Universidade do Minho

Escola de Engenharia

André Feliciano Quintas da Silva Coelho

A web-based tool for user experience
studies

Outubro de 2012

Universidade do Minho

Escola de Engenharia

Departamento de Sistemas de Informação

André Feliciano Quintas da Silva Coelho

A web-based tool for user experience studies

Dissertação de Mestrado
Mestrado em Informática

Trabalho realizado sob orientação de

Professor Rui José

Outubro de 2012

Anexo 3

DECLARAÇÃO

Nome

ANDRÉ FELICIANO QUINTAS DA SILVA COELHO

Endereço electrónico: ANDRESILVACOELHO@GMAIL.COM   Telefone: 253282005   / 914746357

Número do Bilhete de Identidade: 12830608

Título dissertação ☒/tese ☐

A WEB-BASED TOOL FOR USER EXPERIENCE STUDIES

Orientador(es):

PROFESSOR RUI JOSÉ

Ano de conclusão: 2012

Designação do Mestrado ou do Ramo de Conhecimento do Doutoramento:

MESTRADO EM INFORMÁTICA

Nos exemplares das teses de doutoramento ou de mestrado ou de outros trabalhos entregues para prestação de provas públicas nas universidades ou outros estabelecimentos de ensino, e dos quais é obrigatoriamente enviado um exemplar para depósito legal na Biblioteca Nacional e, pelo menos outro para a biblioteca da universidade respectiva, deve constar uma das seguintes declarações:

1. É AUTORIZADA A REPRODUÇÃO INTEGRAL DESTA TESE/TRABALHO APENAS PARA EFEITOS DE INVESTIGAÇÃO, MEDIANTE DECLARAÇÃO ESCRITA DO INTERESSADO, QUE A TAL SE COMPROMETE;

2. É AUTORIZADA A REPRODUÇÃO PARCIAL DESTA TESE/TRABALHO (indicar, caso tal seja necessário, n° máximo de páginas, ilustrações, gráficos, etc.), APENAS PARA EFEITOS DE INVESTIGAÇÃO, , MEDIANTE DECLARAÇÃO ESCRITA DO INTERESSADO, QUE A TAL SE COMPROMETE;

3. DE ACORDO COM A LEGISLAÇÃO EM VIGOR, NÃO É PERMITIDA A REPRODUÇÃO DE QUALQUER PARTE DESTA TESE/TRABALHO

Universidade do Minho, 31 / 10 / 2012

Assinatura: ANDRÉ FELICIANO QUINTAS DA SILVA COELHO

Acknowledgments

Abstract

In a ubicomp environment it is important to evaluate the user experience in interaction with technology and how this technology affects users in several dimensions. To assess those dimensions, what affects them and what causes them to impact users in one way or another, researchers conduct user studies and these aid researchers to answer their questions.

While there are several user study techniques, these techniques are mostly not attuned to the specificities of ubicomp settings. This is where experience sampling proves to be a valuable technique, as it's impromptu and in-situ nature caters to the specificities of ubicomp scenarios.

This work focuses on studying user study techniques, with particular attention to experience sampling and, on the design and evaluation of the platform that we created to empower researchers with a flexible and general purpose user study tool. This tool allows them to easily create and manage studies and, to collect rich contextual data. From the study participants' perspective, the tool strives to be minimally intrusive and to not impose unwanted costs (energetic or monetary).

Evaluation results show that we were successful in achieving our goals in terms of internal validity but the sample size of the test study limits us in achieving generalizability (external validity) for our results.

Resumo

Num ambiente de computação ubíqua é importante avaliar a experiência dos utilizadores na sua interacção com a tecnologia e também, a forma como esta tecnologia afecta os utilizadores em várias dimensões. Para avaliar essas dimensões, o que as afecta e o que faz com que elas afectem os utilizadores, de uma forma ou de outra, os investigadores realizam estudos de experiência e estes, auxiliam no esclarecimento das suas questões de investigação.

Embora haja várias técnicas para a realização de estudos de experiência, maioritariamente dá-se o caso de a técnica escolhida não estar calibrada para lidar com as especificidades de um cenário tecnológico como o da computação ubíqua. Este é um aspecto em que o *experience sampling* se revela como uma técnica valiosa, pois a sua natureza *impromptu* e *in-situ*, faz com que esta se adeqúe a estas especificidades.

Este trabalho foca-se no estudo de técnicas de avaliação para estudos de experiência, com particular atenção para a técnica de *experience sampling* e, no desenho e avaliação da plataforma que criamos, com o intuito de empossar investigadores com uma ferramenta de estudos flexível e de âmbito geral. Esta ferramenta permite de forma fácil, a criação e gestão dos estudos e, permite a recolha de dados contextualmente ricos. Da perspectiva dos participantes em estudos a ferramenta visa ser minimamente intrusiva e não impor custos indesejados (energéticos ou monetários).

Os resultados da avaliação mostram que atingimos os nossos objectivos em termos de validade interna mas, o tamanho da amostra da população do estudo de teste, limita-nos na possibilidade de obter resultados generalizáveis (validação externa).

