

# MAESTRO: Orchestrating User Behavior Driven and Context Triggered Experience Sampling

Alexander Meschtscherjakov  
alexander.meschtscherjakov@sbg.ac.at

Wolfgang Reitberger  
wolfgang.reitberger@sbg.ac.at

Manfred Tscheligi  
manfred.tscheligi@sbg.ac.at

Christian Doppler Laboratory for "Contextual Interfaces",  
HCI & Usability Unit, ICT&S Center, University of Salzburg, Salzburg, Austria

## ABSTRACT

In this paper we present "MAESTRO" a novel concept to study behavior in realistic environments based on the experience sampling method (ESM). It constitutes a variation of ESM by giving researchers the possibility to dynamically and remotely "orchestrate" experience-sampling studies for evaluating usage behavior in various different contexts. MAESTRO allows user behavior driven and context triggered experience sampling by using a client-server architecture. So far a variation of different computerized ESM tools have been introduced capable of delivering questionnaires triggered by contextual information on mobile devices. We propose to enhance ESM by not only using context and user actions as a trigger for ESM questions, but also by exploiting long term user behavior and usage patterns for shaping personalized ESM questions to different types of users. Our concept lays the foundation of measuring user experience not only in a mobile but in a magnitude of different contexts. As a proof of concept we present an implementation of our approach on BlackBerry devices along with initial results of a case study validating this approach.

## Author Keywords

Experience sampling method, in-situ evaluation, user experience measurement, ESM tool, methodology.

## ACM Classification Keywords

H.5.2 User Interfaces: Evaluation/methodology

## INTRODUCTION

The evolution of computing from single user desktop computer located in offices to mobile and ubiquitous devices that permeate the entire environment has sparked the need for in-situ evaluation methods. Theoretical works like Hutchins' "Cognition In the wild" [3] have encouraged

researchers to leave the lab and conduct their studies in the field. Therefore psychological self-reporting methods like the Experience Sampling Method (ESM) [5] have been adopted to a wide range of research fields. ESM is a psychological method to capture participants' experiences in everyday life. This is achieved by triggering self-reports (e.g. diary entries, predefined questions on paper notes) from the participants.

Recently there has been a shift from paper-and-pencil ESM to more sophisticated computerized tools. Le et al. [6] have explored and compared various early ESM tools (ESP, iESP and PMAT) for the behavior research method community. Since ESM is a method for collecting information not only about the content of the daily life of individuals, but also tries to capture different context factors, the requirements for these tools are twofold: on the one hand researchers have to capture input from the users' subjective experience and on the other hand they need to record objective information about the context. Furthermore context information can also be used to trigger dynamic samples. This possibility was first identified by Intille [4] and demonstrated in the Context-Aware Experience Sampling toolkit (CAES). This approach can be referred to as *context-contingent* ESM. Our approach is to enhance ESM by not only using contextual information or user actions as a trigger for ESM questions, but also by exploiting users' usage behavior over a period of time for shaping personalized ESM questions to different types of users. This means that user behavior driven and context triggered ESM enables the researcher:

- to trigger ESM questions based on the recorded user behavior in the past and dynamically adapt these ESM questions whenever the user changes his / her behavior.
- to log and monitor in real-time comprehensively and flexible user behavior together with meaningful context information.

## RELATED WORK

Experience sampling has been applied to various studies of mobile user behavior. Mehl et al. [7] invented the Electronically Activated Recorder (EAR), which automatically recorded audio samples thorough the day in order to reflect people's daily lives. Another recent project

---

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. For any other use, please contact the Measuring Behavior secretariat: [info@measuringbehavior.org](mailto:info@measuringbehavior.org).

is TUMCAT [11] a test bed for user experience for mobile context-aware applications. Beside the previously mentioned CAES [4] different experience sampling tools have been presented. SocioXensor [10] as an example is a toolkit, which exploits hardware sensors and software capabilities located on the mobile device to enable context-contingent ESM. One of the most prominent ESM tools so far is MyExperience [2], which is developed for Windows Mobile devices and requires installation of a Microsoft SQL database. On both tools the logging information and the experience sampling data are deployed *locally* on the mobile device itself. Most of existing tools have been implemented as standalone tools on Windows CE or Palm OS platforms.

