APIs for USER-CONTROLLABLE LOCATION PRIVACY

Norman Sadeh, Ph.D.

Professor, School of Computer Science, Carnegie Mellon University, USA

sadeh@cs.cmu.edu - www.normsadeh.com

Chief Scientist, Zipano Technologies, Inc., USA

norm@zipano.com - www.zipano.com

INTRODUCTION

Over the past decade, the Mobile Commerce Laboratory in the School of Computer Science at Carnegie Mellon University has been piloting a number of context-aware applications and services [1,3,5]. From the very beginning an important part of our work has revolved around reconciling the demands associated with context awareness and privacy [5,10]. This short position paper summarizes some of the main findings of our work and provides a brief overview of how many of the results from our research have been encapsulated in a User-Controllable Privacy Platform for context-aware computing commercialized by Zipano Technologies [4].

Briefly, when it comes to sharing their location with different applications and services as well as with other users, our research has shown that:

- Users often exhibit complex privacy preferences
- Users often do not fully understand the implications of many of their privacy decisions and require better interfaces to help them make better decisions
- Users require auditing functionality that helps them review instances when their location information has been shared. When given this type of functionality, they report significant increases in comfort and a greater sense of control. These factors in turn contribute to the adoption of novel location-sharing applications and services.
- Users often do not manipulate default location privacy settings unless they are given functionality that prompts them to do so.
- Because different users often have different location privacy preferences, one-size-fits-all default policies are generally inadequate. Instead users need functionality that helps them select among a small set of default privacy personas
- Simple dialogues and user-oriented machine learning functionality can also help users incrementally refine their privacy preferences over time

The above findings shed light on the types of location privacy APIs that should be exposed to application developers. Without such APIs, application developers will find it very difficult to capture and enforce the complex location privacy preferences users often have – just as they had found it very difficult to develop location-based applications before the emergence of location APIs.

USERS LOCATION PRIVACY PREFERENCES ARE OFTEN COMPLEX

Figure 1 displays the location sharing preferences of 30 different users, when it comes to disclosing (Green) or not disclosing (Red) their location to other members of the Carnegie Mellon Campus community (see [2,6] for additional details). Each square represents a different user. As can be seen, a small number of users (10%) is never willing to share their locations with others, whereas an equally

small number of users (also 10% in this case) is always happy to share their locations with others. The majority of users falls in between with location sharing preferences that vary from one day to another and with time during the course of a given day.