Index

When designing a data model for the studies conducted using this type of tool, many considerations have to be factored in, like for instance:

Index of figures

Index of tables

Abbreviations

A-ESM Adaptive Experience Sampling Method

A-GPS Assisted Global Positioning System

API   Application Programming Interface

APK   Application Package

ASP   Active Server Pages

BSSID Basic Service Set Identifier

CAES   Context-Aware Experience Sampling

CAPTCHA   Completely Automated Public Turing test to tell Computers and Humans Apart

EPSS   Electronic Performance Support System

ESM   Experience Sampling Method

FNAC   Fédération Nationale d'Achats des Cadres

GPS   Global Positioning System

GSM   Global System for Mobile Communications

IIS   Internet Information Services

LTSI   Licenciatura em Tecnologias e Sistemas de Informação

MVC   Model View Controller

OS   Operating System

rESM   Refined Experience Sampling Method

REST/RESTful   Representational State Transfer

SQLite Structured Query Language Lite

WCF   Windows Communication Foundation

1 Introduction

1.1 Motivation

The experience sampling methods (ESM) [1], [2] are one of the techniques used to capture information, pertaining to real or potential technology usage in peoples' everyday life. This method consists in the use of tools that alert the user to register information, regarding an on-going study, in which he or she is taking part. That registry can consist of: a questionnaire that the user needs to fill-out, a picture or, for example, text stating the user's activity or feelings at that particular moment.

There are two main advantages in using this type of methods. The first is: the ability to collect information about what is happening at that precise instant. This constitutes an important advantage and, one that distinguishes this method from other methods, namely interviews, because these try to obtain information about events that have already passed or, situations in which the person no longer finds herself involved in. As a consequence of this fact, much of the relevant information from the experience will have been forgotten or will not be remembered in the scope of the interview. On the contrary, experience sampling methods are about the here and now which enables us to access information whilst the user is immersed in the situation itself. The second advantage is that the study can unfold without the continuous presence of the investigators; this in turn supplies an additional advantage by enabling more people to be involved in the study and for a longer period of time, than what would have been possible, for instance, by using observation techniques. Experience sampling methods allow the person undergoing a study to maintain a normal daily practice, without the pressure of feeling like they are under study or scrutiny. The only difference being that users will occasionally receive an alert soliciting them to perform the task therein described.

Experience sampling techniques are of particular interest in the scope of Ubiquitous Computing, as the development of systems in this field frequently requires a study of the circumstances under which a certain technology or system is used in the daily lives of the target user base. Many opportunities and limitations of Ubiquitous Computing technologies may only surface in the micro-contexts of everyday life and, as such, will never be detected by studies outside of these contexts. These *in situ* or ecological validity requirements are the reason why ESM, a technique from the field of psychology, was adopted by ubicomp researchers. This

technique was found to be very effective for learning about situations and person-situation interactions which are common-field in ubicomp settings. Additionally, mobile phones [3], [4] provide an ever more appealing platform to sustain this kind of tools, as they facilitate user alert, questionnaire delivery and data capture, when conducting this type of study. This is especially so when one considers platforms like Android and iPhone with substantial data capture and connectivity capabilities, allowing for mobile experience sampling. These platforms represent very enticing and interesting opportunities when regarded as a basis for tools supporting ESM.

Considering how important user studies are in the field of ubiquitous computing it is not hard to argue the case for the existence of a general purpose user study tool. There are clear advantages in arming researchers with a flexible multi-study tool as this will facilitate and accelerate their research work. Not having to develop custom-built tools or learn new ones every time they wish to conduct a study significantly cuts down on the time elapsed from instantiating a technology to testing it. This makes it faster to obtain results, and could even make them more valuable through the addition of contextual data that mobile phones allow us to capture, or even by enabling researchers to collect preliminary data and redirect the focus of research while a study is on-going.

This investigation project encompasses the creation of a Web tool for supporting studies, susceptible to use in several types of studies that are based on ESM, as well as an accompanying mobile application for Android powered devices.

1.2   Objectives

The research and work carried out throughout this investigation project are bound by the following goals:
- Establishing a solid understanding of the evaluation processes based on experience sampling and uncovering the key abstractions that may be used for designing generic tools;
- Devising and implementing a system architecture for the creation of a web-based tool that supports general user experience sampling studies;
- Evaluate the tool in regards to its ability to address the needs of researchers wishing to run user studies based on experience sampling;

2

- Evaluate the platform for mobile data collection in regard to its ability to collect samples and integrate into everyday practices and devices.

With these goals in mind, we are trying to substantiate a platform which we hope will suit various needs and contribute to the help the ubicomp researcher community. The main contributions of this work will be the tool itself and the model which will hopefully allow it to be used in various types of studies.

1.3 Overview of the dissertation

This work covers the study of several common user study techniques, with particular attention to experience sampling. It analyses other efforts into creating experience sampling tools and it also analyses other evaluation tools for user studies. Following this, we analyse work on mobile sensing, which is an area that is a key enabler of the type of study support we are offering. We discuss important aspects of our analysis work and how it informed our platform's design. After this, we propose our model for experience sampling support and next we propose our vision for the platform's architecture and our prototype implementation of this vision, which features a web server and a mobile client application to support the sampling process. Finally, our deployment study for our evaluation process is presented, our results discussed and we conclude upon our findings.

