Context-Sensitive Intelligent Cueing

Julie S. Weber
Computer Science & Engineering
University of Michigan
Ann Arbor, MI 48109,USA
Tel: 1-734-763-0503
weberjs@umich.edu

ABSTRACT

The focus of my dissertation research is on the evaluation of different delivery formats of electronic notifications. In contrast to the breadth of work addressing the question of *when* a notification should be issued, my work addresses the less studied question of *how* to issue a notification. In particular, I explore the question of whether individual preferences for notifications are affected by delivery in different contexts, which differs from prior work, most of which focused on evaluation of notifications with respect to performance. Initial results indicate that (1) on-screen notifications can be categorized as highly or moderately intrusive relative to one another, and that (2) while general preference trends between the presentation of a notification and the associated context are exhibited within the data, the relative desirability of a given visual notification is widely varied among individuals and across contexts. Included in my planned work to address this issue is an in-situ study of both the desirability and effectiveness of various visual and auditory notifications, as well as the exploration of machine learning tools for tailoring the presentation and modality of a notification to a computer user's individual preferences.

ACM Classification: H5.2 [Information interfaces and presentation]: User Interfaces – Evaluation/methodology

General terms: Design, Experimentation, Human Factors

Keywords: Notification delivery, multi-modal notification

INTRODUCTION & RELATED WORK

The objective of my dissertation work is to design an intelligent, interactive cueing system that utilizes contextual information to tailor notifications to its users' individual preferences. There are multiple components to the contextual information surrounding computerized interactions, ranging from the content of an interaction (an alert that it's time for dinner versus a reminder about tomorrow's meeting with the project team) to the user's environmental setting at the time of interaction (whether the user is in the kitchen or in the office at a desk). In this research synopsis, I describe my work in laying a foundation for understanding users' notification-delivery preferences and my plans for the development of an adaptive notification system that is capable of performing preference learning through interaction, rather than by requiring explicit input and/or formulation of these preferences.

Numerous systems have been developed to provide intelligent cueing or notifications, both to people in a work environment, e.g., [1, 6, 8], as well as within the home, e.g., [10]. My work will expand on prior systems by composing different forms of context (notification content and a user's setting information) into a decision about the features of each notification. This differs from prior work that considered only one aspect of context (either some form of notification content or setting), e.g., [10], or a single notification feature (such as visual display or timing), e.g., [3].

Many of today's cueing systems assume that all users maintain the same set of preferences in all situations, or contexts. Some have evolved to consider that different contexts may warrant different forms of preferred interactions, e.g., [5]. However, many of these systems still ignore the possibility that users have individualized preferences for interactions.

My overarching research hypothesis is that [H1] computer users have highly individualized preferences regarding their patterns of interaction with computer systems, and further that these preferences vary across contexts as well as between individuals. Some systems have addressed this by allowing for manual specification of cueing features (such as timing), e.g., [4]. However, as a corollary to H1, I hypothesize that [H2] computer systems can automatically adapt their interactions to users' individualized preferences using techniques adopted from the field of machine learning, and further that this learning can be done unobtrusively, requiring only a small amount of feedback.

To investigate hypothesis H1, I conducted a pair of experimental user studies (Study 1 & Study 2) in an office environment. I first evaluated twenty participants' general acceptance of a set of eight visual reminders (Study 1). I then administered a follow-up study (Study 2) that examined preferences for the same set of visual reminders in the presence of different combinations of notification content and environmental setting information.

UIST'08, October 19-22, 2008, Monterey, California, USA
ACM 978-1-59593-975-3/08/10.

Results of the initial phase of the study indicate that there are certain reminder presentation styles that are generally perceived as more intrusive than others, and that a reminder's level of intrusiveness falls into one of two categories: highly intrusive or mildly intrusive, where the difference in average intrusiveness ratings across reminders in each category was statistically significant (at the $\alpha = 0.05$ level) within the resultant dataset. When reminder presentations were evaluated in the presence of explicit contextual information, I found that there was high variation among users regarding the types of visual reminders preferred in different situations. These results motivate future steps in my research plan outlined below: because preference variation among users and across contexts has been empirically validated, I can move forward in the process of designing a system that will adapt to these varied preferences.

I address my second hypothesis H2, which suggests that systems can be built to accurately learn user interaction preferences, through an additional set of experimental studies. The next study (Study 3) will act as a further examination of the extent to which context affects notification preferences, in a more naturalistic setting than Study 1 and Study 2. In this study, participants will be asked to perform a group of tasks, each of which awards its user a set of points, and users will rate each notification delivered in accordance with a given task. This user feedback will then be incorporated into a user model to be evaluated in a second phase of the experiment.