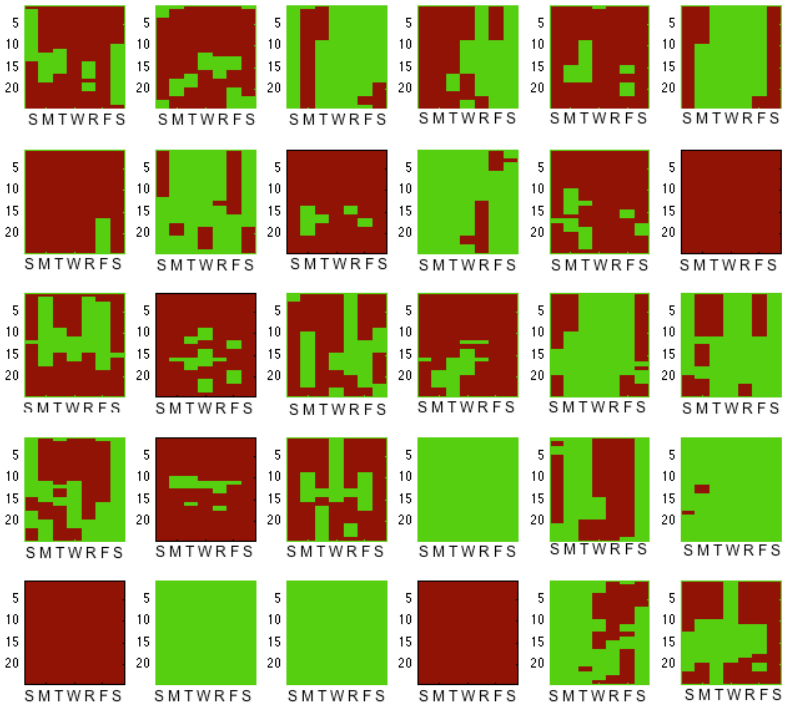

Figure 1. Privacy Preferences for sharing one's location with other members of the Carnegie Mellon Campus Community. Data collected from 30 different users based on day of the week and time of the day. Green means "share" and Red means "don't share". Each square represents a different user.

In studies where users were tracked for extensive periods of time and requested to indicate the conditions under which they would be willing to share their location with others, it became apparent that time of day and day of the week are but two factors in people's location sharing preferences. The user's current location is yet another important factor. A common privacy preference is of the type "I am willing to share my location with my colleagues, but only on weekdays, during working hours and only when I am on company premises". Other important factors have to do with the granularity at which location information is disclosed as well as the frequency of requests.

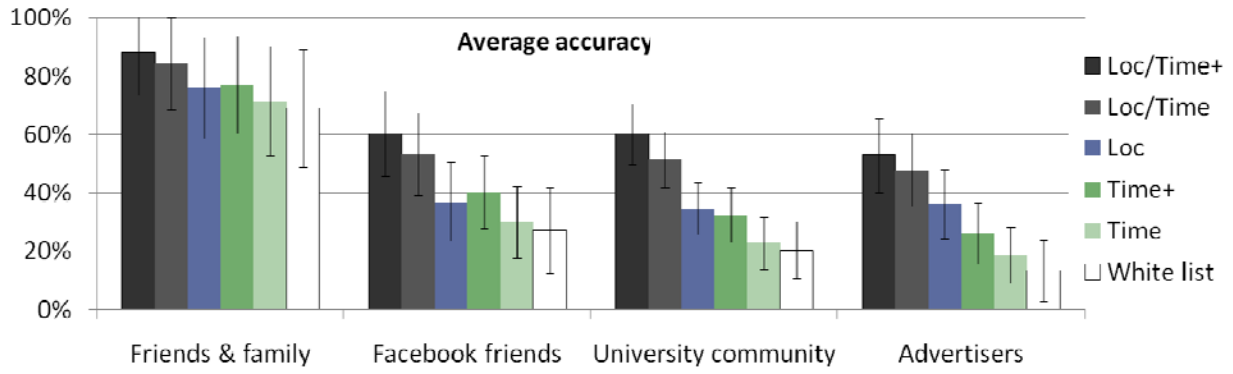

Figure 2. Accuracy with which one can capture a user's location privacy preferences when varying the expressiveness of location privacy settings exposed to that user. Sharing one's location with friends & family, Facebook friends, members of the University community, or advertisers.

Figure 2 summarizes the accuracy with which different levels of expressiveness in underlying policy languages capture people's actual location privacy preferences (see [2] for additional details). This figure differentiates between the following levels of expressiveness:

- White List: unconditional list of entities that can access one's location
- Time: Differentiating solely based on time of the day
- Time+: Differentiating between weekdays and weekends and between different times of the day
- Loc: Restricting access to one's location to situations where one is in a particular location
- Loc/Time: Restricting access to one's location subject to both time of the day and location restrictions
- Loc/Time+: Restricting access to one's location subject to both weekend vs. weekday, time of day, and location restrictions

As can be seen, with the exception of friends and family, simple white-lists are insufficient. More expressive settings have a major impact on our ability to accurately capture people's privacy preferences. This is best illustrated in the case of location-based advertising and sharing one's location with Facebook friends (see [2] for additional details).

QUANTIFYING EXPRESSIVENESS AND USABILITY TRADEOFFS

Given that users are only willing to invest so much time in configuring their privacy preferences, it is legitimate to ask to what extent the above results hold when taking into account user burden considerations.