2   Related work

In this chapter, we analyse work relevant in informing our design decisions regarding the creation of a general purpose study tool that would be based upon experience sampling. Naturally, we start by analysing several user study techniques, their advantages and disadvantages. Following this, we analyse the experience sampling method (ESM) in its many forms, in order to uncover issues surrounding its use as a study technique and to better grasp the essence of ESM. This was an important step in understanding how it wards off the most common downsides to the other techniques under analysis. To better understand the requirements and implications of creating and managing a flexible user experience study tool based on ESM, we studied work that discusses several prior efforts in this research field that resulted in mobile experience sampling tools. We reviewed this work with the intent of laying our foundations on the basis of previous knowledge and experiments. This effort was key in avoiding past mistakes and allowed us to build upon knowledgeable advice on how to approach common issues. Finally, we outlay and assess what are the main issues to be taken under consideration when undertaking the development of such a system for all parties involved in its ecosystem.

At this point, it is pertinent to clarify some terminology necessary to understand some of the points discussed in this section. Having stated that, the following list of terms and their explanation should suffice to provide clarity to the reader:

- The **internal validity** [5], [6] of a study technique is a property that reflects to which extent one can accurately state that the independent variable(s) assessed, produced the observed effect. This means that it is a measure of cause and effect;
- The **external validity** [5], [7] of a study technique pertains to its ability of generating findings that can be generalised to an audience larger than that of the study sample, or that can be generalised across tasks, or even across environments;
- **Ecological validity** [8] is a property of a user study technique by which its results can be deemed to be applicable in the real-world scenario that the test study aims to cover. This implies that the test study's methods, materials and settings would ideally approximate, as much as possible, the real-life situation under study.

2.1   Overview of user study techniques

Here we present a comprehensive range of common user study techniques that we surveyed. We describe what defines them as well as their main advantages and disadvantages. This allows

us to gather knowledge on important issues like validity, and how it is affected by these techniques and by the constraints they impose on the study subjects. The analysis of these aspects is covered in section 2.5.1.

Survey research [9] is a technique that entails aggregating information from a sample of the population, using standardized instruments or protocols, with the purpose of answering a research question. This technique involves administering a questionnaire, conducting field interviews and using diary studies. The main advantage of survey research is that it supports the generalization of results because it enables the gathering of large amounts of data, with relatively little effort.

Diary studies [10] focus their attention on studying events that occur during interaction and while this may be a valid method for extracting knowledge about a user's perspective and context of interaction, it is clearly influenced by the user's self-report bias. Also, this method is constrained by the consistent demand it poses on the participants, requiring them to interrupt their main activity to register things on the diary. One important consideration in diary studies is the media chosen to record events, as this has been shown to effect penalties on the subject's recognition and elicitation.

Observational studies [11] are characterised by a technique called shadowing in which a researcher will follow the subject, and register relevant aspects of the interaction context. This type of studies has multiple advantages: participants are not required to input additional effort in registering data and therefore, data produced is more objective; data registration can be coded and produce quantitative data; this technique can generate large quantities of data in fairly short periods of time. However, this technique offers no assurance that the data collected will be minimally representative. In addition, it is often difficult, intrusive and improper to move around with devices to record the subject's interaction and the contextual elements. Sometimes, pair observation [12] is applied to try and combat this. Pair observation consists of having a friend of the subject act as the researcher. This variation of the technique is susceptible to data capture bias, if the friend happens to be improperly trained.

Cultural probes [13], unlike the techniques referenced up until this point, aim at taking into consideration the subjective interpretation of the research material, on the part of both the researcher and the participant. This study technique consists of giving a number of participants, a set of instructions and a recording device (e.g., diary, Dictaphone, disposable camera, etc.). The instructions are then interpreted by the participants who create self-reports and then, return

the material to the researchers. The strong suit of this approach is that it enables researchers to reach environments that are difficult to contemplate directly, and therefore it allows for capturing aspects of human cognition to which researchers would otherwise remain unaware of. Cultural probes also reveal themselves to be less intrusive than observational studies.

Longitudinal studies [14] are often used as a technique for collecting data, either qualitative or quantitative, at repeated intervals in time. Afterwards, comparisons are drawn on the different time periods and this exposes the dynamic nature of a variable. This technique is limited by some aspects, such as its inability to resolve issues of causal order and, the difficulty in guaranteeing that no intervening factor biased the data collected between periods.

Automatic logging [15] is a study method that consists of gathering data generated or sensed automatically by a user's mobile device, without their notice. It is easy to extrapolate that this is a reliable way of gathering consistent and objective data. Alas, this technique cannot capture the user's context or cognitive states, which poses as a limitation to the quality of data.

Controlled experiments [16] are a method similar to the aforementioned one, in that variables considered to have an impact on the research questions under scrutiny, are reliably logged or manipulated. These experiments are usually conducted in controlled environments where logging and control of disturbing factors is easier. This high degree of control constitutes this method's greatest advantage and herein lays its most criticised flaw. Lab studies are usually criticized for their limited relation to the real world and undetermined level of generalizability.

Quasi-experimentation [17] is a method proposed to complement the main shortcoming of controlled experiments. As such, this technique tries to bring controlled experiments into the actual environments where subjects use mobile technology. It requires that subjects wear a number of micro-cameras and perform structured tasks, designed in a way that controls spurious factors. The result is a technique with better internal validity but reduced external validity because of the constraints imposed on the behaviour of the subject and, the necessary recording equipment.

A case study is "an empirical enquiry that investigates a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident" [18]. Data collected through the use of this technique is grounded in natural settings. It is typically very rich and sometimes gathered without the support of a structured framework, resulting in incomplete or inconsistent datasets. This type of study is particularly

suited to the task of describing and explaining specific phenomena, developing hypotheses, or theories.