The final phase of my research plan (Study 4) involves the design of an adaptive reminding system that incorporates real-time preference feedback into its learned user model. This study will poll a set of electronic calendar users on their acceptance of visual reminders explicitly tied to their calendar events. Similar to Study 3, in one phase of the study, reminders will be generated by an adaptive system that incorporates a learned set of preferences (garnered in an initial phase of the study) into a user model of each participant. Acceptability ratings will be compared to a separate phase in which an alternate model is used.

This will require the use of a learning mechanism that can accept multiple types of feedback, allowing the system to explore a variety of means for eliciting reminding preferences. More specifically, I plan to incorporate a variant of the active learning algorithm that I developed in prior work on schedule preference learning [11]. That algorithm adopted a context-driven approach in its online decision-making process, suggesting a set of scheduling options to its user in an effort to balance efficient learning and user satisfaction. In that work, "context" was associated only with the user's calendar, whereas I will employ this technique to consider other features of a user's setting, as well as the content of the reminder itself. Another form of feedback that may be included in this study is explicit seeding of the learning system: users can indicate those preferences that they know in advance, essentially pre-tuning the system to their liking.

The remainder of this paper will further describe the two studies that have been performed, as well as the design for Study 3, which will introduce the auditory modality and more realistic contextual scenarios.

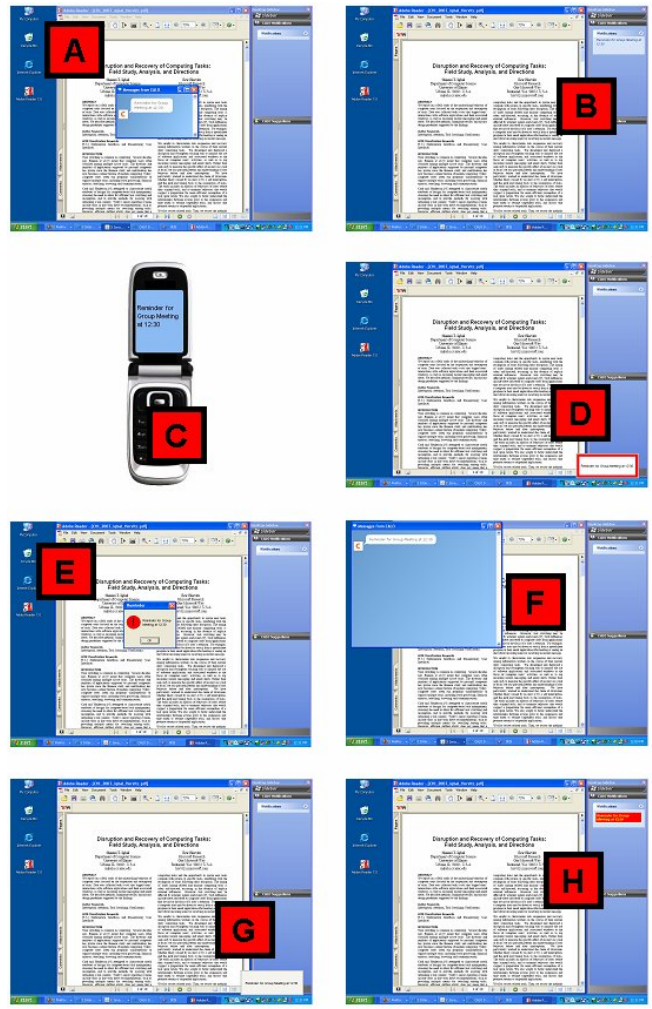

Figure 1: Each of the the eight reminder presentation styles. Style *A* is a centered dialogue window, *B* is a message on the persistent sidebar, *C* is a text message on a mobile phone, *D* is a newly opened message window enhanced with red, *E* is a centered, modal, red dialogue, *F* is a dialogue window in the upper corner, *G* is a newly opened message window, and *H* is a message on the sidebar in red.

REMINDER DISPLAY: INITIAL USER STUDIES

The objective of the first user study I conducted was to examine the relationship between reminder presentation style and its associated level of perceived intrusiveness. My primary hypothesis for this study was that: when a visual reminder is examined completely independent of associated context, different users will ascribe similar levels of perceived intrusiveness to the same reminder presentation style, where presentation style is described in terms of visual display features such as color and location on the screen. A second study was conducted to evaluate the acceptability of various visual reminders in the presence of surrounding contextual information. Here I hypothesized that preferences will be both context-dependent and highly individualized. These studies were part of a larger initiative to learn about general time management preferences [12].

