

# Do I Do What I Say?: Observed Versus Stated Privacy Preferences

Kay Connelly, Ashraf Khalil\*, and Yong Liu

Computer Science,  
Indiana University, Bloomington, USA  
Connelly,yonliu@cs.indiana.edu

\*Abu Dhabi University  
Abu Dhabi, United Arab Emirates  
Khalil.ashraf@gmail.com

**Abstract.** This paper examines the use of surveys in measuring privacy concerns in ubiquitous computing environments. Two evaluation techniques are used to study the privacy concerns of sharing context information: a paper based survey and in-situ questionnaires. Results from the two techniques differ significantly, suggesting that surveys are not reliable in predicting privacy concerns regarding context-aware services. Further, the surveys are not consistently biased; for some information, people shared more in-situ than they predicted they would share in the survey, while for other types, they shared less.

## 1 Introduction

Ubiquitous computing applications, by definition, exist in the complex situation of our daily lives. Not only is an application's behavior dependent on the current context, but a person's reaction to an application often depends on the context as well. These unique characteristics introduce many challenges that make it difficult for surveys and polls to capture the full picture of a user's privacy concerns and preferences.

The use of surveys to study privacy concerns have been questioned by many researchers in the field of consumer market and ecommerce [1-4], where experimental studies have found a poor match between users' behaviors and their privacy preferences gathered by surveys. Bettman, Luce, and Payne attributed this discrepancy to decisions based on heuristics rather than rational consideration of all possible factors at play due to limited processing capacity (bounded rationality) [5]. Acquisti discussed different hypotheses beyond bounded rationality to explain the dichotomous privacy preferences and concerns between reported and actual behaviors [1].

Similar studies have yet to be performed in the field of ubiquitous computing and context-aware services; indeed, researchers have used surveys to report on privacy preferences in this domain [7]. In this paper, we examine the validity of using surveys as tools to study privacy in ubiquitous computing by measuring the differences in the level of privacy concerns reported by participants using traditional surveys and in-situ studies. If the techniques produce similar results, then researchers could rely on surveys because they are much cheaper to conduct than in-situ studies. However, a significant difference implies that surveys do not accurately capture participants' true privacy preferences, requiring the use of the in-situ approach.

## 2 Experiment Design

The test application in our study is a context-aware service aimed at minimizing cell phone interruptions by providing a potential caller with cues about the current context of the receiver; thereby helping the caller make a more informed decision about whether or not to call [6].

People perceive different kinds of personal information with varying degrees of sensitivity or privacy comfort [7]. Patil and Lai attribute the difference in the rate of context disclosure to the level of privacy comfort associated with a particular type of context information [8]. For this study, we identified four types of contextual information that participant's could choose to share with a potential caller: their *location*, their current *activity*, if they were *talking*, and if they were in the *company* of others.

We used the rate of context disclosure between different social relations as a measure of privacy concerns, looking at 6 distinct categories of social relations between caller and receiver based on Olson et al. [9]: *significant other*, *family member*, *friend*, *colleague*, *boss* and *unknown*. A full description of the experimental design and detailed results from the in-situ study are presented in an earlier paper [10].

We measured participants' privacy concerns for this context-aware service using both a survey and in-situ questionnaires. We used a survey format previously used to examine privacy concerns in pervasive computing [11]. For different locations, participants were given a table where the rows listed common activities performed in that location and the columns contained the 6 social relations. Participants were asked to indicate which context information they would share in each cell of the table, giving us data points for unique combinations of location, activity and relationship to caller.

For the in-situ portion of the study, we gave participants a Palm PDA running iESP, a general purpose Experience Sampling Method (ESM) application [12]. Throughout the day, participants received inquiries from the PDA prompting them to choose what context they would like to disclose to a potential caller. Participants were asked to assume the role of the receivers of a cell phone call. The caller assumed one randomly chosen role out of the 6 social relations. In addition to inquiring about participants' willingness to disclose different types of context information, every questionnaire included a list of questions about the current location, activity, if they were in a conversation, number of surrounding people, and the social relationship to the surrounding people. These questions allowed us to compare the in-situ answers to the survey under the same environmental conditions. The study lasted for 10 days during which participants were prompted 13 times a day at random intervals.

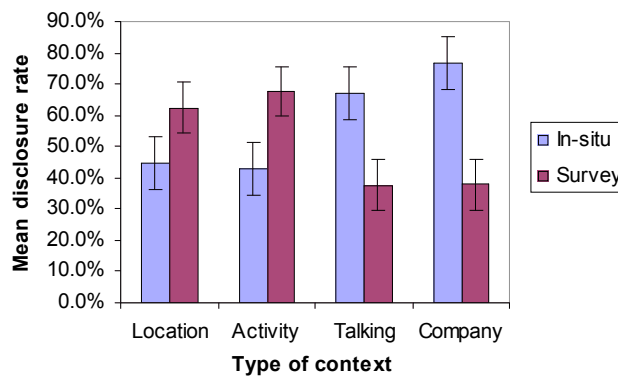
We recruited 20 participants equally divided between genders. Participants were ages 18-51 (average 24), were mostly students and all had either a full-time or part-time job. All had owned cell phones for more than a year (4.4 years on average) and regularly used their cell phones (daily average: made 5 calls and received 4).

