Empirical Models of Privacy in Location Sharing

Eran Toch, Justin Cranshaw, Jason Hong, Patrick Kelly,
Janice Tsai, Lorrie Cranor, and Norman Sadeh

March, 2010
CMU-ISR-10-108

Institute for Software Research

School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

An expanded version of this work appeared in the Proceedings of the 12th ACM International Conference on Ubiquitous Computing, Copenhagen, Denmark, September 2010.

This work is supported by NSF Cyber Trust grant CNS-0627513 and ARO research grant DAAD19-02-1-0389 to Carnegie Mellon University CyLab from the Army Research Office. Additional support has been provided by Google, Microsoft through the Carnegie Mellon Center for Computational Thinking, FCT through the CMU / Portugal Information and Communication Technologies Institute, France Telecom and Nokia.

Keywords: Mobile location sharing technology, mobile social, information disclosure, privacy

Abstract

The rapid adoption of location tracking and mobile social networking technologies raises significant privacy challenges. Today our understanding of people's location sharing privacy preferences remains very limited, including how these preferences are impacted by the type of location tracking device or the nature of the locations visited. To address this gap, we deployed Locaccino, a mobile location sharing system, in a four week long field study, where we examined the behavior of study participants (n=28) who shared their location with their acquaintances (n = 373.) Our results show that users appear more comfortable sharing their presence at locations visited by a large and diverse set of people. Our study also indicates that people who visit a wider number of places tend to also be the subject of a greater number of requests for their locations. Over time these same people tend to also evolve more sophisticated privacy preferences, reflected by an increase in time- and location-based restrictions. We conclude by discussing the implications our findings.

1 Introduction

Location-aware applications are becoming more prevalent on mobile platforms. Development has been spurred by the adoption of GPS-enabled phones and WiFi positioning technologies. As the wireless market grows, so will the ability of users to continuously share their location with people and services. There has also been a recent increase in the use and availability of location sharing applications, allowing people to share their current location with their online social network (e.g. Twitter, Google Buzz, and Facebook). Location sharing is also the basis for new and increasingly popular mobile social networks, such as Foursquare and Gowalla.

Creating systems that enable users to control their privacy in location sharing is challenging. Evidence from several field studies have shown that users have complex privacy preferences, which depend on many factors: the entity that receives information about the location, the context of the sharing, the user's activity and so forth [12, 1, 14, 3]. Here, we opted to take a different approach in investigating this problem, investigating how location sharing preferences are impacted by the actual locations visited by users and by the way these locations are tracked. Also, we approach the problem of privacy from a statistical standpoint, looking for simple models that can predict some elements of people's location sharing preferences.

In this paper, we report on the results of a month-long user study in which these hypotheses were examined in a field deployment of Locaccino, a location sharing system [14]. We deployed Locaccino to a set of participants ($n_1 = 28$) who shared their location with their friends and acquaintances ($n_2 = 373$). We tracked user behavior of the two groups of users throughout the study, including location requests for and by study participants and their privacy settings. We align this empirical evidence with survey information, including a section in which participants provide detailed sharing preferences for a set of locations at which they were observed.

We start by exploring possible links between the characteristics of a location and a user's willingness to share this location with her social network. We show that users are more comfortable sharing their location when they are at places visited by a large and diverse set of people. While people's location sharing privacy preferences are known to be diverse and complex, we find that they can in part be predicted by analyzing characteristics of the locations where they are. We adopt the notion of *location entropy* [6] as a measure the diversity of visitors to a given location. Just as entropy is used to measure bio-diversity, it can be used to cap-

ture the intrinsic diversity of a location without looking at the functionality of that location (e.g., is it a private home? is it an airport terminal)? Our results show that locations with high entropy are more likely to be shared than places with low entropy, and that entropy can be established in urban environments by using a relatively small number of location observations.

We examine how people's location sharing privacy preferences relate to the type of location tracking devices they carry (i.e.,laptop versus cell phone) and the variety of places they visit with these devices. We find that users who are recorded at a large number of unique locations genenrally evolve more complex privacy preferences but also report finding location sharing more useful. The number of unique locations at which users are recorded depends on several factors, including the number of unique locations they actually visit, and the tracking device they carry. Our results suggest that the total number of unique locations visited by a user is a stronger predictor of the complexity of her privacy preferences than the type of device she carries. These more complex privacy preferences are reflected by the introduction of time and location restrictions in the user's location sharing preferences (e.g. "'Only disclose my location to my colleagues when I am between 9am and 5pm on weekdays and only when I am on company premises"').

Contributions Our main contributions are threefold: (1) we show that users are more comfortable sharing high entropy locations than low entropy locations; (2) we show that location tracking patterns, such as the number of locations visited in a day, impacts the overall privacy preferences of users; (3) we show that rich privacy controls, which allow users to restrict location sharing to specific times and locations, helps users define privacy preferences they are more comfortable with.

2 Related Work

Developments in geographical positioning and location management technologies have made it easier for application developers to create new location sharing applications. Locations can be found via GPS, cellular triangulation, and wireless positioning using databases of wireless access points mapped to locations, such as those provided by Skyhook Wireless. Recent developments have also made it possible to develop location sharing application on a multitude of platforms, including mobile computers (using the Skyhook Wireless API or the Windows 7 location API) and mobile phones (using the iPhone SDK or the Android SDK) [15].

2.1 Location sharing and Privacy Research

Several researchers have conducted studies to examine the usage of location sharing applications and the privacy concerns raised by these applications. Some of these studies have employed the experience sampling method (ESM) where users have carried devices to simulate location requests [1, 5], or involved small laboratory experiments where participants had simulated location sharing scenarios [2, 4].

Research has shown that the primary privacy concerns surrounding the disclosure of this information include *context* and *use* [2]. The willingness to share one's location and the level of detail shared depends highly on *who* is requesting this information [5] or knowing who is requesting this information [16], and the *social context* of the request [11]. Privacy concerns can depend on the situation or activity in which the user may be engaged [9]. In addition to the context of a location request, it is the users' own perceptions of the *use* of one's location information that impacts their privacy concerns [5]. Our study builds upon these works, investigating how two new factors, location characteristics and tracking method, impact privacy.

