1.1. This settings page allows the user to view their past decisions for each application that has ever requested access to Location Services and to change their past decisions.

Figure 1.1: On the left, a Location Services prompt, as a user would see it at runtime. On the right, a Location Services settings page.

Apple sets policies regarding the acceptable use of Location Services, and reserves the right to remove applications that violate these policies from the App Store. For example, Apple's guidelines say that applications may not use location information only for analytics. Apple's policy allows applications to request access to location services if the location information is needed to provide some aspect of application functionality or for serving advertising.

Related Work

There is a great deal of prior work on users' attitudes and behavior concerning location privacy [3, 7, 19, 16, 22, 1, 2, 14, 18, 24, 11, 9, 8]. Most prior work has focused on attitudes about location privacy, self-reported willingness to share location information, or behavior during a user study. In contrast, our work measures users' behavior in the field. Our data provides insight into how often iPhone users actually grant access to location information. In addition, some of the early work in this area focused on feature phones or was done before smartphones were as widespread and familiar. One could imagine that people might have grown more comfortable with sharing location information as they gained more experience with smartphones. Our work is able to measure location sharing today.

Methodology

We took a data-driven approach to answering these questions. We wanted to collect the actual privacy decisions people had made. Therefore, we designed a user study to gather this data from a collection of iPhone users.

On Apple's iOS, decisions to allow or deny location access to an app can be reviewed and revised through the Location Services settings page. There is no particularly good way to collect this information programmatically, due to the privacy implications of an untrusted third-party application being able to read (effectively) system privacy settings. Thus, we took a different approach: since all the past decisions on location access can be reviewed through the settings screen, and every iPhone ships with the ability to take screenshots, we decided to have study participants give us screenshots of their Location Services settings page.

We recruited nearly 300 anonymous iPhone users to participate in our study through Amazon Mechanical Turk, a microtask marketplace. Through a website created for the study, we instructed participants to take a series of screenshots of their iPhone's Location Services settings page. We guided them through the task of taking enough

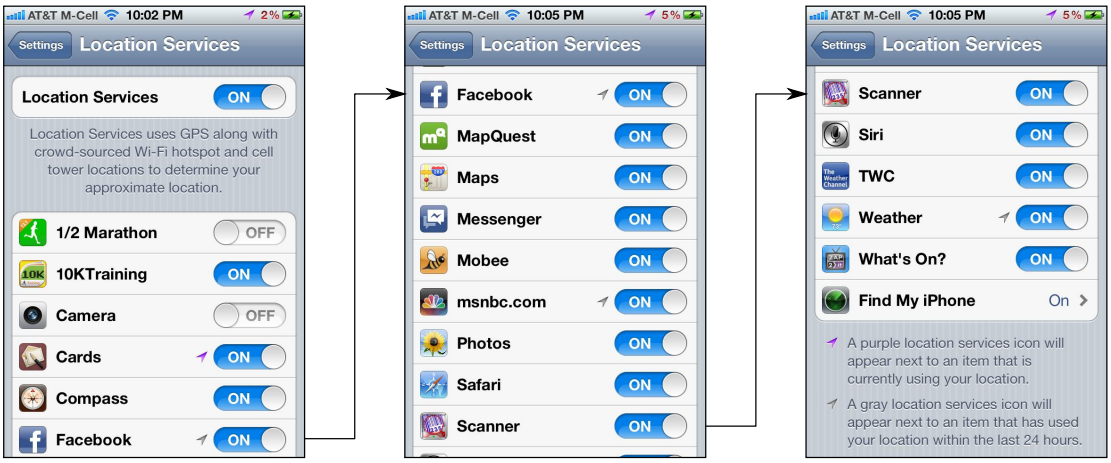

Figure 1.2: Participant 113 uploaded this series of overlapping screenshots, which taken together show the participant's location privacy decisions.

screenshots to cover all applications list on the page, as shown in Figure 1.2. These screenshots reveal, for each app that the user has installed and that has requested location access, whether the participant had most recently allowed or denied that request. We asked participants not to modify these settings until after they completed the task, in case navigating to the Location Services settings page made the participant want to change their settings. Participants then uploaded these screenshots to a server for collection and received a completion code to enter into the Mechanical Turk interface to prove completion of work.