Thus far, all presented techniques follow the principle of logical deduction. As of late, research on mobile technology is progressively adopting the principle of logical induction. As the number of ubiquitous technologies capable of sensing human behaviour rose, so too did the amount of raw data. Datamining [19] techniques are applied to these datasets in the interest of learning of recurring patterns and predicting human behaviour or future status of systems used by humans. The main flaw of this technique is its lack of internal validity, which is offset by its large external validity owed to the vast extent of collected data.

Simulations [20] are processes by which we try to obtain the qualities of (real or planned) reality without reality itself. This is a very versatile technique since it enables one to deal with realities that might exist within a range of possible realities depending on the variation of the values assigned, or taken, by configuration parameters. The principal shortcoming of such a technique is that, results can be invalid or misleading if the underlying model of the reality under simulation is a poor one and does not account for the need of simulating essential aspects of said reality.

Morten S. Andreasen et al. [21] analyse the three methods for remote usability testing. One method involves remote synchronous condition, where they conduct the test in real-time but the test monitor is separated in space from the test subjects and the other two methods (expert testing and user testing) are based on remote asynchronous condition, where test monitor and subjects are separated both in time and space. Remote usability testing is a method that has been studied empirically and which some software companies apply when testing their products. These authors present their empirical comparison of the three methods in contrast to a conventional laboratory methodology and their results point to two conclusions. One is that the remote synchronous method is nearly equivalent to the conventional laboratory method, it even affords for a higher task completion rate and, therefore, it is a potentially good approach for involving broader audiences in usability testing and support novel development approaches. The other is that asynchronous methods, when compared to the conventional one, uncover less usability problems and are more time consuming for test participants. This methodology of remote testing can be useful in crossing organizational and geographical boundaries, however, the boundaries we wish to cross are of a different nature, given that the requirements of our scenarios are entirely different. We have to take into consideration participant mobility

requirements, participant convenience, participant adoption and a whole host of other things that do not apply in the case of this methodology.

2.2 Experience sampling technique

In this section we present and discuss the experience sampling technique and several refinements that extend it, in order to improve the validity of its results. We approach the issues and the benefits of this method and its variants. Finally, we report on previous experience sampling tools, dissecting their approaches while exposing the advantages and shortcomings of said efforts.

2.2.1 Experience sampling methods

Vastenburg, M.H. and Herrera, N.R. [22] state that: "researchers cannot know all questions nor the detailed product use and context of use in advance", therefore, they propose that ESM needs to be adaptive. "Adaptive ESM (A-ESM) aims to enhance the existing experience sampling methods by providing instruments that enable researchers to easily inspect the preliminary findings in relation to context and product usage data and to easily adapt the questions and triggers". This adaptability aims at enhancing the ESM technique so that it can provide data to researchers faster, with better quality, and also allow them to change the study according to their current needs and preliminary findings.

ESM, like many others, is an example of a technique from another field (psychology) that was brought into the field of ubiquitous computing, and that usually results in the technique not being fully adapted to the purpose it was brought in to serve. As a consequence, even though ESM preserves the ecological validity of its measurements, it produces self-reported data which is susceptible to cognitive bias. This led Cherubini, M. and Oliver, N. [23] to propose a refined ESM (rESM). This technique consists of two things: 1) automating mobile device data collection through use of sensory capabilities and 2) triggering data collection on objective user-generated events that are non-self-reported. This is an important consideration to inform the design of an ESM application as it tells us that any and all data that is relevant to a study, and whose collection can be performed automatically, should be collected in just that fashion. If the user's intervention is not required and could possibly taint data due to cognitive bias then the user should be removed from the equation.

Another case of a technique designed to adapt ESM to ubiquitous computing settings is the paratype [24]. This is an inquiry technique based on experience prototyping and event-contingent

experience sampling, designed to survey users in real-life situations, about ubiquitous computing technologies. It is a simulation of a technology or interaction model that is evaluated alongside a real-world experience. Paratypes are very useful in evaluating high-level or implicit interaction where reference to a concrete experience in real life is necessary. Iachello, G. et al. claim that paratypes are particularly useful when it comes to gathering early feedback on mobile and ubiquitous technologies, especially applications that collect information at unexpected times, that provide information when needed, or where interaction is embedded in unplanned social action or everyday routine. This methods usefulness is due to the fact that it avoids high costs in prototyping and, it allows the gathering of relevant information needed for design and/or that is only possible to obtain in social domains that are sometimes not addressed or, do not occur in other study techniques. However, one must keep in mind that paratypes are susceptible to sources of bias like proxy/researcher demographics which can influence respondent/participant demographics and the selection of probed situations. Use of this technique requires overcoming certain obstacles related to the disruption in the flow of social interaction, sampling accuracy, and procedural requirements.

