txt 4 l8r: Lowering the Burden for Diary Studies Under Mobile Conditions

Joel Brandt

Stanford University HCI Group

353 Serra Mall

Stanford, CA 94305

jbrandt@stanford.edu

Noah Weiss

Stanford University HCI Group

353 Serra Mall

Stanford, CA 94305

noahweiss@stanford.edu

Scott R. Klemmer

Stanford University HCI Group

353 Serra Mall

Stanford, CA 94305

srk@cs.stanford.edu

Abstract

We present and evaluate a new technique for performing diary studies under mobile or active conditions. Diary studies play an important role as a means for ecologically valid participant data capture. Unfortunately, when participants are asked to capture data while mobile or active, they are often unwilling or unable to invest time in thorough, reflective entries. Ultimately, this leads to lowered entry quality and quantity. The technique presented here suggests the capture of only small *snippets* of information in the field. These snippets then serve as prompts for participants when completing full diary entries at a convenient time. We describe how our system automates collection of snippets via SMS (text), MMS (picture) and voicemail messages and later presents these snippets for full entry elicitation. We then present results from a preliminary evaluation of this technique.

Keywords

Diary study, field work, mobile data capture, mobile computing, text messaging

ACM Classification Keywords

H.5.2 [Information interfaces and presentation]: User Interfaces – Evaluation/methodology, theory and methods, user-centered design.

Introduction

The diary study technique is commonly employed to capture data *in situ* [5]. In diary studies, participants are responsible for data capture, which has both advantages and disadvantages. Because there is no observer present to affect participant behavior, diary studies potentially increase ecological validity. They also reduce the per-participant burden on the researcher because the researcher does not need to be present for data collection to occur. Shifting the burden of data collection to the participant, however, can lead to omissions in the events reported: occurrences which seem trivial to the participant (but may not be to the researcher) or which occur at inconvenient times may go unreported.

As Palen and Salzman suggest, this problem often arises when participants are asked to complete diary entries under mobile or active conditions [4]. Motivated by a desire to perform a need-finding study in the mobile computing domain, and inspired by current work on improving diary study techniques by employing various medias and tools [1, 2, 4], we have developed and evaluated a technique for lowering the *in situ* burden of performing diary studies.

The remainder of this paper proceeds as follows. We first describe our technique in detail, and discuss the implementation of a system which supports this technique. We then present our evaluation criteria, and discuss the methodology used to evaluate our technique. Finally, we present the results of our evaluation, and discuss future work.

The Snippet Technique

Our technique centers on the *in situ* capture of *snippets*: bits of text, audio, or pictures captured in a matter of seconds. Participants record and transmit these snippets to a server using standard mobile phones through SMS or MMS messaging, or by leaving a voicemail message. Then, at a convenient time, participants access a website to review their snippets and complete thorough, structured diary entries. A diagram of our system is shown in figure 1, and a view of the structured web interface with several snippets is presented on the first page.
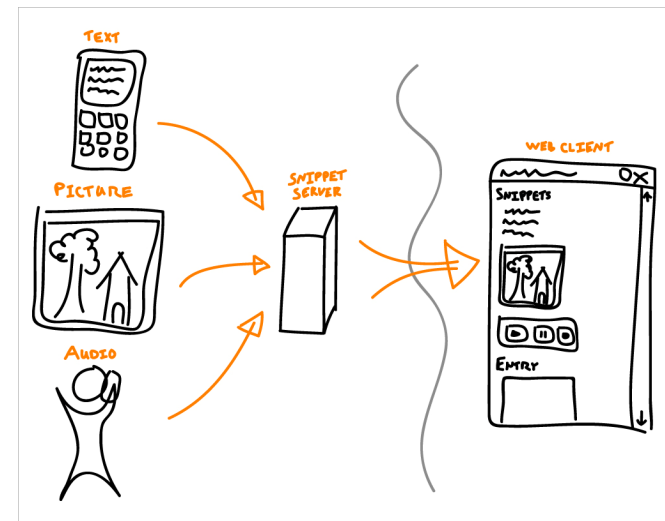
figure 1. The system diagram for our implementation of the snippet technique.