We obtained IRB approval for this study. To protect participants' privacy, we gathered data from them anonymously. Participants took an average of 4 minutes to complete the study and we paid them $4 for their time.

After data collection, we threw out data that was not relevant to the study (spam uploads from Turkers), uploads with duplicate sha1 hashes (to prevent accidental

double-uploads from having undue weight in the dataset), and data from users where the user had not uploaded enough screenshots to reassemble the entire Location Services settings page. In total, we received 273 valid submissions.

The screenshots also contain information about whether Location Services are globally enabled or disabled and which apps had actually accessed location information in the past 24 hours. We did not use any of that information in this study.

Once the sets of screenshots were collected, a single graduate student coded all of the data from the screenshots into a machine-readable matrix with participants as rows and applications as columns. Each cell contained either "grant," "deny," or "no decision," to represent the user's choice.

Our study design gathers data on iPhone users' past decisions in response to location prompts. It does not provide any visibility into *why* users made the decisions they did, merely *what* decisions users did make. Nonetheless, we believe it is a reasonable way to find out the actual privacy decisions that the users made.

Our dataset encompasses data on 1126 unique apps from 273 unique users. There was considerable variation in the number of apps requesting access to location installed per participant: the average was 19.8 location-using apps per user, but one user had installed 88 apps which requested access to location (the most in our dataset) and granted the permission to all of them. Figure 1.3 shows a histogram of the number of apps installed per study participant.

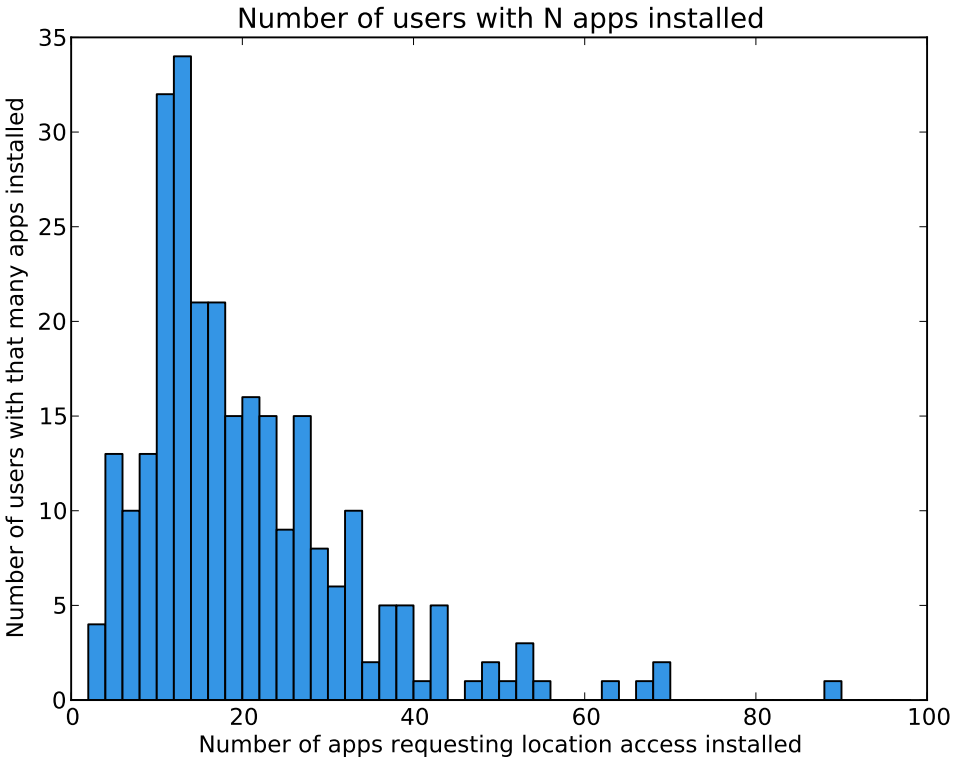

Figure 1.3: A histogram of the number of applications requesting location access installed per participant. The x-axis indicates a number of location-using apps. The height of the bar above the number N indicates the number of study participants who had N location-using apps installed on their phone.