Our approach follows a different paradigm. We propose a shift from the client (the mobile device) to the server. We therefore only need a lightweight client software and no database installed on the mobile devices. This allows developing relatively easily lightweight clients on a multitude of mobile platforms in various contexts. A similar solution is Momento proposed by Carter et al [1]. Contrary to our approach Momento does not support ESM questions triggered by an analysis of previous user behavior. Reimers and Steward [9] used SMS text messaging as a simple means for responding in ESM studies. This approach has the advantage that no client at all is needed on the other hand neither user behavior or context triggered ESM is possible.

## **USER BEHAVIOR DRIVEN AND CONTEXT TRIGGERED EXPERIENCE SAMPLING**

### **The MAESTRO Concept**

The above-mentioned approaches have successfully shown that it is possible to trigger ESM questions on certain contextual parameters. The contextual trigger has two purposes: (a) it determines the point in time, when an ESM question is asked and (b) it chooses an adequate context dependant ESM question. (a) gives researchers the possibility to react to a user interaction immediately and therefore minimizes the recall bias. (b) provides researchers with the possibility to ask context and usage behavior specific questions.

The MAESTRO approach includes both and adds another dimension. We want to be able not only to react to singular specific context information or an actual usage behavior but to respond dynamically to different user behavior patterns. To learn these patterns it is necessary to log the usage behavior over a period of time, and then dynamically assign appropriate ESM questions contingent on a long-term user behavior. It is therefore possible to provide different questions to different user groups, without knowing - a priori - which participant belongs to which group (e.g. frequent SMS users vs. infrequent SMS users). It is possible to dynamically adjust to a change in usage behavior (e.g. the participant used to write on average one Email per day and now he writes on average 25 Emails per day). It is also

possible to predict the probability of a certain user behavior and then react with corresponding ESM questions. Another advantage of this approach is the possibility to conduct a multi-phase study, where the results of the first part of the study can impact the setup of the subsequent parts of the study (even dynamically). Using traditional tools would require collecting all mobile devices used for the ESM study in order to update the study setup.

### **The MAESTRO Architecture**

The MAESTRO concept follows a client-server architecture to master user behavior driven and context triggered experience sampling. Logged events are not stored in a local database but are immediately sent to a web server, which is one of the key differences to other ESM tools (e.g. MyExperience [2]). Each of these events can serve as a trigger for the user behavior driven and context triggered ESM. Additionally, the researcher can define rules whether the specific event triggers an ESM question and - based on the user history - determine which questions are asked.

Generally the MAESTRO concept can be applied to various contexts. In the following we describe an example from the mobile context: MAESTRO has logged a certain user behavior over a period of time. E.g. the user has written on average three short messages (SMS) per week and is therefore categorized as an infrequent SMS user. One day the user behaves different than before and sends 35 short messages on one day. Whenever the user sends a short message a "SMS\_sent"-event is sent to the server. Based on predefined rules (e.g. change of SMS usage behavior) it is decided whether an ESM question is sent to the user. The appropriate question is then displayed on the client accompanied with an acoustic and tactile signal. The user is prompted to answer the question immediately. This has the advantage of minimizing recall bias. Since most events on the mobile device are user triggered, questions are likely to be asked in moments when the user is interacting with the device. This minimizes the burden for the user and can increase the answer rate. If the user answers the question, this answer is immediately sent to the server. Depending on the answer an additional question may be asked.

### **BLACKBERRY IMPLEMENTATION OF MAESTRO**

For the purpose of validating the feasibility of our approach we implemented MAESTRO on BlackBerry mobile devices. The implementation consists of a light-weight client software and a web application including a database on a web server (see Figure 1). The client was built in Java 2 Micro Edition especially adapted for BlackBerry devices. It is designed as a system module, which starts automatically with the BlackBerry OS. The client's task is twofold: firstly it sends predefined events via GPRS / EDGE to the web server. This is done in the background and happens without explicit user interaction. If at the time of the event no connection is available (e.g. too weak signal strength), the data is stored internally and will be sent as soon as possible together with the next event. The second

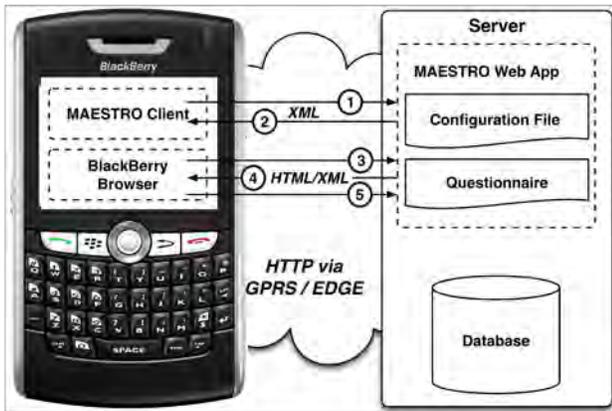


Figure 1. The MAESTRO Architecture.

task is to direct the BlackBerry's internal web browser to a web site to display the questions defined by the researcher, which allows the user to answer them.