We believe this approach has a number of advantages which make it amenable to participant data capture under mobile conditions. First, it lowers the *in situ* data entry burden by an order of magnitude (from minutes

to seconds). Second, it leverages a device that most participants have with them whenever mobile, and that they know how to use—no special software or carrier services are required. Third, the various input modalities (text, picture, audio) allow participants to choose a media that they feel most comfortable with and/or fits the situation most appropriately.

System Implementation
In our implementation, SMS and MMS collection takes place using the NowSMS gateway and a Sierra Wireless 860 GSM modem. Voicemail collection takes place through Skype and a custom voicemail application written using the Skype Java API for call answering, Virtual Audio Cables for sound I/O, and Lame for MP3 encoding. All of these services provide the sender's phone number, which we use for automatic routing of the snippets. The snippets are stored in a MySQL database, and the web front end is written in PHP running on Apache. A custom Flash application allows streaming playback of voicemails in the web browser.

Evaluation

Through two studies, we evaluated the following four hypotheses:

H1 Our snippet technique leads to longer entries in a shorter time when compared to current *in situ* techniques for mobile data capture.

H2 Entry accuracy is not significantly affected through the use of the snippet technique.

H3 Participants are willing to contribute more "trivial" events when using the snippet technique than when using current *in situ* techniques.

H4 Media choice for the snippet will be dictated by both content and context of entry.

Each study was two weeks long, and in both studies, 15 participants were recruited. Nine participants (5 female, 4 male) completed study one, and 14 participants (9 female, 5 male) completed study two. All participants were undergraduate or graduate students at our institution. 187 entries were collected in Study 1, and 255 entries were collected in Study 2. All participants were asked to complete structured diary entries about mobile computing needs. At the conclusion of both studies, we conducted 15-minute exit interviews with each participant to elicit feedback about the study techniques employed.

Study 1 – Comparing the snippet technique to in situ methods
In this study, participants were randomly assigned to groups *a* and *b*. Initially, group *a* was asked to use an *in situ* technique that allowed them to complete structured entries using any variety of media they desired: written text, voice, or pictures. The only restriction was that the entire entry had to be completed *in situ*. Group *b* was asked to use our technique. After one week, the groups swapped techniques.

Study 2 – Effects of media in snippet recording
In this study, all participants began using our snippet technique, but were limited to recording text snippets. After one week, participants were allowed to use all medias for snippet entry.

Results

H1 – ENTRY LENGTH AND TIME SPENT
Both groups from Study 1 wrote statistically significantly longer entries when using our snippet technique compared to their *in situ* entries (*In*

Situ/Web difference *P*-value: 0.028, one-sided paired *t*-test, 8 *d.f.*). Study 2, which used our technique exclusively, also created entries with an average length longer than the *in situ* entries from Group 1a and 1b.
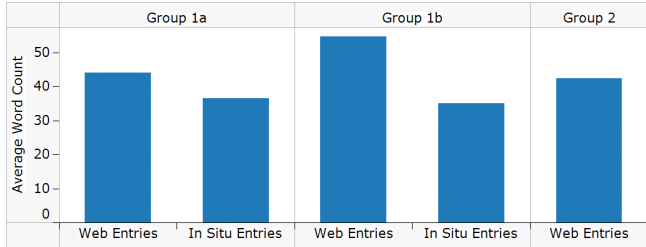
figure 2. Average word count for paper and *in situ* entries, broken down by group.

While the difference in entry completion time is not statistically significant—likely due to the large variance across users—our results coupled with qualitative feedback suggest that web entries take no longer to complete than *in situ* entries (*In Situ*/Web difference *P*-value: 0.16, two-sided paired *t*-test, 7 *d.f.*).
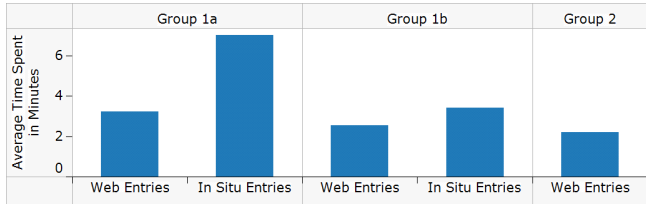

figure 3. Average time spent for paper and in situ entries, broken down by group.