There was also considerable variation in the prevalence of individual apps. Each app requesting location was installed on an average of 4.81 participants' phones, but some apps were installed on many participants' phones, and there was a very long tail: 717 of the 1126 apps had only been installed by one user in our dataset. Of the 25 apps appearing most frequently in our dataset, 11 are first-party applications that come pre-installed on the phone. Figure 1.4 shows the prevalence of apps with multiple installs observed in our study.

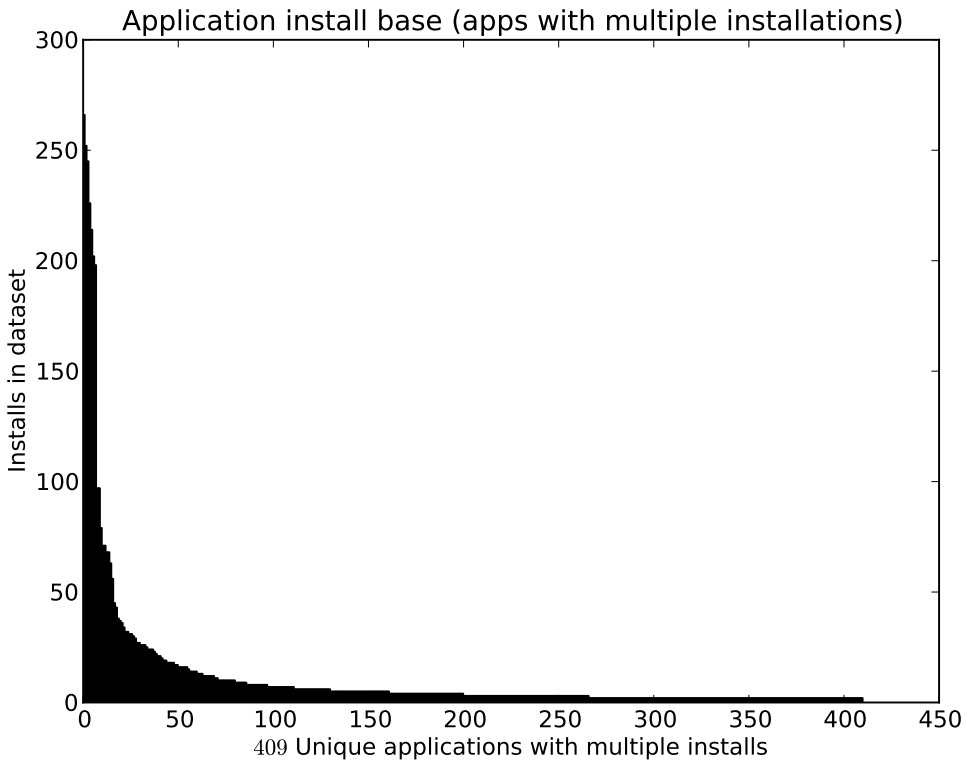

Figure 1.4: The number of installations for each of the 409 applications in our dataset that had been installed by more than one user. We saw 717 apps that were installed by only a single participant.

Results

Do users routinely grant access to location?

We found that, while many users grant location access to most or all of their applications, there are also some users who deny access to many of their applications. Users make a wide variety of privacy decisions. Most users granted location access to at least two-thirds of the apps that requested it, and 40 users granted location access to every app that requested it. However, a significant number of users denied location access to more than half of the apps that requested it, and one user denied location access to every app that requested it. Figure 1.5 shows a histogram of the fraction of location-requesting apps to which the user granted access.