In our previous studies, we have investigated peoples privacy preferences in location sharing, using lab studies and ongoing field studies [14, 3]. In this paper, we focus on privacy preferences as a variable of location characteristics and tracking. We address the question using a new methodology, which combines a longitudinal field study using two types of location tracking devices, and a detailed location privacy survey. Furthermore, we apply methods from mobility analysis to the realm of privacy.

2.2 Mobility Patterns Analysis

Several interesting results demonstrate the potential of using mobile location technologies to study human behavior. In a series of papers, González et al. observed a large group of mobile phone users over six months, showing that phone users' mobility behavior falls into a small set of identifiable patterns [8]. Eagle et al. used location data to analyze patterns of human mobility and behavior [7]. Cranshaw et al. [6] introduced the concept of *location entropy* as a way to analyze the social context of physical interactions of users in a location sharing social network. Location entropy reflects the diversity of a location, by measuring the proportions of visits by unique visitors for a given location.

The use of location entropy has been suggested in another field: anonymity in

location-based services. Papers by Xu et al. [17] and Xue et al. [18] explore k-anonymity algorithms for hiding location origins in spatial queries. These works calculate the entropy of all requests for a location-based service, determining the resolution of the location sent to the service. On the contrary, we investigate how entropy impacts user perceptions of location privacy.

3 Overview of Locaccino

(a) Locaccino Web Application (b) Smartphone Locator

Figure 1: The Locaccino web application (on the left) and the mobile locator installed on a Nokia N95 phone (on the right). The web application is embedded within Facebook, and allows to request friends' locations, set up privacy settings, and receive privacy-related information. The phone locator reports the user's location and enables the user to request their friends' locations.

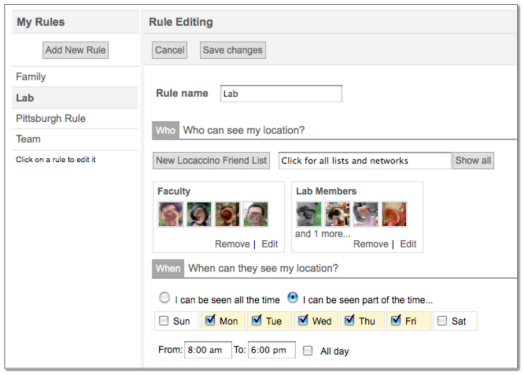

Figure 2: The Locaccino privacy settings user interface. This interface allows users to create *rules* specifying who can locate them, and in which conditions. Conditions include the time of the request and the user's location at the time of the request. This rule allows friends from the Faculty and Lab Members groups on weekdays between 8:00 AM and 6:00 PM.

Locaccino, our mobile location sharing application[1], leverages user's existing social networks on Facebook to facilitate sharing their location. Users add the Locaccino Facebook application and are able to request the location of Facebook friends who have added the Locaccino application and installed as Locator software. Users define location disclosure rules which determine the exact circum-

[1] `http://locaccino.org`

stances under which location information is disclosed. Locaccino is comprised of two main user-facing components:

- *Web application:* In the Locaccino web application, users can request friends' locations, set up privacy rules, and get privacy-related information. The user interface is available as a Facebook application accessed through a Web browser.
- *Locator software:* Users can install the Locaccino Locator on their laptop computers or phones. The software transmits the user's location to the Locaccino database every five to ten minutes.

3.1 Web Application

The Locaccino Facebook application contains three main pages: "Home", "Privacy Settings", and "Friends' Views."

3.1.1 Home Page

The "Home" page is the first page the user sees when entering the application (Figure 1 (a)). The page contains a map and a list of the user's Locaccino friends (i.e., Facebook friends who have enabled the Locaccino Facebook application.) Each friend in the list is marked with a visual sign when the friend is locatable, helping users see who they can locate. Clicking on a friend's name generates a location request. If the friend is locatable, their profile picture is displayed on the map. Users are not locatable if their locator is offline, if their rules do not permit the requester (the user who generated the request) to see their location, or if they are in "hidden mode" (which blocks all requests). The requester does not get a message regarding the reason for the deny, giving the requested user plausible deniability.

3.1.2 Privacy Settings

The "Privacy Settings" page, illustrated in Fig. 2, allows users to create, edit, or delete location disclosure rules. When first adding the application, the default disclosure policy is to deny all requests. Rules can allow access according to three criteria that we refer to as "restrictions":

- *Group (Who):* Group restrictions specify individual Facebook friends, groups thereof, or whole Facebook Networks (e.g. "Carnegie Mellon University)" with whom a user wants to share his or her location information.
- *Time (When):* Users can define the days of the week and a single time window for these days during which they wish to allow others access to their location information. For example, users can set a rule that will take affect in all workdays between 9:00 AM and 5:00 PM.
- *Locations (Where):* In the Privacy Settings interface users can select a geographic area where they wish to allow themselves to be locatable.

3.1.3 Friends' Views

The Friends' Views tab in the Locaccino interface allows users to review the complete log of requests made for their location. This page also allows users to see who can see their current location at that moment.

3.2 Locators

We developed two types of locators, targeted at laptop computers and Symbian-OS phones. The laptop locator determines the user's location using WiFi positioning, pulling information from Skyhook Wireless and our own database of the university's WiFi access points. The position is accurate within 20-30 meters range.

The mobile client (see Figure 1(b),) tracks the user's location using both WiFi and GPS positioning, and determines which one to use according to availability and battery life considerations. The client allows users to query their friends' locations, either by selecting a specific friend or by showing all nearby friends according to user-defined radius.

Both locator clients allow users to get additional information, including who has viewed their location in the last 24 hours and who can currently view their location. Users can also toggle "hidden" mode, which instantly hides their location from all their friends.

4 User Study Design

To investigate the question of the impact of mobility and tracking on location sharing, we deployed Locaccino in a month-long field study during July and August of 2009. The study includes an analysis of two types of participants:

- **Primary participants** ($n_1 = 28$): Participants who were directly recruited and compensated for the study. The participants were using Locaccino on a laptop or a mobile smartphone for a period of a month, and filled surveys.
- **Secondary participants** ($n_2 = 373$): Participants who were invited by the primary participants, and were able (but not requested) to locate primary and secondary participants, and to install a locator.

4.1 Method

Primary participants were recruited from the university population using fliers and posts on electronic message boards. Participants were compensated $30 for their participation in the study. Participation consisted of four phases, the pre-study survey, the installation of Locaccino, Locaccino utilization, and the post-study survey.