2.2.2 Experience sampling tools

Fischer, J.E. [25] reviews three popular ESM tools, ESP/iESP [2], CAES and MyExperience. The article aims to expose the shortcomings of each tool and make recommendations on what would be a better approach to a general purpose ESM tool. ESP/iESP is a tool composed of a Windows or Linux, browser-based, desktop application and a mobile Palm Pilot client. The desktop application is used as a means for creating and managing the logic of the study, its content, and the timing at which the questionnaires are triggered. This tool's most prominent deterring factor is its mobile client which is tied to Palm Pilot, an out-dated device, with no connectivity and location or acceleration sensors. CAES [26] (Context-Aware Experience Sampling) was a tool developed in an MIT project that was designed to enable questionnaires to be triggered by contextual clues, derived from the computation of data captured by the device. The tool ran in Microsoft PocketPC PDAs and was developed in a modular fashion, so as to allow it to capture data from context-sensing sensors. This would allow researchers to use context-sensitive sampling to trigger specific questions in relation to specific user actions. The proponents of this tool aimed to improve ESM by introducing contextual-awareness and thus, minimize the interruption annoyance of the ESM technique by sampling only during, or just after, an activity of interest. This project joined efforts with the team from MyExperience. This last tool

is an open-source one that runs on devices with Windows Mobile 5.0. The MyExperience client can collect sensor data, allow for time-based (random or scheduled) or sensor data-based questionnaire triggering and also, provide phone usage data through logs. Researchers making use of this tool require some degree of technological knowledge as the creation of studies is done by editing XML files. However, that is not the only limitation as MyExperience's default approach translates into a black box for the researcher, this is because all the data is stored on the devices themselves. This means that all the data must be aggregated at a later stage and also, that there is no form of telling whether or not the users are engaged in the experiment, while it is on-going. This can be overcome by configuring a server to post data to, the inconvenient aspect being that it is not an out-of-the-box experience.

In MobXamp [3] and Track your happiness [4] we find mobile experience sampling tools that take advantage of a popular mobile OS (iPhone), the former is designed to evaluate contextual economic decision making and the latter is designed to track a user's happiness and, what factors in to it. As such, both of these tools were custom-built for a specific case study whereas we intend to build a more general purpose tool.

2.3   Other evaluation tools for user studies

User studies are generally rigorous undertakings that end up being a trade-off between sample size, time frame and monetary costs. As Aniket Kittur, Ed H. Chi and Bongwon Suh [27] detail, micro-task markets such as Amazon's Mechanical Turk are an example of a tool that can be used to quickly and easily engage a large number of users for diminished time and monetary costs. These micro-task markets can be particularly useful for surveys, rapid prototyping and any quantitative performance measurement.   However, very careful task design is necessary so as to guarantee that the results are usable. As this investigative work shows, two different approaches to task design lead to significantly different results and the effect of this is particularly noticeable if the desired measurements are of qualitative nature.

In [28] Fogg et al. detail their web-based system for conducting online user studies. The authors used iterative design to develop a system that introduces participants to studies, ensures their informed consent and presents them with the experiment, while taking measures and storing them in a database. This system allows researchers to quickly run studies and access a wide-range of participants (around the world). This is not dissimilar to what we aim to achieve but our focus is geared towards ubicomp settings, which have different requirements and context. Their results show that this method offers advantages in deployment and study flow speed and,

study sample reach, which allows for quicker findings. However, this type of approach was designed to cater to a specific segment of user studies, mostly pertaining to usability in HCI. We intend to serve a much broader spectrum of study types.

SOUSA [29] is a general purpose user study tool for sketch collection and verification. Sketch recognition is a field of study that works towards integrating into computer systems the ability for them to understand hand-written symbols and diagrams, in order to facilitate user interaction. This is a field the claim to be riddled with multiple tools for the same effect but instantiated to the particular needs of one study, so the authors decided to build and make available (online), one tool that could serve anyone trying to perform this type of study. This tool would also make available the datasets gathered from previous studies to anyone in need of that data. The results from their verification studies point towards the system being a good tool for collecting sketch data easily and efficiently but it appears that it needs stronger security since the tool has no control over who creates studies and manages studies. While this tool aims for being general purpose, and like others reviewed benefits from the possibility of having extended reach through online presence, it was built with a specific niche of studies in mind. Ours aims to provide more flexibility in choosing what the study will be trying to evaluate, there are however lessons to be learned here, regarding the security of studies and also, regarding the possibility of adding post-study questionnaires for gathering additional data that can be valuable.

KidTools [30] is an electronic performance support system (EPSS) that allows teachers and children to customize easy to use templates, whose purpose is to have children to interact with the system in order to experience cognitive-behavioural interventions that will allow them to gain internal control over problematic behaviours in the classroom environment. The tools are organized into several types of self-management procedures and feature colourful graphics, text-with-audio directions and generally simple formats. They can be used independently by kids, or teachers, or kids can be assisted by their teachers in the process. The authors' evaluation concludes that the tool has real benefits in helping in children to overcome the issues already discussed and, additionally, it offers good design guidelines, such as gathering feedback from the target audience to help guide tool design and development, this is a critical notion as they highlight, because children think differently from adults. Another important consideration is how important it is for the tools to be easy to use overall, teachers will not have much time to read manuals and troubleshoot software and children will have limited technical capabilities. This is true of any user study tool, people will not accept tools if they are cognitively taxing for no good

reason and, if they present burdensome obstacles to them in the task they have to perform. The tool should just present the task and generally facilitate user performance.

2.4   Mobile sensing

Yoshida et al. detail the creation and deployment of a localization system based on Wi-Fi fingerprinting [31]. The paper focuses on the effect and efficiency of the method used for acquisition of fingerprint data and the influence that it could have on the accuracy of the localization results produced. Even though this work has different objectives, it has informed us about the need to devise a data collection protocol that could accommodate for the variation in the network landscape.