We conclude that users are making use of the iPhone location privacy controls. Most users have denied location access to one or more application. Our data suggests that users are not reflexively clicking OK to grant location access to every application that requests it. Also, our data suggests that users have varying sensitivities to location privacy, and they make decisions accordingly.

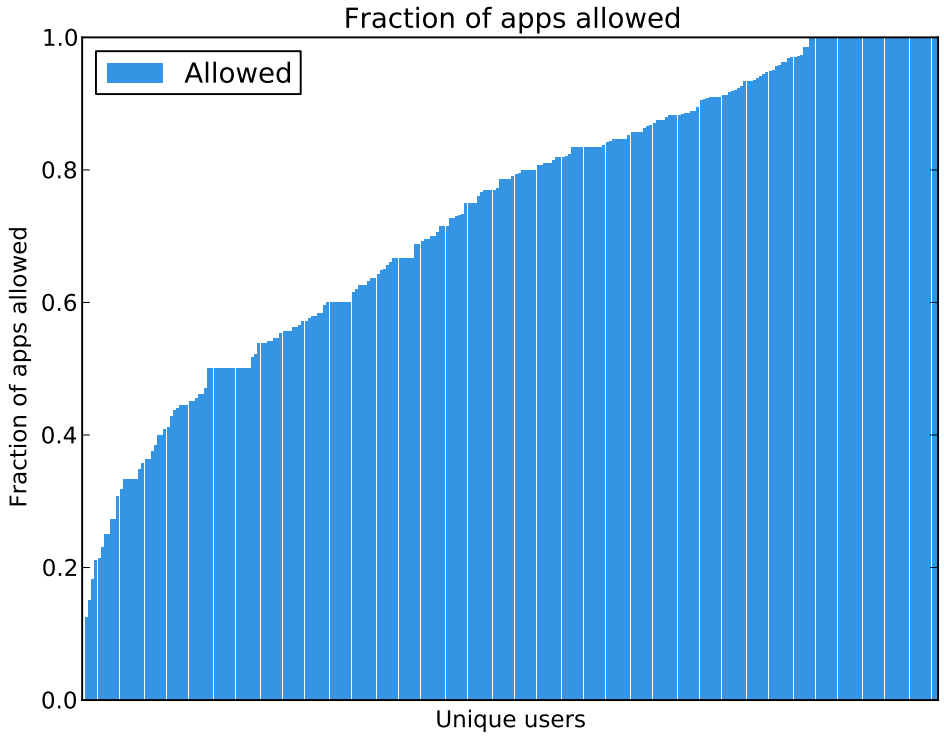

Figure 1.5: Users make varying privacy decisions. In this diagram, each vertical bar represents one of 273 unique users. The height of the bar indicates the fraction of location-requesting apps installed on that user's phone to which the user granted location access. The bars are sorted by this grant rate.

Do apps receive different amounts of trust from users?

We were curious whether users appeared to trust some apps more than others. In other words, did users grant some apps access to location information at a higher rate than others? Our dataset provides partial information about this question.

The answer to this question is not immediate from the data we collected. Our dataset has a long tail, with data on 717 of 1129 apps appearing for only one user each.

Thus, it is impossible to say precisely what fraction of users would grant each of these apps access to location information.

Nonetheless, even among the top 25 apps (the ones that were installed on the most phones within our dataset), we found wide variation in the fraction of users that granted that app access to their location. For instance, of the 97 users (53%) who had installed Shazam, 51 (53%) granted it access to location information. In contrast, 245 of 253 users (97%) who had installed the Maps application granted it access to their location. This suggests that users are more willing to grant location access to the Maps application than to the Shazam application. Figure 1.6 shows the rate at which users grant access to location information for the top 25 apps in our dataset.