To join the study, potential primary participants were asked to complete a pre-study questionnaire in which their eligibility to participate in the study was evaluated. Primary participants were required to be members of the university community, to be users of Facebook, to regularly use a portable computer or mobile phone, and to be current customers of either AT&T or T-Mobile cellular services, a necessary requirement is for the operation of the phone locator. The pre-study survey included additional questions asking potential participants about issues such as their technical expertise, demographics (see Table 1), and initial attitudes towards privacy. Eligible primary participants were randomly assigned to a device: mobile phones or mobile computers. 2 participants had used Locaccino beforehand, at a previous user study, and the rest did not use Locaccino beforehand.

In the installation phase, all participants added the Locaccino Facebook application. Participants who were using phones were instructed to use the given phone as their primary phone, installing their personal SIM card on the new phone. We provided assistance to participants who encountered difficulties operating their phones. Phone participants were required to have an active data communication plan, and were compensated an additional $15 for their data usage.

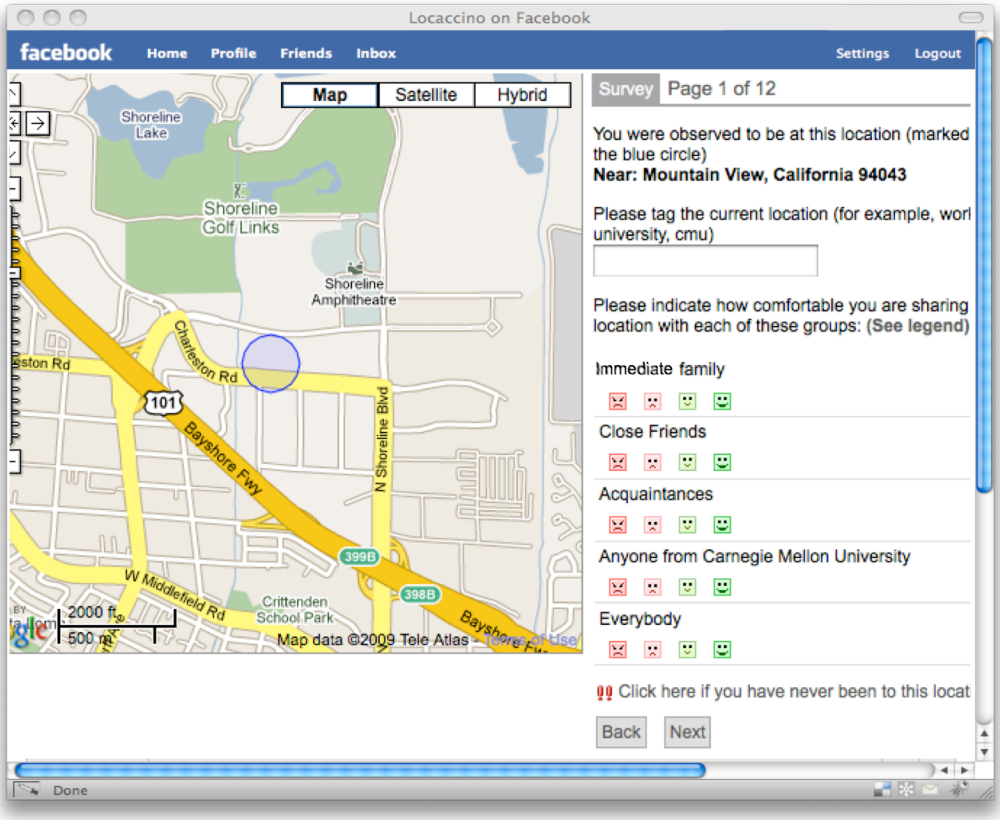

Figure 3: The location privacy survey interface. Participants were asked to tag 12 locations they were observed at, and to indicate how comfortable they were with sharing that location with different social groups.

At the beginning of the study, primary participants were instructed to invite at least 10 friends who were on Facebook to add the Locaccino application. Table 1 contains the number of friends who accepted the invitation. In the utilization stage, primary participants used Locaccino on their respective devices for a period of 4 weeks. All participants had started and finished the study at the same day. Primary participants were instructed to have the Locator running for at least an average 5 hours a day. All primary participants were asked to audit location requests in the Facebook application 3 times a week on non-consecutive days. Participants received email reminders if they did not follow these instructions for more than 2 days. At the end of the study, participants were asked to fill in an exit

Item		
Gender	22 male / 6 female	
Affiliation	25 student / 3 staff	
Device assigned	16 laptop / 12 smartphone	
	Mean	**SD**
Technical expertise	5.72	1.27
Number of friends	12.86	10.07
Number of locatable friends	8.38	7.30

Table 1: Study details and demographics. Technical expertise is measured between 1 - min and 7 - max. Locatable friends are secondary participants who installed the Locaccino locator.

survey and phone participants returned their phones.

The post-study survey included a location privacy survey, shown in Fig. 3, in which participants were asked to rate their comfort level with sharing 12 specific locations randomly sampled from among the locations where they were observed, using a uniform distribution. For each of these locations, participants were asked to assign a semantic tag to the location (e.g. "home" or "school") and to indicate how comfortable they were sharing the location on a 4-point scale ranging from *very uncomfortable* (1) to *very comfortable* (4), with 5 social groups: immediate family, close friends, acquaintances, anyone from the university population, and everybody[2]. The Likert scale, phrasing and presentation of the survey were similar to the auditing interface on the Friends' Views page. Users could indicate if the location was inaccurate (a total of 2 locations were marked as inaccurate). In addition to the surveys, we analyzed user actions such as location requests, privacy setting updates, and system usage.

[2]Some users were not observed in 12 distinct locations, and, therefore, were asked about all of their distinct observed locations.