The MAESTRO web application is implemented with PHP and MySQL. To setup a study, the researcher has to define questions, answers and structure of the questionnaire, and a configuration file, which holds the triggers for questions. When the server receives an event from the mobile client it stores it automatically in the database (1). The server determines whether the event triggers a question and which question. The particular question ID is sent back to the client (2), where the BlackBerry internal browser is opened and directed to request a URL containing the question ID from the web application (3) and then displays the question to the user (4). The user's answer is send back to the server and stored in the database (5). Depending on the answer the next question is sent to the user.

### CASE STUDY

We now present a case study to illustrate how MAESTRO can be used to measure the usage behavior and the corresponding experience of users, while interacting with their BlackBerry devices. The study included 20 participants (10 m, 10 f, age between 22 and 56 years). The study lasted for eight weeks. During this time period each participant was provided with a BlackBerry device. On each device the client software was installed. Using a BlackBerry Enterprise Server (BES) it is possible to install the software without having the participant physically come to the lab. This reduces the effort for the researcher and the participant alike. Setting up the client involves selecting events, which the researchers want to be logged during the study. This is achieved by editing a configuration file on the client, which can also be remotely updated during the study over the air. We balanced the number of ESM questions to avoid to be too burdensome for the participants. We decided to react to events three times a day – in the morning, in the afternoon, and in the evening. Once the user answered the first question, he was presented with five to

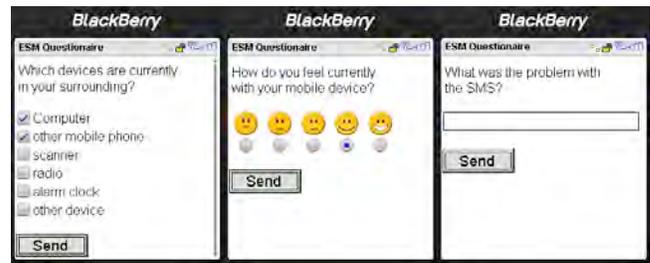


Figure 2. ESM questions as displayed on the BlackBerry.

six follow-up questions. Due to the fact that we only asked questions as a reaction to an event on the mobile device, we were able to present mobile device usage context specific questions. During the study we analyzed the logged data and the collected answers. This gave us the possibility to react to problems and adapt questions depending on the usage behavior.

From a user's perspective a typical experience sampling, where the context trigger was activated by sending a text message looked like the following: After the user sent the text message at 10:30 a.m. the MAESTRO client logged contextual information (e.g. time, user id) and sent it to the server. Based on different rules (e.g. user history) it was decided whether a sample was triggered or not. Initially all users had the same set of rules. The MAESTRO toolkit gave us the possibility to react to certain user behavior patterns in the past and ask different questions based on this behavior. This possibility clearly sets MAESTRO apart from any other existing tool. One rule was for example that only a single set of questions were asked between 9 and 12 a.m. Another rule concerned the last time a text message related question was already asked. Another rule was whether the user behavior has changed significantly or not. If all the rules applied the set of questions were displayed on the BlackBerry device. Sample questions are illustrated in Figure 2.

The results of this study are out of the scope of this paper and can be found in [8]. To give you an impression about the power of MAESTRO we report that we logged all in all 127.255 events (e.g. email sent, contact updated, keyboard unlocked etc.). The number of actual fully answered set of questions was 2.041. Altogether we had answers for 10.216 single questions, which means that we got in average nine answers per day per participant. This huge amount of data was analyzed by using prepared SQL-scripts.

### DISCUSSION

Our results show that MAESTRO enables user behavior driven and context triggered ESM studies. We showed that it might not only log usage data but also measure subjective user experience during actual everyday mobile device usage. Both the possibility to react to a certain user behavior over time and the context triggering proved to be a very fruitful approach. Our user tailored ESM approach also reduced the burden on the user to answer too many

questions since user behavior driven and context triggered ESM minimized the number of questions being asked. Individual interviews with the participants conducted after the study showed that the study with this setup as no burden. Some participants even said, that it would not have bothered them, if the study had continued. Only one participant stated the he was happy that the study ended. Based on our experiences in this case study, we now discuss advantages and limitations of MAESTRO in relation to other ESM approaches:

*User behavior driven ESM:* MAESTRO gave us the possibility to exploit long term usage behavior for shaping personalized ESM questions to different types of users.