Furthermore, it appears that users grant access more often to apps where location is central to the purpose of the app than to apps where location is a more optional feature or where it is less clear what benefit the user gets from sharing their location. For example, 28 of 29 users (97%) granted the foursquare app (in which users check in to various locations nearby) access to their location, a similar approval rate to that of Maps. In contrast, 40 of 68 users (59%) who had installed the IMDb app (which can use location to display nearby movie showtimes) granted it access to their location, and Shazam (which allows users to tag where they heard a song) was granted access by only 53% of users. This suggests that users are (at least in part) paying attention to the nature of the apps that are requesting access to their location and making their privacy decisions based on the expected value they will derive from sharing their location.

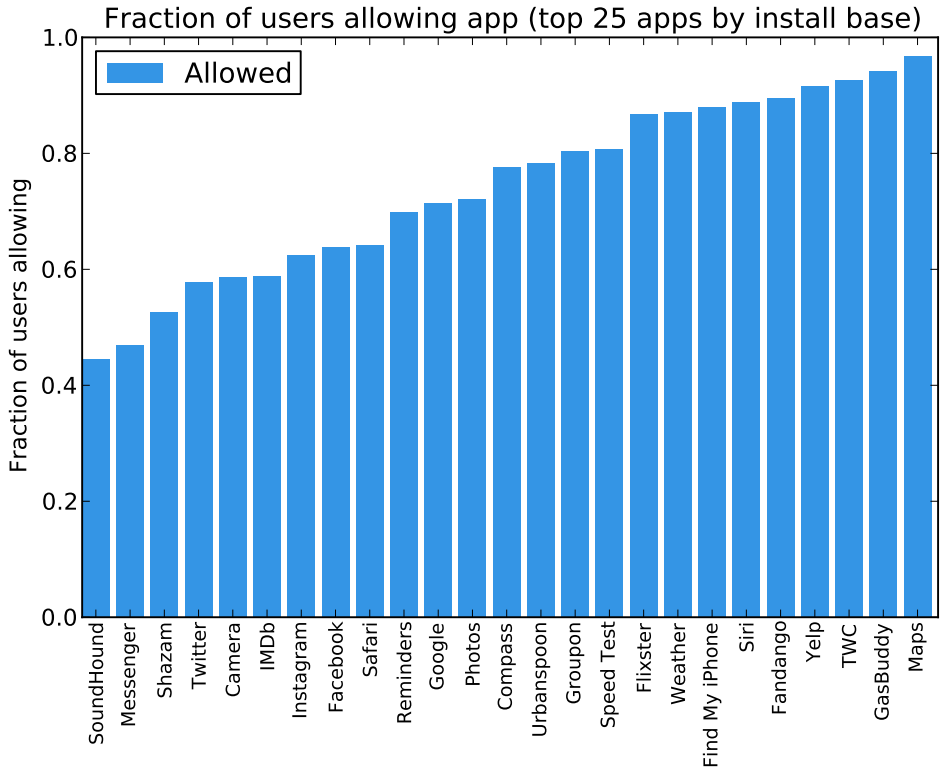

Figure 1.6: Apps receive different treatment by users. Users' willingness to grant location access seems to depend upon how the app uses location information.

It would be interesting to explore this observation in greater detail in future work to see if it holds in general and to understand why users made the decisions they did.

The intuitive explanation of these results is that Apple's design is working as intended to empower users without becoming so burdensome that they ignore the prompts.

Can knowledge of users' past decisions and previous decisions given about an app give good predictions on future choices?

Since we know that different users make different privacy decisions, and different apps elicit different privacy decisions, there may be latent factors in the data that can be used to predict what particular users would decide about particular apps, given knowledge of the user's preference and the app's reputation.

This is very similar to the problem of recommendation systems, which have been studied broadly. In that context, matrix factorization techniques are widely used to identify latent factors that help predict user preferences. There are various algorithms that have been developed to perform this sort of clustering. Common techniques include principal component analysis (PCA), k-means clustering, and singular value decomposition (SVD).