Zandbergen [32] describes a study to estimate accuracy of positioning techniques, using a similar methodology. A mobile phone (iPhone 3G) was used to collect location data (A-GPS, Wi-Fi and Cellular positioning) at several distinct metropolitan locations, and test the accuracy of each of the iPhone's positioning methods against a benchmark location (ground truth). We aimed to test the accuracy of a delay-tolerant location calculation (with Wi-Fi and GSM data), against the location data provided by the device on-site, also in relation to ground truth.

PlaceLab [33] is a positioning system that allows users to locally (on their device) calculate a location, based on the BSSIDs and signal levels of nearby Wi-Fi access points. This location is calculated by crossing the data gathered in real-time by the device with information stored locally on a database, to which the user previously subscribed. While this may address some of the connectivity issues we have identified, it would have important disadvantages in terms of deployment, given the need to install additional client-side software and database.

Herecast [34] allows users to determine their symbolic location, e.g. building floor. Information about locations is kept in a database that is maintained and accessed by the community. While an alternative for cases in which information only needs to be generated occasionally and in familiar locations, this is not suitable for frequent location logging.

BikeNet [35] supports the collection of data related to performance, environmental and health aspects of cycling and provides an example of a sensing systems that demonstrates the effectiveness of the overall approach of relaying sensing data back to a server to address specific requirements of the sensing process.

2.5 Analysis

2.5.1 User study techniques

The single most important criticism of the qualitative set of methods under analysis (survey research, diary studies, observational studies and pair observation, cultural probes and, longitudinal studies) is that, their involvement of the user in the process of data capture makes them permeable to potential biases introduced by the subject, his personality, his expectations and personal preferences/biases when coming into the study. This does not imply that the validity of these methods is necessarily compromised by this fact but it is a factor that must be taken under consideration. There can be instances when averting this is paramount in the pursuit of valid results or even, where the combination of various methodologies can be beneficial or required for extracting the intended information. This is where the quantitative set of study methodologies such as automatic logging, controlled experiments, quasi-experimentation, case study, datamining, simulations and remote usability testing come into play as all of them seek to combat the phenomenon under exposure.

Experience sampling presents itself as a technique that appears to combat most of the issues with other study techniques. This is not meant to say that it is perfect, naturally, as any other technique it has its issues. However, ESM seems to be particularly well suited for studies in ubicomp scenarios, considering its impromptu and in situ nature, which allows it to capture not only the relevant information regarding occurrences but also, their context at the time. As such, ESM can be used as a formative technique to reveal where ubicomp solutions might be appropriate and to help define their requirements, it is an ecologically valid user study technique that can be used as a summative technique. However, Cherubini, M. and Oliver, N. [23] remit to the fact that probing peoples' mental models on interaction with technologies in a given context is hard if not impossible; ESM is a better away of achieving this but still, answers can suffer from a self-report or precision bias. In these situations rESM provides better results by combining automatically generated data with user queries prompted at a time when the answer is most relevant. This is clearly a design guideline that underlines the importance of ESM tools adhering to programmatically controlled event-based triggering rather than user controlled event-based triggering, so as to avoid these biases. One other important aspect of ESM, in which it can differ and improve upon other techniques, is in its ability to become an adaptive technique (A-ESM). This can be a significant advantage for researchers since it allows them to accompany the study's findings, gather preliminary results and, if necessary, adjust the study in order to better

answer the research questions or answer new ones that may have arisen from the preliminary findings.

2.5.2 Tool design

When designing an ESM tool there are two different perspectives whose essence needs to be analysed and captured, the researchers' and the study participants' perspective. On the researchers' side there are several issues and nuances that are critical to the adoption of a tool. As Fischer, J.E [25] suggests, all ESM tools require researchers to be familiar with their method of creating questionnaires. This translates into having to choose a tool carefully and with the certainty that it will cover all the necessary bases because, having to change tool means that there will be a learning curve to master a new tool. Depending on how steep the learning curve is, this can be time consuming and will take time away from the primary task of conducting a study. Taking this into consideration, a staple feature of design emerges, ESM tools should be designed for authoring as easy-to-use creation of studies is paramount for adoption by the researcher community. On the participants' side design will mainly be influenced by factors such as convenience (both not having to carry extra apparatus, and not having to learn something that is complex to use), interruption during critical processes and the annoyance inevitably associated with that.

2.5.3 Study sample size

Another critical point for researchers, when conducting a study is: how many people can they recruit? A bigger target audience means there is more chance to collect meaningful, representative data, reason for which opting for a platform with a native application or, forcing participants to carry an additional device reduces the recruitable user base and can increase costs. Considering this, it seems there is a seemingly simple solution to both these obstacles. Develop a web-based generic client that is accessible from any mobile phone. Making use of peoples' own devices is less expensive and it increases the likelihood that people will actually carry the device when prompted to interact. It would seem then that this is the ideal approach. However, sensory data is evermore invaluable in ubicomp user studies and it is hard to obtain, today, without native applications on mobile devices.

There are other possible approaches. A mixed approach could be considered, if one designs for different levels of study complexity then it allows us to target a larger user base by adjusting the way the tool works, to the capabilities of each device. Notwithstanding the validity of this

solution, it incurs a higher penalty for the designers of the platform as it adds a lot of complexity to the system and it is also important to point out that as Iachello, G. et al. [24] mention, study sample size is usually limited by the size of the proxy's interaction network. The compromise solution seems to be the targeting of popular mobile OS platforms with a client-server architecture. This means the design of the tool should: separate logging and questionnaires enabling questionnaires to be accessed online and phone usage or sensory data to be handled through client software and, it also means that designers should make wise client choices when picking which target language/platform to use and regarding requirements like background processes for triggering questionnaires, range of sensory apparel needed as well as other considerations that might influence choice.