In exit interviews, participants reaffirmed the conclusions from our data analysis. When asked which technique was less time consuming, participant s9 said, "I was much faster at responding to the questionnaire online than on paper." S9 expressed his early skepticism about the potential invasiveness of our technique, but went on to say, "It definitely did not interfere with my life, even though I thought it would initially." He then offered this explanation for why he made longer entries with our technique: "It was great to have the option to fill it in online later, because I could do it anytime I wanted and make a fuller entry." Both our usage data and participant interviews support our hypothesis that our technique would take less time and result in longer entries than *in situ* techniques.

H2 – ENTRY ACCURACY
One of the primary arguments for making complete entries *in situ* is that it eliminates the time window where participants potentially forget important details. Our results suggest that prompting with snippets helps mitigate this concern. At the end of each web entry, participants responded to the Likert Scale prompt: "All of the data contained in this entry is accurate." Participants selected either "I strongly agree" or "I agree" on over 87% of the entries they completed.
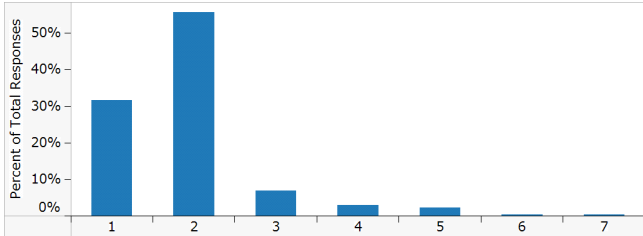
figure 4. For each Likert Scale option, the percentage of entries participants selected that option.

During the exit interviews, we observed that participants developed very individualized snippet recording strategies. However, several themes did
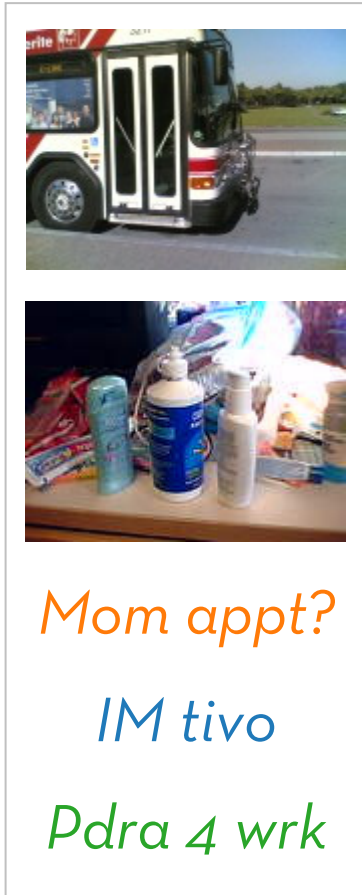
figure 5. Examples of text and picture snippets left by study participants.

emerge. For example, several participants often used text and photos together. Photos were used to establish context and text was used to summarize the specific need. S28 said that this "combination of photos and text helped my memory the most." Another common technique was to create a keyword system to give meaning to short text snippets. S9 described this as, "For the texts, I would usually have keywords that I entered and the needs were separate enough that I could differentiate them based on the keywords. Even though the snippets were small…I would remember what the exact need was." Decreased entry accuracy with our technique was one of our principal concerns, but participant feedback indicates that remembering the intent of snippets when completing the web entry was a non-issue.

H3 – "TRIVIAL" CONTRIBUTIONS

While it is difficult to assess the "trivialness" of contributions in a quantitative way, qualitative evidence supports the hypothesis that subjects are willing to submit more "trivial" events when using the snippet technique. S3 remarked, "A couple [of entries wouldn't have been made without snippets.] Smaller things were too trivial to open up a diary for." A frequent complaint about doing entries *in situ* was that many active situations made it difficult to complete a full entry. S15 said, "Some of the [*in situ*] entries I couldn't do at the time because I couldn't stop the car, or if I was talking to someone." S9 echoed s15's complaint: "Most of the time, I would write the paper entry after it happened because most of my mobile needs happen when I'm active, like when I'm walking or driving." It is clear from our interviews that our technique, in comparison to writing full entries *in situ*, lowers the threshold for creating entries in situations where either the need is

so small that writing a full entry "isn't worth it" or if the need occurs while the participant is active.
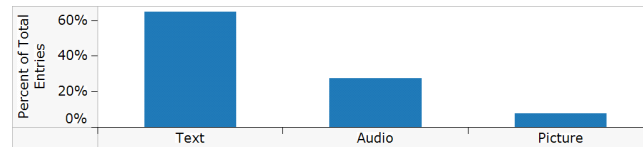
H4 – MEDIA CHOICE FOR SNIPPETS