We applied these methods from the recommendation systems literature to see if we could predict the location privacy decisions made by iPhone users in our study. We attempt to predict the decision made by a single user for a single app, based upon information about other decisions made by this user for other apps and decisions made by other users for this app. Informally, we would expect that the rate at which this user approved location requests for other apps is predictive, as is the rate at which other users have approved location requests for this same app.

We evaluated these methods on our data set using cross-validation. We used the following procedure, which simulates a prediction algorithm that attempts to guess the outcome of the next user/app pair request based upon prior decisions from this user and decisions from other users.

First, we form a matrix representing the observations in our data set. Each row corresponds to a user, each column to an application, and the entry in each cell is either $+1$ (representing a grant of location access), -1 (a denial of location access), or 0 (indicating that this user has not been asked by this app for access to location information). Next, we select a user and application pair for which we want to make a prediction. The corresponding cell in our matrix of ground truth data must be nonzero, or we have no ground truth to which to compare our prediction. We erase this value from the matrix (by replacing it with a 0), then use PCA on the resulting matrix to produce a set of row factors and column factors. We multiply the row pertaining to the particular user by the column pertaining to the particular app for which we are trying to make a prediction. This yields a scalar. If this number is positive, our predictor guesses that the user would allow the app access to location information; if it is negative, the predictor guesses that the user would deny the app access to location information. The greater the absolute value of this number, the greater our confidence in the accuracy of the decision.

We can pick a threshold value for this confidence to restrict the set of user/app pairs that the algorithm will make a prediction on to situations where the classifier is

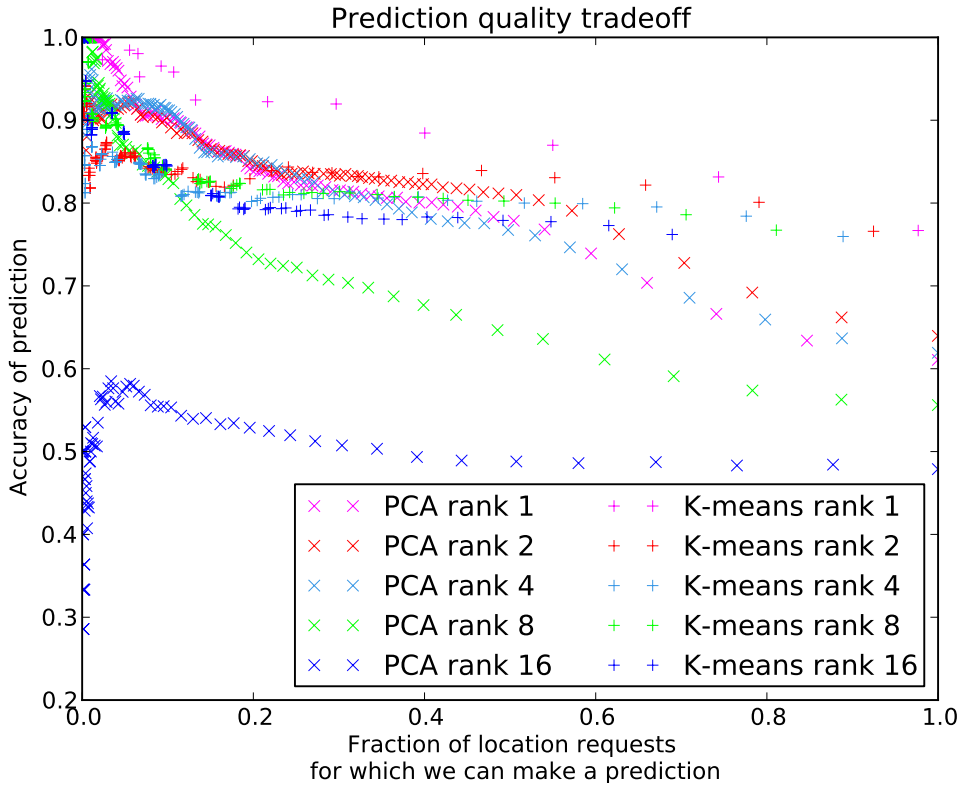

Figure 1.7: Accuracy/applicability tradeoff

more likely to be successful, in exchange for not being able to make predictions on chunks of the dataset. This allows tuning of the tradeoff between accuracy and applicability, as illustrated in Figure 1.7.