*Remote Installation:* With the use of BES it is possible to install the MAESTRO client remotely onto BlackBerry mobile devices over the air.

*Remote Setup Configuration:* Once the client has been installed the experience sampler setup and the context logger setup can be configured via the server. This makes it easy to include a large number of participants and to change the study setup for follow-up studies.

*Real Time Data Management:* Since MAESTRO stores study data into a database on the server as it is generated the researcher has immediate access to the data. Thus the data can be backed up and analyzed in real time. This enables the researcher to assess whether the mobile device is still in use and if the MAESTRO software is running. Additionally the users' compliance with the study can be verified.

*Limitations:* The most severe limitation of MAESTRO is the need for a constant wireless network connection to be able to send logged data to the server and trigger questions. If no connection is available, the events are logged locally, but no questions can be triggered. The big amount of data traffic potentially leads to high connection fee, and may be particularly problematic when the user travels abroad. The constant network connection also reduces the battery life time cycle.

## CONCLUSION AND FUTURE WORK

We have presented MAESTRO a user behavior driven and context triggered experience-sampling approach along with an implementation on BlackBerry devices. Overall our approach gives the researchers the possibility to dynamically "orchestrate" experience-sampling studies for evaluating usage behavior. At the moment MAESTRO focuses on the specific mobile device usage aspect of the entire context. MAESTRO lays the foundation of logging contextual information and measuring user experience in various different contexts. For the future we plan to get a holistic view of the situation and therefore be able to trigger questions even more precisely. For example within the car context it will be possible that the user is asked an ESM question dependant on different car relevant data (e.g. while

waiting in front of a red traffic light). Another context in which we will deploy our concept is the factory. Dependent on operator behavior we aim at understanding reasons for certain operator decisions. Furthermore we will port MAESTRO to different platforms and release an open source version.

## ACKNOWLEDGEMENTS

The financial support by the Federal Ministry of Economy, Family and Youth and the National Foundation for Research, Technology and Development is gratefully acknowledged (Christian Doppler Laboratory for "Contextual Interfaces").

## REFERENCES

1. Carter, S., Mankoff, J., and Heer, J. Momento: support for situated ubicomp experimentation. In *Proc. CHI 2007*, ACM Press (2007), 125-134.
2. Froehlich, J., Chen, M. Y., Consolvo, S., Harrison, B., and Landay, J. A. MyExperience: a system for in situ tracing and capturing of user feedback on mobile phones. In *Proc. MobiSys 07*, ACM Press (2007), 57-70.
3. Hutchins, E. *Cognition in the wild*. Cambridge, MA: MIT Press (1995).
4. Intille, S. S., Rondoni, J., Kukla, C., Ancona, I., and Bao, L. A context-aware experience sampling tool. In *CHI '03 Ext. Abstracts*, ACM Press (2003), 972-973.
5. Larson, R., Csikszentmihalyi, M. The experience sampling method. *New Directions for Methodology of Social and Behavioral Science*, 15 (1983), 41-56.
6. Le, B., Choi, H. N., Beal, D. J. Pocket-sized psychology studies: Exploring daily diary software for Palm Pilots. *Behavior Research Methods*, 38 (2006), 325-332.
7. Mehl, M. R., Pennebaker, J. W., Crow, M. D., Dabbs, J., Price, J. H. The Electronically Activated Recorder (EAR): A device for sampling naturalistic daily activities and conversations. *Behavior Research Methods*, 33 (2001), 517-523.
8. Meschtscherjakov, A., Obrist, M., Mirlacher, T., Lankes, M., and Tscheligi, M. Evaluating the emotional attachment to mobile devices & services. Tech. rep., HCI & Usability Unit, ICT&S Center, University of Salzburg, 2008.
9. Reimers, S., Stewart, N. Using SMS text messaging for teaching and data collection in the behavioral sciences, vol.41 (2009), 675-681.
10. ter Hofte, G. H. Xensible interruptions from your mobile phone. In *Proc. MobileHCI '07*, vol. 309. ACM Press (2007), 178-181.
11. Vermeeren, A., and Kort, J. Developing a testbed for automated user experience measurement of context aware mobile applications. In: *Proceedings of the COST294-MAUSE Workshop NordiCHI '06*, (2006).