Initially, we thought that participants would choose snippet media based on the content and context of the entry. Our usage data and interviews instead point to individual preferences as the primary motivating factor when selecting snippet media type. When given the choice of using text, photos, or audio, no participant used all three options. One-third of the participants used exclusively one type of media, with five using only text and two using only audio. As figure 6 depicts, participants usually had a strong preference for one type of media.
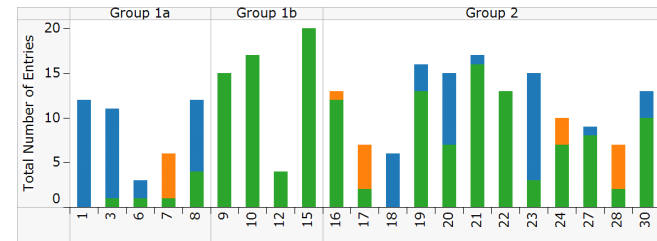
figure 6. The number of snippet entries, by group, user, and media type. Photos, Audio, Text

In the exit interviews, participants often felt strongly about the media type they preferred. S16 said, "Text was easier and faster. I only had to write one or two words to remind myself." S23, on the other hand, hated text: "Voice was so much easier. I hate the T9 and I'm slow with the original text." Although there were proponents of audio and photos, overall media usage was heavily weighted towards text.

Figure 7. Distribution of snippets media choice for Study 2.

Participants were reluctant to use audio often because they felt awkward leaving a "strange sounding" voicemail in front of other people. S28 described this when he said, "I didn't send voicemails because I felt goofy doing it." The lack of picture usage seems primarily due to the difficulty of sending picture messages and the poor photo quality of most camera phones. S15 said, "I didn't do picture messaging because the camera quality on my phone is atrocious." But in the future, many participants said they could see themselves using photos when the technology improves. S15 continued to say, "If the quality was better, I would have used it more."

It is also interesting to note that when full *in situ* entries were required in Study 1, only one entry (out of 122) used a media other than written text.

Future Work
We believe that the techniques presented here show a great deal of promise for both lowering the work load and raising the validity of field work involving participant data capture in a variety of disciplines. Indeed, several researchers at our institution have expressed interest in using our system in their upcoming studies. We plan to use this opportunity to perform a researcher-side evaluation of our technique.

To this end, our next goal is to make our system generalizable and deployable. To address generalizability concerns, we are working closely with potential users of our system to identify desirable features. One researcher, for example, is interested in adapting our technique to perform a snippet-based experience sampling method (ESM) study [3]. To address deployability concerns, we currently exploring the use of VMware Virtual Appliances to enable turn-key deployment.

Additionally, we are currently preparing to use this method for a larger scale need-finding study in the mobile computing domain. As we will be using a larger, more diverse population for this study, we will gain further participant feedback on this technique that will likely augment the data presented here.

References
[1]   Carter, S. and J. Mankoff, Momento: Early Stage Prototyping and Evaluation for Mobile Applications, Technical Report CSD-05-1380, UC Berkeley 2005.

[2]   Carter, S. and J. Mankoff, When Participants do the Capturing: The Role of Media in Diary Studies, in Proc. of SIGCHI. 2005.

[3]   Consolvo, S. and M. Walker. Using the Experience Sampling Method to Evaluate Ubicomp Applications. IEEE Pervasive Computing 2(2). pp. 24-31, 2003.

[4]   Palen, L. and M. Salzman, Voice-mail diary studies for naturalistic data capture under mobile conditions, in Proc. of CSCW. 2002.

[5]   Rieman, J., The Diary Study: A Workplace-Oriented Research Tool to Guide Laboratory Efforts, in Proc. of SIGCHI. 1993.