We tried several algorithms for matrix factorization: PCA, k-means, and SVD, each with several different rank values (1, 2, 4, 8, and 16). We observed the best accuracy from the k-means algorithm with rank 1. There exist several techniques that also take into account the sparseness of the matrix in their predictions [17], but we have not yet tried them.

Unfortunately, our results show only modest accuracy at prediction: for example, we can predict with $\approx 88\%$ accuracy if we only make a prediction about 40% of the time (but the remainder of the time, we make no prediction at all). Our data set is fairly small. We do not know whether more data from a large collection of users would significantly improve accuracy.

Possible applications

User attention is a valuable and limited resource. Given the ability to predict users' decisions, there exist a multitude of ways this knowledge could be leveraged to reduce the demand for user attention. Here's an example of how a system might modify behavior depending on how confident the system is:

- If the system has very low confidence of what the user will decide, then we can show the prompt as usual, and perhaps show what other users typically decide.
- If the system has moderate confidence, we could highlight the likely choice in the UI as the default.
- If the system has high confidence, we could simply take the expected action, and show the action taken in a non-modal notification which briefly allows for the reversal of the decision if the user interacts with the notification.

Conclusion

Users behave in a manner that is consistent with them desiring varying amounts of location privacy. Some apps are more trusted than others with location data. Knowing past decisions by a user or by other users about a particular app can inform whether a particular user will grant location access to a particular app.

It may be possible to leverage this information to reduce the number of privacy decisions users make, to select sane defaults based on past choices, or to use other designs to reduce the amount of attention demanded from users.

Acknowledgements

I thank my collaborator Leah Dorner and my advisor David Wagner for their efforts, advice, and tireless discussions. Additionally, I thank Nathan Good, Serge Egelman, and Björn Hartmann for helpful comments and feedback on this work. This research was supported by Intel through the ISTC for Secure Computing and by a gift from Google.

Bibliography

[1] Denise Anthony, David Kotz, and Tristan Henderson. Privacy in location-aware computing environments. *IEEE Pervasive Computing*, 6(4):64–72, 2007.

[2] Louise Barkhuus. Privacy in location-based services, concern vs. coolness. In *Workshop on Location System Privacy and Control at MobileHCI*, 2004.

[3] Louise Barkhuus and Anind Dey. Location-based services for mobile telephony: a study of users' privacy concerns. In *International Conference on Human-Computer Interaction*, 2003.

[4] Brian X. Chen. iPhone tracks your every move, and there's a map for that. `http://www.wired.com/gadgetlab/2011/04/iphone-tracks/`, April 2011.

[5] Erika Chin, Adrienne Porter Felt, Vyas Sekar, and David Wagner. Measuring user confidence in smartphone security and privacy. In *Proceedings of the 8th Symposium On Usable Privacy and Security (SOUPS)*, 2012.

[6] Nicolas Christin, Serge Egelman, Timothy Vidas, and Jens Grossklags. It's all about the Benjamins: an empirical study on incentivizing users to ignore security advice. In *Proceedings of the 15th international conference on Financial Cryptography and Data Security*, FC'11, pages 16–30, Berlin, Heidelberg, 2012. Springer-Verlag.

[7] Sunny Consolvo, Ian E. Smith, Tara Matthews, Anthony LaMarca, Jason Tabert, and Pauline Powledge. Location disclosure to social relations: why, when, & what people want to share. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI)*, 2005.

[8] Dan Cvrcek, Marek Kumpost, Vashek Matyas, and George Danezis. A study on the value of location privacy. In *Proceedings of the 2006 Workshop on Privacy in an Electronic Society (WPES'06)*, 2006.