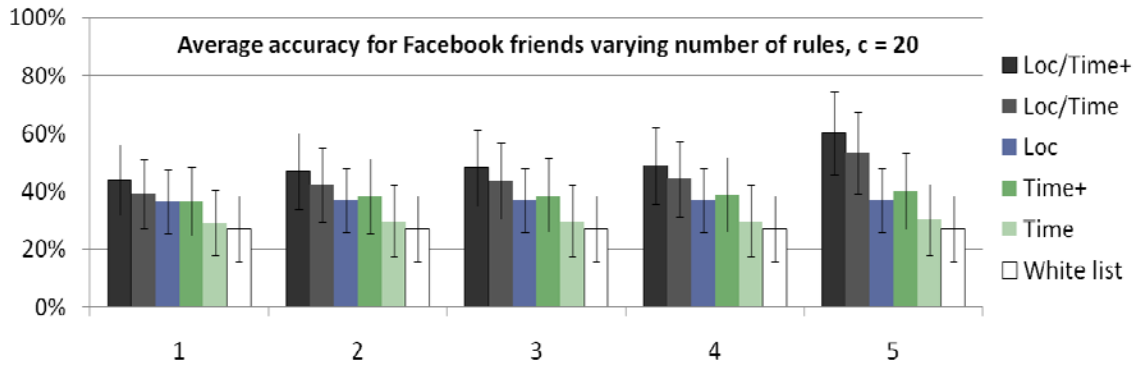

Figure 3. Quantifying accuracy versus user burden tradeoffs. The vertical axis shows the accuracy of the location privacy policies defined by users when varying the number of rules they can define (horizontal axis).

Figure 3 shows similar results when users are only willing to define a small number of location privacy rules. As soon as users define at least 2 privacy rules, the accuracy with which one can capture their location privacy preferences nearly doubles. The increase is even more dramatic for users willing to define 4 or 5 rules (see [2] for additional details).

DEFAULT PRIVACY PERSONAS CAN HELP

While a single one-size-fits -ll location privacy policy is often inadequate, presenting users with a small number of privacy personas they can choose from can help them take advantage of expressive settings without requiring them to explicitly select from a large number of options.

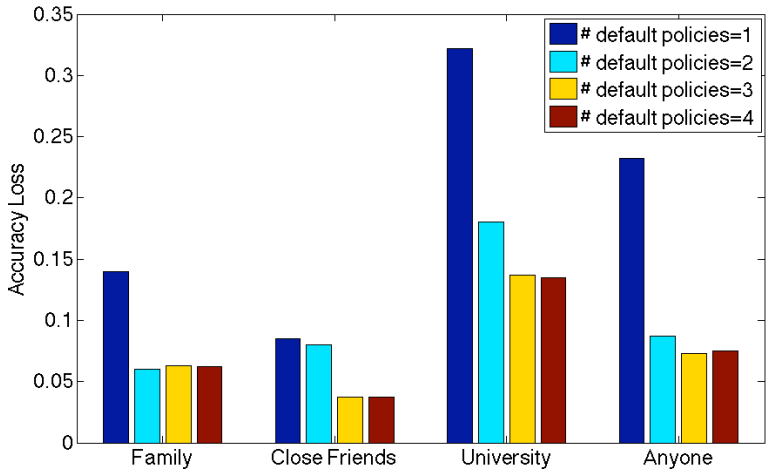

Figure 4. Accuracy obtained by varying the number of privacy personas (referred to here as "default policies") presented to users when it comes to configuring their location sharing privacy preferences for four different groups, namely family members, close friends, members of the university community, and anyone.

Figure 4 shows how presenting users with just two location privacy personas for configuring their privacy preferences when it comes to sharing their location with family members is sufficient (accuracy of 95%) and adding a third or fourth persona does not add any value. For close friends and members of the university community, a third persona helps increase accuracy over a situation where the user can only select between two personas. All in all, these results show that a small number of privacy personas is often sufficient to capture people's location sharing preferences, even if these preferences themselves are possibly fairly complex – additional details on how to learn canonical personas that are understandable by users can be found in [6].