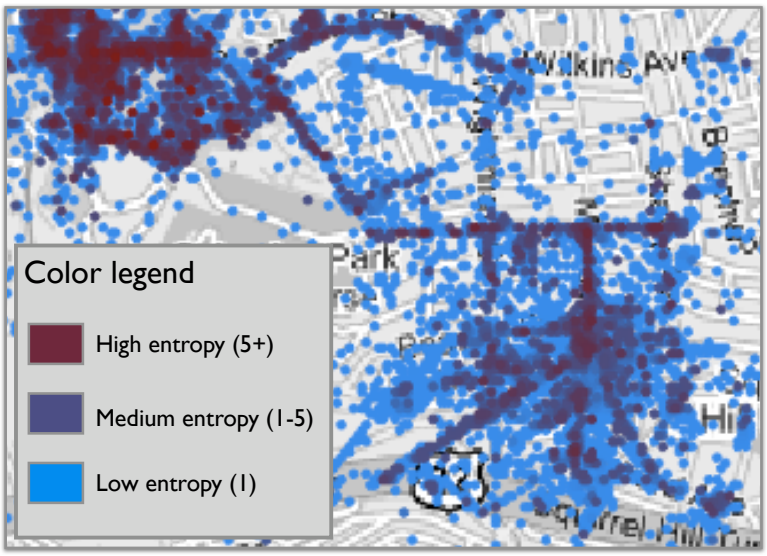

Figure 4: A map of a residential neighborhood the adjunct university campus. Each point on the map represents a location observation in our database, colored according to their level of entropy. Places such as campus building and commercial streets are visibly noticeable by their higher entropy.

4.2 Limitations

Running a long-term field study such as our own has several limitations. First, the participant pool contains mostly students. While being a common practice in ubicomp, we agree that it limits the generalizability of our results. It is important to note, though, that the study revealed distinct differences between the participants, even though the population was homogenous. Second, the study was carried out only at the limits of a single city, missing several important scenarios such as travel, vacations etc.

4.3 Data Analysis

Our motivation in analyzing the location data is to examine the privacy characteristics of locations using a straightforward mathematical model. Here we adapt the location entropy measure defined in Cranshaw et al. [6] to study the privacy of a location. Entropy measures the proportions of visits by unique visitors for a given location, assigning higher values to places which are visited evenly by

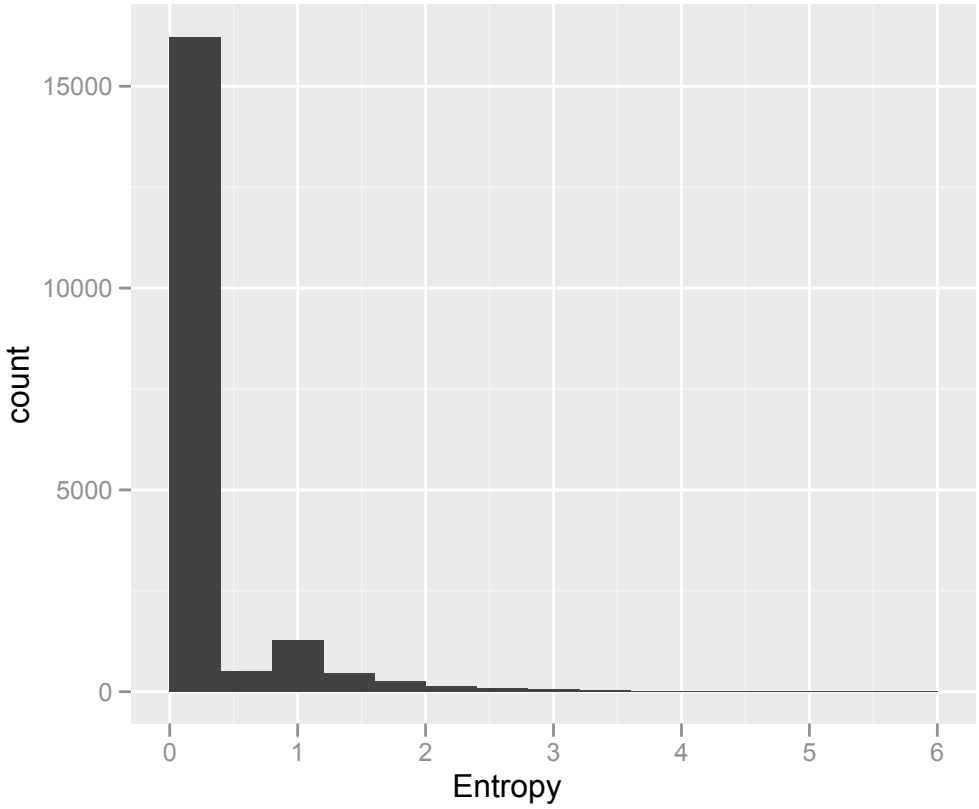

Figure 5: A histogram of the locations' entropy values. A heavy tail distribution can be observed. Over 82% of the locations have an entropy of 0 (a single person had visited the location.) 5% of the locations have an entropy around 1, reflecting a location shared by two people. 6% of the locations have a higher entropy, reflecting a shared location.

many users. Thus, we expect a particular user's home to have low entropy, while common areas like a university campus are likely to have high entropy.

Let U be the set of distinct users, L_u be the set of all location observations of a user $u \in U$, and L be the set of all locations: $L = \bigcup_{u \in U} L_u$. For $l \in L$, we define $\rho_{r,l}$ to be the set of all observations across all users that are within radius r from l: $\rho_{r,l} = \{l' \in L : d(l, l') < r\} = \bigcup_{u \in U} \rho_{r,l,u}$, where $d(l, l')$ is the distance from l to l'. Given this, we define $\rho_{r,l,u}$ to be the set of observations of user u within radius r from l: that is, $\rho_{r,l,u} = \{l'_u \in L_u : d(l, l'_u) < r\}$. We can then define by

$p(u, l, r) = \frac{\rho_{r,l,u}}{\rho_{r,l}}$, the fraction of observations within radius r from l that belong to user u.

Definition 1 *The* Entropy *of a location l with respect to radius r is given by:*

$$H(p(u, l, r)) = -\sum_{u \in U_L} p(u, l) \log p(u, l)$$

Intuitively, a location will have a high entropy value if many users were observed at the location with equal proportion, and a low entropy if it is dominated by a small number of users visiting the location unevenly. See Figure 4 for a concrete illustration of the difference between locations with high and low entropy. The university campus, visited by most of our users on a daily basis, have high entropy values on average. On the other hand, private residence have low entropy values.