[9] George Danezis, Stephen Lewis, and Ross Anderson. How much is location privacy worth? In *Proceedings of the Workshop on the Economics of Information Security Series (WEIS 2005)*, 2005.

[10] Serge Egelman, Lorrie Faith Cranor, and Jason Hong. You've been warned: An empirical study of the effectiveness of web browser phishing warnings. In *Proceedings of the twenty-sixth annual SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, pages 1065–1074, New York, NY, USA, 2008. ACM.

[11] Serge Egelman, Adrienne Porter Felt, and David Wagner. Choice architecture and smartphone privacy: There's a price for that. In *Workshop on the Economics of Information Security (WEIS)*, 2012.

[12] Adrienne Porter Felt, Kate Greenwood, and David Wagner. The effectiveness of application permissions. In *Proceedings of the 2nd USENIX conference on Web application development*, WebApps'11, Berkeley, CA, USA, 2011. USENIX Association.

[13] Arik Hesseldahl. Apple and Google answer tough questions from senators on the location brouhaha. `http://allthingsd.com/20110510/liveblog-apple-and-google-testify-before-congress-on-the-location-brouhaha/`, May 2011.

[14] Giovanni Iachello, Ian Smith, Sunny Consolvo, Mike Chen, and Gregory D. Abowd. Developing privacy guidelines for social location disclosure applications and services. In *Proceedings of the 2005 Symposium on Usable Privacy and Security (SOUPS)*, 2005.

[15] Apple Inc. *iPhone User Guide for iOS 5.1 software*, 2012.

[16] Patrick Gage Kelley, Michael Benisch, Lorrie Faith Cranor, and Norman Sadeh. When are users comfortable sharing locations with advertisers? In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*, 2011.

[17] Raghunandan H. Keshavan, Sewoong Oh, and Andrea Montanari. Matrix completion from a few entries. In *Proceedings of the 2009 IEEE international conference*

on Symposium on Information Theory – Volume 1, ISIT'09, pages 324–328, Piscataway, NJ, USA, 2009. IEEE Press.

[18] Scott Lederer, Jennifer Mankoff, and Anind K. Dey. Who wants to know what when? Privacy preference determinants in ubiquitous computing. In *CHI Extended Abstracts on Human Factors in Computing Systems*, 2003.

[19] Janne Lindqvist, Justin Cranshaw, Jason Wiese, Jason Hong, and John Zimmerman. I'm the mayor of my house: examining why people use Foursquare – a social-driven location sharing application. In *ACM CHI Conference on Human Factors in Computing Systems*, 2011.

[20] Nathan Olivarez-Giles. Path apologizes for saving iPhone contacts, wipes data, updates app. `http://articles.latimes.com/2012/feb/08/business/la-fi-tn-path-ceo-dave-morin-we-are-sorry-20120208`, Feb 2012.

[21] Sarah Perez. Android phones track your location, too. `http://www.readwriteweb.com/archives/Android_phones_track_your_location_too.php`, April 2011.

[22] Norman Sadeh, Jason Hong, Lorrie Cranor, Ian Fette, Patrick Kelley, Madhu Prabaker, and Jinghai Rao. Understanding and capturing people's privacy policies in a mobile social networking application. *Personal and Ubiquitous Computing*, 2009.

[23] Joshua Sunshine, Serge Egelman, Hazim Almuhimedi, Neha Atri, and Lorrie Faith Cranor. Crying wolf: an empirical study of SSL warning effectiveness. In *Pro-*

ceedings of the 18th conference on USENIX security symposium, SSYM'09, pages 399–416, Berkeley, CA, USA, 2009. USENIX Association.

[24] Jason Wiese, Patrick Gage Kelley, Lorrie Faith Cranor, Laura Dabbish, Jason I. Hong, and John Zimmerman. Are you close with me? are you nearby?: investigating social groups, closeness, and willingness to share. In *Proceedings of the 13th International Conference on Ubiquitous Computing (UbiComp)*, 2011.