AUDITING FUNCTIONALITY IS KEY TO USER COMFORT

A key to empowering users to effectively manage their location privacy preferences involves giving them access to simple auditing functionality. Using this functionality, users are able to review when their location has been shared and with whom. This in turn helps them better appreciate the behaviors their current privacy policies give rise to and empowers them to more effectively refine their preferences over time. Studies have shown that users report being much more comfortable accessing location-based applications when given auditing functionality [7]. In addition, experiments conducted with location sharing applications show that, when given this functionality, users tend (on average) to selectively relax their privacy preferences over time, eventually leading to more sharing [1,7] (see Figure 5 below) . In social networking contexts, where the value of an application derives from the amount of sharing it gives rise to, auditing functionality makes applications more valuable. The same has been measured for location-based advertising applications [2].

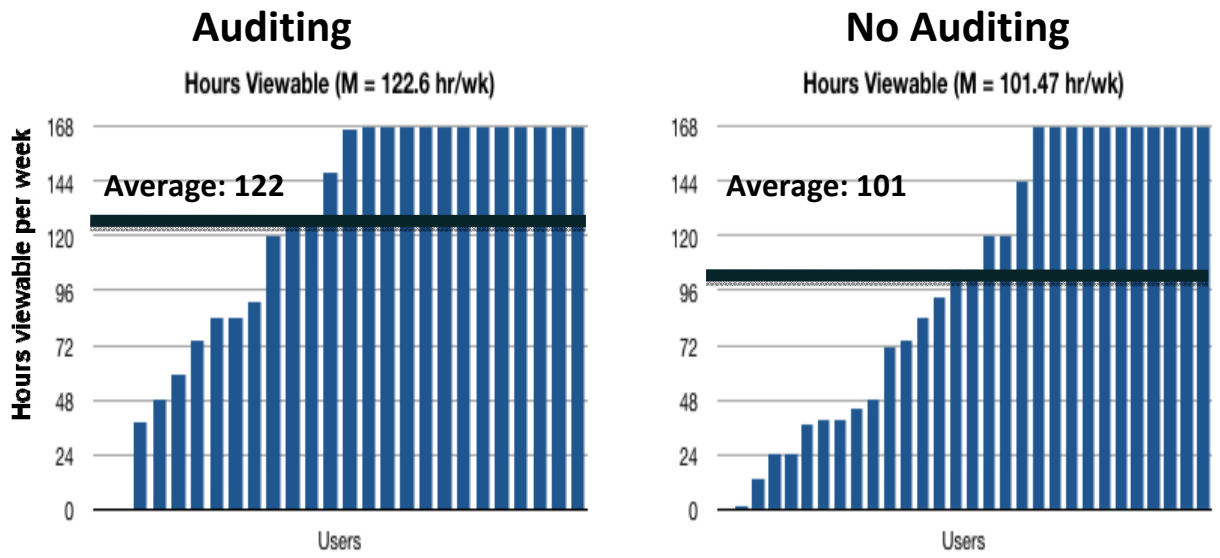

Fig. 5. Providing users with access to auditing functionality increases user comfort and also empowers users to more selectively relax their privacy preferences. Results from a month-long Facebook location sharing pilot involving nearly 60 users split into two conditions, one with auditing and one without.