We analyzed 4,150,171 location observations from 644 users. 570 of the participants were using laptops and the observations were collected from all Locaccino users, including this study, two other studies and other users who used Locaccino. Due to processing complexity, the observations were grouped into discrete 30 meters $\times$ 30 meters bins. Figure 5, which is the histogram of the locations' entropy, shows that entropy is distributed according to a heavy tail distribution. This pattern enables discovery of high entropy locations even with a small number of users. First, the number of high-entropy locations is small compared to low-entropy locations. Second, the probability of a random user to visit a high entropy location is inherently high as high-entropy locations are visited by many more unique visitors.

5 Results

We analyze data collected during the course of our study, along five dimensions: the privacy attached to locations, location tracking patterns, location requesting behavior, privacy preferences, and surveyed data.

5.1 Location privacy

We investigate the relationship between entropy of a location and the user's comfort in sharing that location. We now analyze the post-study survey results, in which users were asked to indicate how comfortable they were in sharing the locations in which they were observed in. We compare 4 measures for modeling

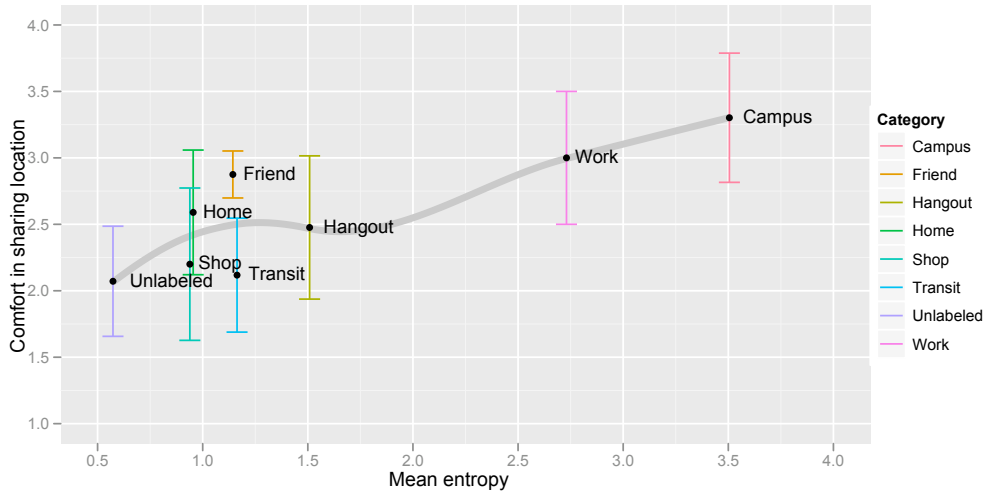

Figure 6: Comfort in sharing location with an average of scores for university population, acquaintances and everybody, according to entropy. Each points is the average of the category (e.g., "Campus" for university location, "Friend" for a friend's home, "Hangout" for restaurants and coffee shops, and so forth. Comfort in sharing location is based on a four point scale, where 1 is "very uncomfortable" sharing a location and 4 is "very comfortable." The blue line depicts the moving average using local polynomial regression fitting.

locations: entropy, number of unique visitors, number of overall visits, and the number of the visits by a given user.

Table 2 shows that users tended to feel less comfortable in sharing low entropy locations (e.g., home, friend's house, a shop) than high entropy locations (e.g., university campus). The impact was stronger when it comes to sharing location with distant social groups (university population, acquaintances, and everybody) than with friends and family. Analysis of variance (ANOVA) shows that entropy is significant for both analyzing location sharing with friends and distant social groups. Counting the number of overall visits to a location was also found to be significant only for distant social group, and is overall weaker than entropy. The number of visits to the location of the user who performed the ranking was not found to be significant.

The number of unique visitors was not found to be significant, a finding we explain by the impreciseness this measure has in low entropy locations. For example, if a user invites another user to her home, the number of unique visitors

for that location grows, even though home is still considered private to the homeowner. Entropy, on the other hand, takes into account the proportion of visits by unique users, which is a robust method in differentiating between seldom visited and often visited locations. This explanation is backed up by the correlation between unique visitors and comfort in sharing location. The correlation between unique visitors and comfort in sharing locations is significant for places with entropy above 1 (Pearson's product-moment correlation, $\rho = 0.235$, $t = 2.45$ and $p < 0.015$), but insignificant for places with entropy below 1 (Pearson's product-moment correlation, $\rho = 0.14$, $t = 1.27$ and $p < 0.2$).

Item	**Friend**		**Distant**	
	F	p	F	p
Entropy	5.46	**0.02**	15.57	$<$ **0.001**
Unique visitors	0.48	0.48	1.03	0.30
Overall visits	1.87	0.17	11.58	$<$ **0.001**
User's visits	0.0002	0.98	1.53	0.22

Table 2: Analysis of variance (ANOVA) results, showing the impact of different location characteristics on the comfort in sharing location. The results including comfort in sharing location with two types of social groups: friends and distant social groups (the average of sharing with the university population, acquaintances, and everybody). Entropy is significantly related to comfort in sharing location for both groups. Counting the overall visits is also significant, but only to the distant social groups. ANOVA is based on linear regression, where F is the ratio between the estimate of between-group variance to the estimate of within group variance, and p is the significance level. Significant p values are highlighted. The degree of freedom (df), is 1 to 179 for all rows.

In addition to the entropy of the location, the social context of sharing - who the participants are sharing their location with - also had an affect on their willingness to share the location. Figure 7 shows the relation between comfort in sharing location and entropy, for 5 different social groups of requesters. Unsurprisingly, users were more willing to share their location with friends and family than with university population, acquaintances and everybody. However, it is interesting to note that there are hardly any differences between friends and family.

Now, we will move on to a qualitative analysis of the locations visited by the participants and their respected privacy. In the survey, participants were asked to tag their locations (open comment). We then grouped these tags into 8 categories, shown in Fig. 6. The categories "Home," "Work" and "Campus," are the most common. The category "Transit" represents local travel. The category "Hangout" represents locations which are outside of home and are used for leisure (restaurants and coffee shops form the majority). The category "Friend" represents a friend's house. The category "Unlabeled" represents locations which were not tagged by participants[3].

We then examined the privacy values assigned locations in different categories. Fig. 6 allows us to analyze location sharing preferences according to the location types. The graph aligns the sharing comfort score the distribution of for university, acquaintances and everybody to the entropy. Campus and work, the most common categories, were also the ones that users were most comfortable sharing. To our surprise, home was not the most private place, and several categories, including shop, transit, and unlabeled were considered more private.