Study 1—Experimental Design

The study included 20 participants (15 male, 5 female) between the ages of 18 and 55. As a result of various constraints on the presentation of reminders through the different interfaces that were considered for this study, a set of eight *reminder presentation styles*, including a reminder delivered as a text message on a mobile phone, were chosen for inclusion and displayed in Figure 1. Please see [13] for more details about the features of these presentation styles.

From initial informal interviews with potential system users, it was found that annoyance is the primary factor that determines whether or not the user of an intelligent personal assistant actually continues to interact with (and accept the interactions of) the system; this is consistent with the claims of [7]. Consequently, I make use of an annoyance scale to evaluate alternative interaction patterns. For this initial study, all of the participants were asked to consider eight screen shots of reminder presentation styles (see Figure 1) and rate each presentation style by the amount of disruptive annoyance that it would cause. Disruptive annoyance was explained to participants as the degree of disruption caused in a user's current task; this is in contrast to the annoyance that would be caused by not noticing a reminder. The annoyance scale was presented to each participant as a number line from 0 to 10. Lower values represented lower levels of annoyance.

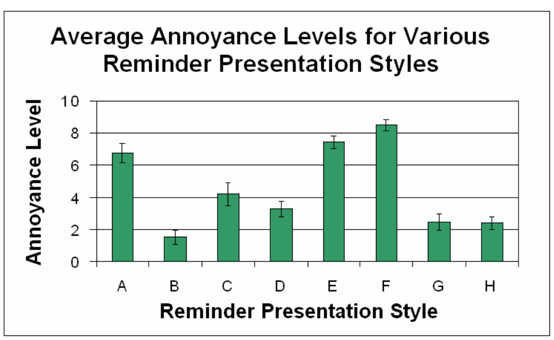

Figure 2: Mean annoyance ratings for each of the eight reminder presentation styles.

Study 1—Results

The mean annoyance values for each of the eight reminder presentation styles are displayed in Figure 2. Upon detailed analysis of these results using a repeated measures analysis of variance test and post-hoc pairwise evaluation with a Bonferroni adjustment, I found that instead of merely creating a continuous scale of increasing annoyance, the reminder presentation styles formed a pair of equivalence classes based on pairwise variations in annoyance ratings. More specifically, reminder presentation styles A, E and F formed one equivalence class, in which relatively high levels of annoyance were ascribed, while B, C, D, G and H formed a second class with lower levels of ascribed annoyance, or intrusiveness, with $\alpha = 0.05$.

Study 2—Experimental Design

An additional study was conducted to measure the acceptability of various reminder presentation styles in the presence of contextual information. The same set of eight screen shots depicting various reminder presentation styles was used again for this phase of the study.

The conditions for this experiment include the eight reminder presentation styles and a set of eight scenarios, designed to introduce a variety of tasks in which a user might be engaged, and in a variety of environmental settings, at the time a reminder is issued. Each scenario represents a particular level of utility (comprised of urgency and importance of a reminder and its associated event) and focus required for the suggested current task. For example, the following scenario: "You are sorting through old emails in your inbox when you are reminded about your meeting with the Director of [your company] in five minutes" would be classified as having high importance (meeting with the Director), high urgency (five minutes from now), and a low level of focus on the given task (email sifting); whereas another scenario, such as: "You are in your office, meeting with your boss, when a reminder arrives for a meeting next week that you had no interest or intention of attending" would be classified as low importance (the meeting is of no interest), low urgency (it is happening next week), and high focus (you're meeting with your boss).

Study 2—Results

A test of model effects was conducted on the data acquired in this study, and findings were as intended: user ratings are significantly affected by both scenario (and the associated contextual features) and reminder presentation style ($p < (0.001, 0.000)$). Further the estimated marginal means over scenarios shows a distinct trend in average user rating of reminder presentation styles between the first (high-utility, low-focus) scenario and the last (of low-utility, high-focus). This summarizes the percentage of positive ratings for each scenario, where the first (Scenario 1) represented high importance, high urgency and low current task focus, and the last (Scenario 8) represented low importance, low urgency and high focus. However, individual preferences over the eight reminder presentation styles were widely varied within scenarios and between participants. In other words, while there was high overall acceptance of visual reminders in certain contexts when compared with others, there was no way to predict an individual user's preferences for a specific reminder presentation style in a given context.

Study 3—Experimental Design

The objectives of my upcoming experiment are to (1) conduct a controlled laboratory study that assesses the effect of audio and visual notification display techniques in different contexts on people's preferences for such notifications, (2) automatically learn these context-based preferences for the different types of notification delivery styles, and (3) compare with prior work focused on notification timing: to analyze the impact of notification display techniques on annoyance or irritability, and on performance on a set of goal-driven tasks.