DIALOGUE AND USER-ORIENTED MACHINE LEARNING CAN HELP USERS REFINE THEIR LOCATION PRIVACY PREFERENCES

Supplementing auditing functionality with dialogues and user-oriented machine learning technology to suggest to users how they could possibly improve their current privacy settings can help further improve user comfort by enabling users to converge towards privacy settings that better capture their true preferences [8]

ZIPANO'S USER-CONTROLLABLE PRIVACY PLATFORM (UCCP)

The above findings are encapsulated in a *User-Controllable Privacy Platform* (UCPP^{TM}) commercialized by Zipano Technologies. The UCPP exposes to application developers a layer of interoperability in the form of a set of RESTful web services APIs. It serves three primary roles that can be customized and selectively integrated into different platforms:

1. **A policy engine:** The essence of the UCPP is the storage, maintenance, and enforcement of rich user privacy policies.

2. **An interactive audit log:** The UCPP maintains detailed logs of all transactions. These logs can be used to help users refine their policies and generally help them get a better feel for who is requesting their location information and under which conditions (e.g. when or how often) . These logs should only be maintained for short periods of time (e.g. one or two weeks) to minimize privacy risks – just long enough to help users evaluate the impact of their current policies.

3. **A clearinghouse for context data:** The UCPP can accept and consolidate multiple concurrent streams of contextual data (e.g. user location or status), intelligently deciding which is the most current and relevant for a given user.

This functionality is packaged in the form of a highly scalable, low latency solution with APIs for privacy preference editing, location request processing, user request auditing, and flexible/open authentication (Fig. 6).

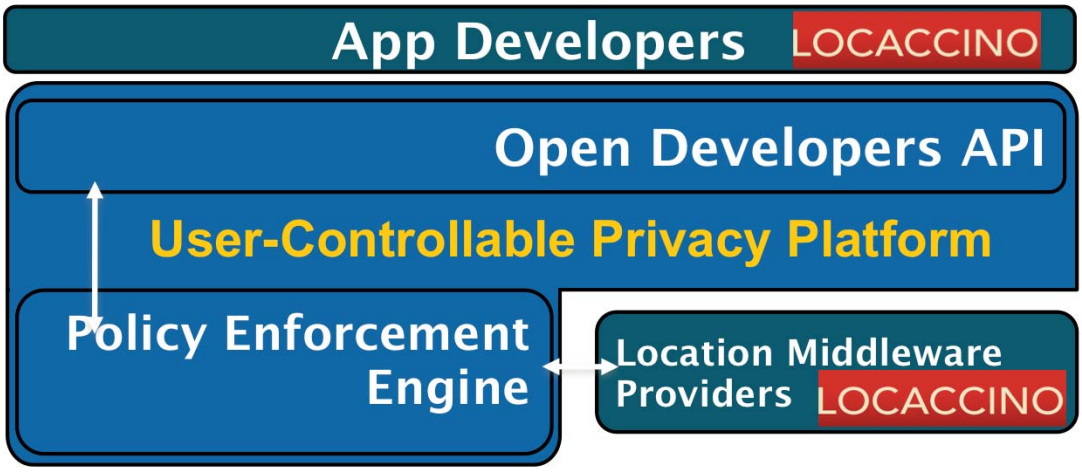

Figure 6 shows where the UCPP sits in the LBS development stack.

Locaccino [3] is an example of a Facebook application built on top of Zipano's UCPP. The application has been downloaded in over 130 countries. It runs on WiFi-enabled laptops (Windows and Macs) and is also available on Nokia's Ovi store (for Symbian phones) and on Android Market (Android phones).

SHORT SELECTION OF REFERENCES

[1] 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", Journal of Personal and Ubiquitous Computing, Vol. 13, No. 6, August 2009. http://www.normsadeh.com/file_download/8/People+Finder+PUC.pdf

[2] M. Benisch, Patrick Gage Kelley, Norman Sadeh, Lorrie Faith Cranor, "Capturing Location Privacy Preferences: Quantifying Accuracy and User Burden Tradeoffs", Carnegie Mellon University Technical Report CMU-ISR-10-105, March 2010, http://reports-archive.adm.cs.cmu.edu/anon/isr2010/CMU-ISR-10-105.pdf