2.5.4   Tool adaptability and flexibility

Yet another crucial point for researchers is tool flexibility and adaptability, to the possibly changing conditions of the study. Intille, S.S et al. [26] indicate that exploratory stages of product design require a dynamic experience sampling approach. User behaviour, changes in context use and research interests, are all motives for the need to have an adaptive ESM tool. This creates key design challenges in: detecting unexpected behaviours, detecting that anticipated behaviours do not occur, and providing the ability to view the findings of the experience sampling study in relation to the configuration in time. Enabling researchers to adapt the ESM protocol in accordance with changes in the usage context, user behaviour and research interest is a key factor in improving ESM studies.

Designers should create a modular tool that supports study orchestration. The client-server-based approach factors in to this by allowing for the possibility of a researcher to monitor the progress of the study and motivate participants to engage more, alter the content and structure of the questionnaire. There is an emerging pattern here in that all design choices affect either power or availability of the tool as a whole. There is a delicate balance between power and availability, the shift of that balance is inexorably linked to the type of data the researcher intends to collect on the client side.

2.5.5   Quality of data, user perception and, intrusion/disruption minimisation

As with any user study, the quantity and quality of data is reflected on the amount of participants and on how participants perceive the tool. As Intille, S.S et al. refer [26] a common challenge when using ESM is to maximize the quality and quantity of the samples while

minimizing interruptions and maintaining the motivation of the participants. Vastenburg, M.H. and Herrera, N.R. [22] indicate that peoples' behaviour and their response to technology are critically linked to the environment and context in which information is presented or requested. What this means for us is that, in designing a tool that intends to be used in the context of everyday life to sample user experience, we should strive to have it be as little intrusive as possible, and to minimize the interruption annoyance of the ESM technique. This can be achieved through contextual awareness as it allows for the acquisition of richer information and enables the sampling to be done at an appropriate time or even, for data to be acquired automatically.

Furthermore, the design of the questionnaires itself can be made to reflect this need to lessen intrusion. Chained response-driven sequences of questions, this implies answers presented in multiple-choice fashion, and aggregation of questions can work to minimize disruption by diminishing time spent answering. On the other hand, sometimes it might be more useful to collect a statement from the user or another form of input so, additional methods may apply too, such as: pictures, audio and/or video recording. The subject of interruption is one that should not be taken lightly, it has come up often when presenting and discussing the creation of an ESM platform with researchers and potential study participants alike. Often times it was mentioned that there should be some form of either stopping the system from issuing notifications or simply opting-out or delaying response. As Fogarty et al. [36] shows us, especially in the work environment people consider themselves to be in a highly non-interruptible state most of the time, which translates into the necessity of devising mechanisms to evaluate a person's interruptibility. Additionally, their work shows us several models based on simple sensors that can be used to gauge someone's interruptibility, some of those sensors can nowadays be found or developed using mobile phones. Whether or not interruptibility considerations come into play in the prototype deployment, any ESM system should take great care in its design to accommodate for this reality.

An important perception issue relates to the way the application deals with answered questionnaires, Consolvo S., and Walker M., [2] state that immediate transfer of responses could create a feeling that the participants are being watched. This can be a potential point for leading participants to provide less data, inaccurate data or quit the study.

All of this informs us of the importance of having an intelligent application design that can help to minimize the issues discussed.

2.5.6 Considerations on sensing

Sensing is critically important for this type of system, which relies to a large extent on the ability to perceive the physical environment and the activities taking place in it. Mobile phones, with their already substantial data capture and connectivity capabilities, have a unique potential to become powerful sensing devices for uncovering new knowledge about the realities of our world and the patterns of Human behaviour [37]. In particular, considering their widespread use and their continuous presence in people's lives, they represent a major resource for location-based data collection. For example, to study mobility patterns within cities, there is a need to collect traces of users moving across the city in their daily life. In experience sampling studies [1], [2], there is a need to register, either implicitly or as part of an explicit user action, events as they occur in people's daily lives and annotate them with location information that will normally be crucial for their interpretation.

Location-based logging is fundamentally shaped by the need to combine frequent device positioning with the consequences that the process can have for users. These processes normally involve recruiting people to run the data collection applications on their own mobile phones and as part of their normal daily activities. This is crucial for generating realistic data and enabling larger scale studies. However, if the data collection implies significant energy, communication or privacy costs for users, it will become a severe obstacle to large scale use and volunteer recruitment. Therefore, in our location logging processes we have included two important design principles: the first is avoiding the use of GPS. We would always need an alternative solution because we need to consider indoor locations, but the key issue is that the continuous use of the GPS would necessarily have a very high cost in terms of power consumption [38]; the second implication is to avoid depending on connectivity. In part this is also important to save energy, but since many people will not have a flat-rate data plan, there is also the issue that they will not accept the potential costs associated with data communications. The independence from connectivity would also allow us to perform the positioning without having to wait for the availability of a network connection.

3   Requirements for experience sampling tools

In this chapter we describe the infra-structure necessary to support an ESM system along with the requirements and design goals one should meet when approaching the problem of creating said infra-structure. Careful analysis of related work allowed us to uncover common challenges and requirements posited by the creation and management of a platform of this type.