## 3 Experiment Results

There were 2422 completed in-situ questionnaires, and 2520 entries from the survey. Given that we had no control over the location and activity of the participants

during the in-situ portion of the study, we extracted data points where the situations had the same conditions for both the survey and in-situ questionnaires, i.e. the same participant, location, activity, and caller role. There were 422 of these.

Out of the 422 answer pairs, only 31% of the answers from the two studies were identical. Each of the remaining answers contained at least one type of contextual information (location, activity, talking, or company) that a participant disclosed in the survey but did not disclose in-situ, or vice versa, given the same conditions. The maximum number of mismatched contexts was 4, which means that the participant incorrectly predicted what she or he would be willing to share for all four types of context. A one-sample t-test on the average number of mismatched contexts shows that the difference between the answers from survey and in-situ study is statically significant (mean = 1.59,  $t_{19} = 24.51$ ,  $p < 0.001$ ). This means participants gave different answers to the same questions depending on the type of evaluation technique.



**Fig. 2.** Disclosure rates for four types of context using the two different evaluation techniques

Figure 2 shows the disclosure rate for each type of context information that participants revealed using both the survey and the in-situ techniques. Disclosure rate for a particular context is the frequency at which participants released their context information relative to the total number of times they were able to release it. For each particular type of context, we found that participants' answers to the survey questions were significantly different from their answers in the in-situ study. Participants were much more willing to reveal their company ( $M_i = 76.8\%$ ,  $M_s = 37.8\%$ ,  $t_{19} = 5.74$ ,  $p < 0.001$ ) and talking ( $M_i = 67.1\%$ ,  $M_s = 37.7\%$ ,  $t_{19} = 4.34$ ,  $p < 0.001$ ) information in the in-situ study than they had predicted to disclose using the survey<sup>1</sup>. However, they predicted they would disclose more information in the survey than they did in the in-situ study for activity ( $M_i = 42.7\%$ ,  $M_s = 67.7\%$ ,  $t_{19} = 4.91$ ,  $p < 0.001$ ) and location ( $M_i = 44.6\%$ ,  $M_s = 62.5\%$ ,  $t_{19} = 3.05$ ,  $p < 0.01$ ). Our results suggest that, in the survey, participants tended to overestimate their privacy concerns associated with disclosing company and talking contexts and underestimate their privacy concerns associated with disclosing location and activity contexts.

<sup>1</sup>  $M_i$  is the mean disclosure rate in in-situ study.  $M_s$  is the mean disclosure rate in survey

## 4 Conclusions

We empirically demonstrate that surveys are not able to accurately reflect participants' true privacy concerns when using context-aware services. For some contexts, surveys overestimate user disclosure behavior, while for others, they underestimate. The in-situ technique may not be the ideal tool for measuring privacy preferences, but we expect it to fare better when compared to the survey tool. Our aim is to contribute to the understanding of using surveys as tools to study privacy concerns. The discrepancy between privacy attitudes (measured by the survey technique) and privacy behavior (measured by the in-situ technique) is well-documented in the field of online commerce and market research [1]. Our results extend such findings to the field of context-awareness and ubiquitous computing.

## Acknowledgments

This work was supported by NSF grant EIA-0202048 and a grant from the Lilly Endowment.

## References

1. Acquisti, A. and J. Grossklags, Privacy Attitudes and Privacy Behaviour: Losses, Gains and Hyperbolic Discounting. *The Economics of Information Security*, 2004.
2. Berendt, B., O. Günther, and S. Spiekermann, Privacy in E-Commerce: Stated Preferences vs. Actual Behavior. *Communications of the ACM*, 2005. **48**(4).
3. Junglas, I.A. and C. Spitzmueller. A Research Model for Studying Privacy Concerns Pertaining to Location-Based Services. In *38th Hawaii Inter. Conf. on System Sciences*. 2005.
4. Jensen, C. and C. Potts, Privacy Practices of Internet Users: Self-report versus Observed Behavior. *International Journal of Human Computer Studies*. 2005.
5. Bettman, J., M.F. Luce, and J.W. Payne, Constructive Consumer Choice Processes. *Journal of Consumer Research*, 1998. **25**: p. 187-217.
6. Pedersen, E.R. Calls.calm: Enabling Caller and Callee to Collaborate. In *CHI 2001*. 2001.
7. Goffman, E., *The Presentation of Self in Everyday Life*. 1959, Garden City, NY.
8. Patil, S. and J. Lai. Who Gets to Know What When: Configuring Privacy Preferences in an Awareness Application. In *CHI 2005*. 2005.
9. Olson, J., J. Grudin, and E. Horvitz. A study of preferences for sharing and privacy. In *CHI 05*. 2005. Portland, OR, USA.
10. Khalil, A. and K. Connelly. Context-aware Telephony: Privacy Preferences and Sharing Patterns. In *Computer Supported Collaborative Work (CSCW)*. 2006.
11. Barkhuus, L. and A. Dey. Location-Based Services for Mobile Telephony: a study of users privacy concerns. In *Interact 2003*. 2003.
12. Consolvo, S. and M. Walker, Using the Experience Sampling Method to Evaluate Ubicomp Applications. *Pervasive Computing Mobile and Ubiquitous Systems*, 2003. **2**(2): p. 24-31.