There will be a single "base task" that is configured to require either high or low attentional focus (dependent upon the number of subtasks to be performed) at any point in a session of the study. There will be two "interrupting tasks" that will become active at various points. A participant will be allotted a fixed amount of time T for each interrupting task. At any given moment, an interrupting task may be activated for M minutes (with $M \geq T$); a participant will have the option of performing the task at any time within the M minute window, and the task will be deactivated after T minutes, or at the end of the time window (whichever comes first). A schedule of activation times for the interrupting tasks will be provided to each participant at the beginning of the study.

When an interrupting task becomes activated, a notification will be issued to alert the participant about the upcoming activation time. Participants will acknowledge each notification by closing it either via a 'thumbs up' or 'thumbs down' symbol (indicating positive or negative feedback, respectively), in place of the traditional close button. A participant's objective is to maximize his or her total score by performing the base task, which is always active, and the interrupting tasks whenever they are active. Feedback will be used to associate preferences for notification delivery with notification content and contextual features of a user's environment.

RESEARCH PLAN

I will carry out the two additional studies (Study 3 and Study 4) in the coming months. I will analyze the results of both studies by comparing participants' acceptability ratings for each form of the adaptive system with their acceptability of the initial, non-adaptive system. I will also evaluate the chosen learning mechanisms. In particular, once the initial user (training) data is acquired, I plan to compare a regularized regressive algorithm such as LASSO [9] to other regressive techniques as well as classification algorithms (which would require an alternate learning-problem formulation) to evaluate whether the formulation and algorithm originally selected are most successful. At the time of the UIST conference, I will have acquired a training set of user data, as well as have performed a complete pilot study, and preliminary analysis of the data will have been completed. After the conference, I will complete the user studies and analysis and will spend the following months writing up the results of each study and completing my dissertation. I expect to defend my thesis next summer.

CONCLUSION

People likely have different preferences for receiving computerized notifications in different contexts. These preferences may differ between people and/or across contexts. If this is indeed the case, a procedure for learning people's contextual notification preferences would be a useful tool for today's computerized notification systems. Further, if the timing of a notification cannot be adjusted, then it may be particularly important for such a system to have the ability to modify the manner in which the notification is displayed on the screen or via the audio channel.

REFERENCES

1. Dey, A. K., and Abowd, G. D. 2000. Cybreminder: A context-aware system for supporting reminders. Second International Symposium on Handheld and Ubiquitous Computing 1927:172–226.

2. Gajos, K. and Weld, D. S. 2005. Preference Elicitation for Interface Optimization. In Proceedings of *User Interface Software and Technology (UIST*), Seattle, WA.

3. Gluck, J.; Bunt, A.; and McGrenere, J. 2007. Matching attentional draw with utility in interruption. In Proceedings of CHI 2007, 41–50.

4. Growl. 2007. http://growl.info.

5. Horvitz, E.; Breese, J.; Heckerman, D.; Hovel, D.; and Rommelse, K. 1998. The lumiere project: Bayesian user modeling for inferring the goals and needs of software users. In In Proceedings of *Uncertainty in Artificial Intelligence (UAI)*, 256–265.

6. Lamming, M., and Flynn, M. 2005. Forget-me-not: Intimate computing in support of human memory. In Proceedings of *User Interface Software and Technology (UIST)*, Seattle, WA.

7. J. Levitt. Internet zone: Good help is hard to find. Information Week: Listening Post, 2001. http://www.informationweek.com/835/35uwjl.htm.

8. Myers, K.; Berry, P.; Blythe, J.; Conley, K.; Gervasio, M.; McGuinness, D.; Morley, D.; Pfeffer, A.; Pollack, M.; and Tambe, M. 2007. An intelligent personal assistant for task and time management. AI Magazine 28(2):47–61.

9. Tibshirani, R. 1996. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Socity, Series B 28(1):267–288.

10. Vastenburg, M. H.; Keyson, D. V.; and de Ridder, H. 2007. Considerate home notification systems: A field study of acceptability of notifications in the home. Personal and Ubiquitous Computing.

11. Weber, J. S., and Pollack, M. E. 2007. Entropy-driven online active learning for interactive calendar management. In Proceedings of *Intelligent User Interfaces* (*IUI*) 2007.

12. Weber, J. S., and Yorke-Smith, N. 2008. Time Management with Adaptive Reminders: Two Studies and Their Design Implications. *ACM SIGCHI Workshop on Usable Artificial Intelligence*, Florence, Italy.

13. Weber, J. S., and Pollack, M. E. 2008. Evaluating User Preferences for Adaptive Reminding. *Work-In-Progress, ACM SIGCHI*, Florence, Italy.