5.2 Location Tracking Patterns

In this section we scale our analysis of empirical privacy models from sharing single locations to sharing large sets of locations over time. We look at the number of daily observable locations as a simple measure for different usage scenarios for location sharing usage. Applications that have the potential of disclosing more locations, which may contain more sensitive locations, may be perceived as riskier by users. To investigate this hypothesis, we have divided the users into two groups: *high visibility users* and *low visibility users*. We measure visibility by the average number of unique locations observable by the system in a day. These locations were not requested necessarily by the other users, and are not necessarily regulated by the user's privacy preferences. High visibility users are users who were observable for each day an average number of location which is higher than the overall median (3.4 unique locations per day). Unique locations are defined as locations which are at least 500 meters away from each other. Low visibility users were classified if they visited less than the median number of unique locations each day on average.

[3]The Unlabeled category included 8 locations by 5 users. We have ruled out the possibility of the location to be either the home of the user or the university campus. The category had very low sharing comfort scores, which leads us to hypothesize that participants were reluctant to tag places which were extremely private.

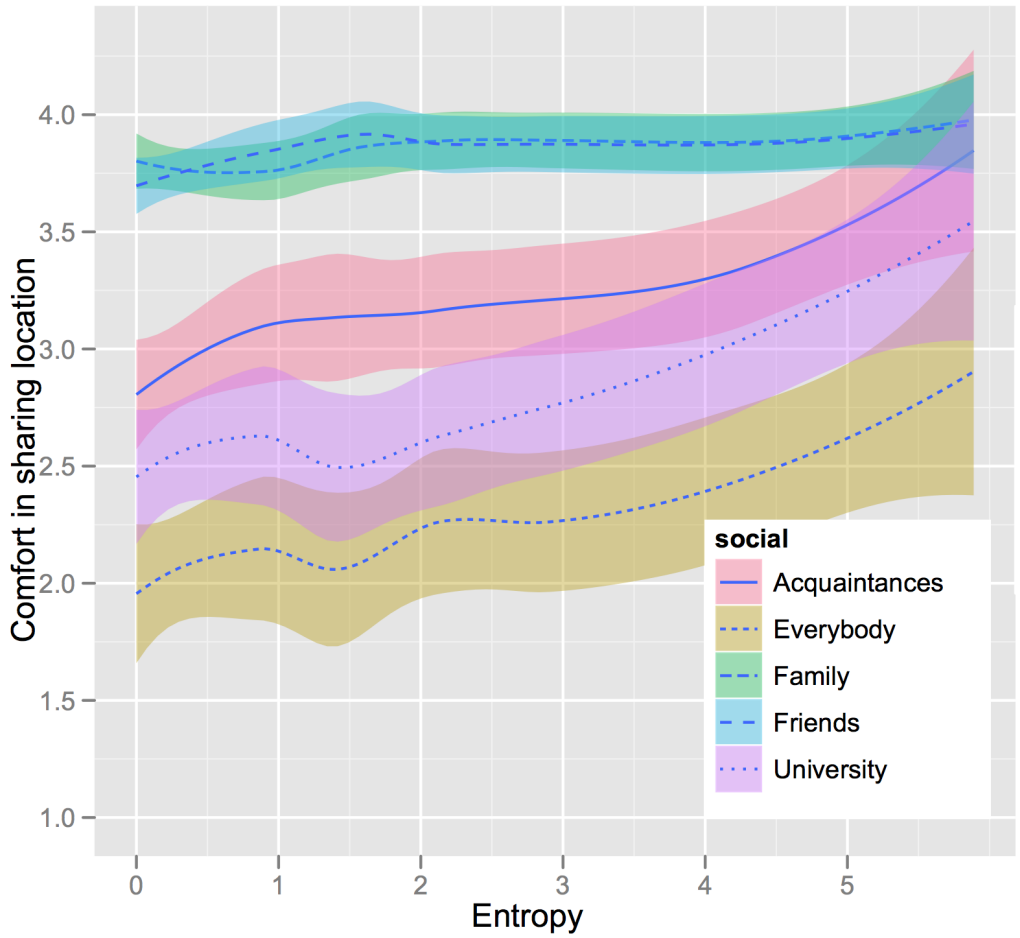

Figure 7: Comfort in sharing location versus entropy, for 5 social groups. Comfort in sharing location is based on a four point scale, where 1 is "very uncomfortable" sharing a location and 4 is "very comfortable." Lines represent moving averages, based on local polynomial regression fitting. Colored areas are error bounderies. The lines for Friends and Family averages overlap. We can see that for the acquaintances, everybody, and university groups, sharing comfort is strongly correlated with entropy.

In our study, we are interested in observed mobility, which is the potential visibility of a user's location to other users. Naturally, the device used for tracking users is tightly related to visibility. While mobile phones are nearly always on, and

therefore always transmitting a user's location, laptops transmit the locations only when they are on and when they are connected to the internet. Laptop users, on average, should be as mobile as cell phone users, however they are much less likely to use their laptop at a restaurant, a bar, or in transit between locations. 5 laptop users and 9 smartphone users were categorized as high visibility users, while 11 laptop users and 3 smartphone users were categorized as low visibility users.

The privacy associated with different location categories provides us with some interesting distinctions between continuous and sporadic tracking. Locations with low entropy are also the ones that are provided mostly by high mobility users. The "Home," "Work," and "Campus" categories were reported by users in both groups. Places such as "friend," "transit," "hangout," "shop," and "unlabeled" are visited
mostly by high mobility users. These locations were also the ones that users were least comfortable sharing. Therefore, continuous tracking has the potential of revealing more locations that people are less comfortable to share.

5.3 Location Requests

Participants' locations were requested a total of 848 times during the four weeks study. This number represents an average of 32 requests per user or an average of 1.35 requests per person per day. The number of requests is tightly related to users' visibility. Users who were more mobile had also received significantly more location requests (ANOVA: $F = 14.713$, $df = 1$ and 24 and $p = 0.00079$), even though the number of friends is not statistically related to the number of requests.

Figure 8 depicts the trend in requests for participants of the two visibility groups defined above. While the two groups were requested similarly at the beginning of the study, the difference increases with the progress of the study. We hypothesize that friends who requested the location of low mobility users did not learn much from the request and gradually decreased the number of requests. On the other hand, friends who requested high visibility users had found the information useful or interesting, and had increased their requests over time.