[3] Locaccino Location Sharing Application – www.locaccino.org

[4] Zipano Technologies website – www.zipano.com

[5] Norman Sadeh, Fabien Gandon and Oh Buyng Kwon, "Ambient Intelligence: The MyCampus Experience", Chapter 3 in "Ambient Intelligence and Pervasive Computing", Eds. T. Vasilakos and W. Pedrycz, ArTech House, 2006. (Also available as Technical Report CMU-ISRI-05-123, School of Computer Science, Carnegie Mellon University)

[6] R. Ravichandran, M. Benisch, P. G. Kelley, and N. Sadeh, "Capturing Social Networking Privacy Preferences: Can Default Policies Help Alleviate Tradeoffs between Expressiveness and User Burden?", Proceedings of the 2009 Privacy Enhancing Technologies Symposium, August 2009. http://www.cs.cmu.edu/~CompThink/probes/papers/captsocnetwork.pdf

[7] J. Tsai, P. Kelley, P. Hankes Drielsma, L. Cranor, J. Hong, N. Sadeh "Who's Viewed You? The Impact of Feedback in a Mobile Location Applications", in Proceedings of the 27th annual SIGCHI Conference on Human Factors in Computing Systems (CHI 2009), April 2009. http://www.andrew.cmu.edu/user/jytsai/papers/paper0691-tsai.pdf

[8] P.G. Kelley, P. Hankes Drielsma, N. Sadeh, and L.F. Cranor, "User-Controllable Learning of Security and Privacy Policies", First ACM Workshop on AISec (AISec'08), ACM CCS 2008 Conference. Oct. 2008. http://www.normsadeh.com/file_download/61/CSS+AISec2008+camera+ready.pdf

[9] J. Tsai, P.G.Kelley, L.F.Cranor and N.M. Sadeh, "Location Sharing Technologies: Privacy Risks and Controls", to appear in "I/S: A Journal of Law and Policy for the Information Society". A shorter version of this article was presented at TPRC 2009. http://www.normsadeh.com/file_download/37/TsaiKelleyCranorSadeh_2009.pdf

[10] Gandon, F. and Sadeh, N., "Semantic Web Technologies to Reconcile Privacy and Context Awareness", *Journal of Web Semantics*. Vol. 1, No. 3, 2004. http://reports-archive.adm.cs.cmu.edu/anon/2003/CMU-CS-03-211.pdf

[11] J. Cranshaw, E. Toch, J. Hong, A. Kittur, N. Sadeh, "Bridging the Gap Between Physical Location and Online Social Networks", (2010). Conference Article, Accepted Collection: Proceedings of the Twelfth International Conference on Ubiquitous Computing.

[12] E. Toch, J. Cranshaw, P.H. Drielsma, J. Y. Tsai, P. G. Kelly, L. Cranor, J. Hong, N. Sadeh, "Empirical Models of Privacy in Location Sharing", (2010). Proceedings of the 12th ACM International Conference on Ubiquitous Computing. Copenhagen, Denmark, Sept 26-29, 2010

[13] Jialiu Lin, Guang Xiang, Jason I. Hong, and Norman Sadeh, "Modeling People's Place Naming Preferences in Location Sharing", Proc. of the 12th ACM International Conference on Ubiquitous Computing, Copenhagen, Denmark, Sept 26-29, 2010.

ACKNOWLEDGEMENTS

The research reported herein has been in part supported by NSF Cyber Trust grant CNS-0627513 and ARO research grant DAAD19-02-1-0389 to Carnegie Mellon University's CyLab. Additional support has been provided by Microsoft through the Carnegie Mellon Center for Computational Thinking, Google, FCT through the CMU/Portugal Information and Communication Technologies Institute, and through grants from FranceTelecom and Nokia. The author would also like to acknowledge the many members of the User-Controllable Security and Privacy for Pervasive Computing project and the Mobile Commerce Laboratory at Carnegie Mellon University who have contributed to research summarized herein and whose names appear in the publications cited in this paper.