High visibility participants were also making more requests for the locations of their friends than were their low visibility counterparts. These results are not statistically significant (96.72 versus 40.31, $p = 0.143$ according to the Mann-Whitney test), but they are illustrative of increased usage by high visibility users. We conclude from this that the users requesting the locations of our study partic-

Figure 8: Location requests made to study participants by their Facebook friends, throughout the course of the study. Requests for low visibility users decline steadily while requests for high visibility users increased more than twofold over the course of the study. Local minimum values in requests for high visibility users correspond with weekends. We did not see correspondence with weekends with low visibility users, perhaps because the average number of requests did not allow enough variance in requesting behavior.

ipants perceived considerably more utility and value when the results were from mobile phone location trackers, and this encouraged usage.

5.4 Privacy Preferences

In this section, we analyze the rules that study participants defined to express their location sharing privacy preferences. The average number of rules per participant is 1.51. In analyzing users' privacy policies, we look at several properties: the expressiveness of the policies (as reflected by the number and diversity of restrictions used across all rules) and the way users change their policies.

Analyzing the types of rules that participants created shows that the number of rule restrictions rises with the user visibility (ANOVA: $F = 5.63$, $df = 1$ and 27 and $p < 0.025$). Users with high visibility have used, on average, 60% more restrictions in their rules than users with low visibility. The overall number of restrictions, across all rules, reflects the expressiveness of the policy, and the effort in refining it. When analyzing the types of restrictions used by participants, the impact of visibility is even more significant. High visibility users are 4 times more likely to use location restrictions and 7 times more likely to use time based rules.

13 of the 28 study participants had changed their rules at least once after initially creating them. The average number of updates for those who had changed their rules at least once is 4.82. Similarly to the number of restrictions, users with high visibility have higher number of rule edits (ANOVA: $F = 10.75$, $df = 1$ and 27 and $p = 0.0028$).

We have tested if the rule changes had relaxed or tightened the privacy policy by checking retroactively if the outcome of a pre-change request would have been different due to the rule update. In 7 out of 13 rule update instances, the outcome of at least one request would be disclosed instead of denied. There were no instances that there were cases in which the outcome changes from disclose to deny.

5.5 Survey results

The participants were asked to complete both a pre-study survey and a post-study survey, giving us additional insight into the usefulness of location sharing, privacy concerns and privacy preferences. First, let us examine usefulness (see table 3 for mean values and significance). Overall, participants found the application useful. 75% of the participants said that they are going to keep using Locaccino, while 25% said that they would not. 10 out of 28 users were eventually using the system a month after the end of the study (i.e., had the locator running or performed any other type of activity in the system).

Item	**Value**	F	p-value
Overall Locaccino usefulness	4.74	4.54	**0.043**
Friends rules usefulness	5.48	4.68	**0.04**
Networks rules usefulness	5.33	0.68	0.41
Time rules usefulness	4.74	5.14	**0.03**
Location rules usefulness	5.14	4.15	0.052
Combination usefulness	5.22	3.86	0.060

Table 3: Analysis of post study survey results. Value is the mean of the survey answers, on a 7-point Likert (1 stands for not useful and 7 for extremely useful). The other two columns show the results of analysis of variance (ANOVA) test of the relation between the survey answers and the visibility of the participants (average number of unique locations per day). High visibility users consider Locaccino to be more useful. At the same time, they find different types of rule restrictions more useful than low visibility users. All types of restrictions, except Facebook network based rules were below or near the significance level. ANOVA is based on linear regression, where F is the ration of the estimate of between-group variance to the estimate of within group variance, and p is the significance level. Significant p values are highlighted. The degree of freedom (df), is 1 to 28 for all rows.

High visibility users were more concerned, on average, with sharing their location with different social groups. When assessing how comfortable participants would be sharing their locations with different social groups, we asked participants about their concerns in sharing locations under different conditions: anytime, at times they have specified, and in locations they have specified. For all social groups, high visibility participants were more concerned about sharing their location than laptop participants. The difference ranged from 0.7 points for family and 1.28 points for university (we did not ask about the 'everybody' social group). The difference is statistically significant for the university group ($p < 0.043$, using a one-sided t-test).

Comparing the survey results to location sharing results presented in Figure 6 reveals that while the average sharing comfort for specific locations was unrelated to visibility, we witnessed lower comfort for unique locations, primarily visited by high visibility users. The survey shows that the relatively small number of unique locations plays an influential role in the overall privacy concerns. One possible explanation is that even if users visit a unique (and private) location once, the expectations of reporting this location to undesired acquaintances is enough to influence general concern.

Our results show how location sharing controls can be used to cope with privacy concerns. The concerns users had when sharing their locations anytime were drastically reduced when restricted by social group, time or location. Visibility is linked to friends restrictions and time restrictions and location and combination restrictions are not far from being statistically significant. These results support the empirical evidence from the rules utilization, which showed that visibility is related to the richness and refinement of the privacy policy.

6 Discussion and Implications for Design

In this field study, we found that location characteristics and visibility characteristics are significantly related to privacy preferences in location sharing scenarios. The results reveal that locations have an inherent privacy characteristic, which can be predicted by the entropy of the location and to some degree by the frequency in which users visit the location. This phenomena can be grounded by several theories related to information value and privacy. Cognitive models of information processing show that people assign higher value to unique and irregular information [10]. In our case, the potential of being located in rare and unique places may reveal more information than being located in public places.

The impact of location tracking patterns on privacy concerns can further be explained by a theory of online privacy and disclosure presented by Palen and Dourish [13]. Disclosing locations in public places helps users maintain their public persona by associating themselves with these places. On the other hand, disclosing locations in private places can reveal too much information and compromise the user's ability to control her public persona. For example, disclosing a location when the user is within the university's sport facility maintains the user's public image, which is not the case when revealing a friend's home.

Understanding people's privacy preferences when building location sharing applications can lead to more trustworthy systems. In the following sub-sections, we present several design implications we believe can benefit designers.

6.1 Privacy Controls

Designers should examine the location context of their location aware services. The types of locations their users visit, the entropy of the locations, and the users's visibility pattern may have a profound impact on the amount of privacy control their users will require. Our research can be used to derive default privacy preferences associated with different places based on the entropy of those places. Possible future work could include identifying the groups of users one might be willing to share a low entropy location with (e.g. perhaps other people who also visit this location - for instance, one is likely to be willing to share their home location with their spouses and their work location with their colleagues).

6.2 Location Entropy

We encourage designers to use entropy as measures for building more privacy preserving location-based services. While we have shown how entropy is related to privacy, we believe it can capture other attributes, such as the social properties of a location. It is important to note that entropy can be calculated using general data, without requiring specific information about the given user.

6.3 Client platforms

Mobile appliances are tracking our movements in an increasingly intimate manner, due to advancements in network usage, battery consumption, multi-tasking mobile operating systems and so fourth. Users will likely request richer privacy settings as this trend advances. This in turn could lead to increased user burden

levels, unless one can identify good default policies. Our results show how the entropy of locations relate to the types of privacy preferences people have when they are at a location. This can be used as a first step in designing clients that preserve user privacy.

7 Conclusions

This research presents the findings of a study examining the impact of entropy and number of places users are seen at on their privacy preferences. Our study involved conducting a four-week field investigation of a live location sharing application, in which participants shared their real-time location with actual friends and acquaintances. Our findings are as follows:

- High entropy locations, namely locations frequently visited by the a diverse set of unique users, are considered less private by users.
- Highly mobile users (as recorded by the system) receive significantly more location requests than less mobile users, and report finding location sharing more useful overall.
- Location sharing privacy settings that enable users to restrict location disclosure to particular times and places, seem to play an important role in capturing people's privacy preferences, especially those of more mobile users.

8 Acknowledgment

This work is supported by NSF Cyber Trust grant CNS-0627513 and ARO research grant DAAD19-02-1-0389 to Carnegie Mellon University CyLab from the Army Research Office. Additional support has been provided by Google, Microsoft through the Carnegie Mellon Center for Computational Thinking, FCT through the CMU / Portugal Information and Communication Technologies Institute, France Telecom and Nokia. Some of Locaccino's WiFi-based location tracking functionality runs on top of technology developed by Skyhook Wireless. The views and conclusions contained herein are those of the authors and should not be interpreted as representing the official policies or endorsements, either express or implied, of one or more of the project sponsors. We would like to thank David Eggerschwiler for developing the Symbian mobile locator. We would also like to thank Lujo Bauer and Kami Vaniea for their feedback on this research.

References

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

[2] L. Barkhuus, B. Brown, M. Bell, S. Sherwood, M. Hall, and M. Chalmers. From awareness to repartee: sharing location within social groups. In *CHI '08*, pages 497–506, 2008.

[3] M. Benisch, P. G. Kelley, N. Sadeh, and L. F. Cranor. Capturing location privacy preferences: Quantifying accuracy and user burden tradeoffs. Technical Report 10-105, Carnegie Mellon University, Institute for Software Research (ISR), March 2010.

[4] B. Brown, A. Taylor, S. Izadi, A. Sellen, J. Kaye, and R. Eardley. Location family values: A field trial of the whereabouts clock. In *Ubicomp '07*, pages 354–371. Springer-Verlag, 2007.

[5] S. Consolovo, I. Smith, T. Matthews, A. LaMarca, J. Tabert, and P. Powledge. Location disclosure to social relations: Why, when, & what people want to share. In *CHI '05*, 2005.

[6] J. Cranshaw, E. Toch, J. Hong, A. Kittur, and N. Sadeh. Bridging the gap between physical location and online social networks. In *Ubicomp '10*, Pittsburgh, PA, 2010.

[7] N. Eagle and A. S. Pentland. Reality mining: sensing complex social systems. *Personal Ubiquitous Computing*, 10(4):255–268, May 2006.

[8] M. C. González, C. A. Hidalgo, and A.-L. Barabasi. Understanding individual human mobility patterns. *Nature*, 453(7196):779–782, June 2008.

[9] G. Iachello, I. Smith, S. Consolovo, G. Abowd, J. Hughes, J. Howard, F. Potter, J. Scott, T. Sohn, J. Hightower, and A. LaMarca. Control, deception, and communication: Evaluating the deployment of a location-enhanced messaging service. In *Ubicomp '05*, pages 213 – 231. Springer-Verlag, 2005.

[10] P. Ingwersen. Cognitive perspectives of information retrieval interaction: Elements of a cognitive IR theory. *Journal of Documentation*, pages 3–50, 1996.

[11] A. Khalil and K. Connelly. Context-aware telephony: Privacy preferences and sharing patterns. In *CSCW '06*, 2006.

[12] C. Mancini, K. Thomas, Y. Rogers, B. A. Price, L. Jedrzejczyk, A. K. Bandara, A. N. Joinson, and B. Nuseibeh. From spaces to places: emerging contexts in mobile privacy. In *Ubicomp '09*, pages 1–10, 2009.

[13] L. Palen and P. Dourish. Unpacking "privacy" for a networked world. In *CHI '03*, pages 129–136, New York, NY, USA, 2003. ACM.

[14] N. Sadeh, J. Hong, L. Cranor, I. Fette, P. Kelley, M. Prabaker, and J. Rao. Understanding and capturing people's privacy policies in a mobile social networking application. *Personal and Ubiquitous Computing*, 13(16):401 – 412, August 2009.

[15] J. Tsai, P. Kelley, and L. C. ad Norman Sadeh. Public perceptions of the risks and benefits of location-sharing technologies. In *37th Research Conference on Communication, Information, and Internet Policy (TPRC)*, 2009.

[16] J. Tsai, P. Kelley, P. H. Drielsma, L. F. Cranor, J. Hong, and N. Sadeh. Who's viewed you? the impact of feedback in a mobile-location system. In *CHI '09*, pages 2003–2012, 2009.

[17] T. Xu and Y. Cai. Feeling-based location privacy protection for location-based services. In *CCS '09: Proceedings of the 16th ACM conference on Computer and communications security*, pages 348–357, New York, NY, USA, 2009. ACM.

[18] M. Xue, P. Kalnis, and H. K. Pung. Location diversity: Enhanced privacy protection in location based services. In *LoCA '09: Proceedings of the 4th International Symposium on Location and Context Awareness*, pages 70–87, Berlin, Heidelberg, 2009. Springer-Verlag.