3.1   Ecosystem

The system's infrastructure will inhabit an ecosystem where there are different key agents, participants and researchers, who interact mostly at different points of the system, for different purposes. The role of researchers dictates that interaction with the system will occur while stationary and in deliberate fashion, while participants in their majority, will interact with the system mostly when prompted to do so and possibly while on-the-go. Immediately it is clear that there is a fundamental divide amongst user types, given that they have very different mobility circumstances and requirements, just as well as other types of requirements. These constraints point us towards a well-known architectural design reference, the client-server approach.

Just as there is a divide in user types, there is also one in user needs. Researchers, for instance, will privilege flexibility, ease of managing studies and accessing study data for analysis, whereas participants will be more interested in not being too encumbered by the system, given that it will be sampling events during their everyday life, fact which potentiates all number of different environments and contexts where such a system can be a burden on the participant. Researchers will not only be interested in sampling events in everyday life but they will also be interested in attaining valuable information that complements user-generated data. This information would allow them to complement the sampled events with context, enabling them to enrich the results of their studies. Another key factor for researchers will be critical mass, in general, there is little information to be gained in conducting studies with few participants.

These requirements, in conjunction with the mobility requirements of participants, mean that smartphones are a good choice for deploying a client, not only due to their vast arsenal of sensors but also because they are personal devices we carry around with us every day and therefore are less likely to forget. This immediately knocks down a significant barrier in user adoption. Choosing smartphones to deploy the client, taking into consideration the requirements already discussed, means that it is genuinely important to target a popular mobile OS that both fits system needs and allows for a significant user base to be available to researchers.

Participants, on the other hand, will value not only the commodity of not having to carry around an extra device (which is also good in terms of platform design, as it minimizes the heterogeneity of the ecosystem) but also, other aspects, such as not being constantly interrupted, not having diminished battery life on their devices, not incurring in monetary costs with data. Evidently, ecosystem needs impact both sides of the architectural model and this is something we will discuss in detail in 3.2 and 3.3, where we will see how the issues discussed here affect design principles.

3.2   Web support

As previously stated, a client-server architecture fits perfectly into the needs of our platform from the perspective of all of the agents involved. This means that the underlying ecosystem would feature a server that would allow for study creation/management, perform data storage and user enrollment. In tandem with this, we would have a mobile data collection device running a client application, which we discuss in section 3.3. In more concise terms, this would encompass, in architectural terms, the creation of a modular server that would fulfil the following requisites:

- Adaptability in the form of allowing to changes study parameters;
- Ease of authoring, the tool should provide easy-to-use creation/management of studies;
- Exporting data in a format that would make it easy to process.

3.3   Mobile support

In keeping with the chosen architectural model, support on the mobile side would be provided by a client application. Developing a native application means having to choose a mobile OS. The choice of mobile OS for supporting an ESM system should take under consideration the needs of the ecosystem in all the dimensions previously discussed and, as such, abide by the following requirements:

- Background processes for alert generation and handling and, for having guaranteed access for performing automatic data collection even in periods when the user is not interfacing with the client;
- Background GPS and Bluetooth capabilities;
- Standby power management relating to GPS and Bluetooth functionality that ensures that these radios do not massively impact battery life;
- Data connectivity through 3G and/or Wi-Fi for relaying data back to the server.

Requirements regarding ecosystem agents also apply, as such, researchers would value that the client:

- Provides flexibility in adding new components and adaptability to contextual or, other changes, during the study;
- Separates questionnaires and logging, making questionnaires available online and handling phone usage data in client software.

Conversely, study participants would value a client that can meet the following requirements:

- Minimize intrusion and user disruption when sampling, through careful questionnaire design (chaining questions, keeping questions simple, keeping questionnaires short) and automatic data collection which also helps to combat possible self-report biases;
- Allow users to control which type of connection will be used to synchronise data;
- Optimizing resource usage to keep battery life impact to a minimum.
- Configurable time period for alerts;
- Delaying location estimation and off-loading it to the server, so as not to incur in energetic or monetary costs that would impede participation for some users.

A crucial issue when supporting experience sampling on a mobile device is related to how the client supports sensing. As we discussed in 2.5.6 there are several important aspects that can impact user adoption and experience, and that also affects the value that a tool may have for a researcher, as getting users to adopt a platform is of high importance if results are to be meaningful. To address these issues, we introduced the concept of delay-tolerant positioning (more on this in section 0). In most location-based services, location is normally part of an interactive feature and thus needs to be immediately available. On the contrary, in location-based logging, location information is needed to annotate an event. Therefore it is possible to just store the information needed to determine location, and leave the actual location calculation to some later point in time. For location-based logging applications, this means that frequent positioning records can be generated without forcing the device owner to incur in significant power or network costs.

4   A system for web-based experience sampling

In this chapter, we present the platform we envisioned for our web-based experience sampling system. Its architecture is modelled on a client-server approach as discussed in chapter 3. Here we present our vision of what a platform such as this one should be like. Further on in this chapter, we present our implementation of the prototype system that was deployed for the purpose of validating the platform and answering our research questions. We detail the inner-workings of the system and its components, explaining how they perform their functions and interact with one another, additionally we expose difficulties that we had to overcome during development. Finally, we explain our work regarding the support for sensing in the mobile client, in particular we detail how delay-tolerant location works. This is a key process of our platform, considering the way in which it fundamentally removes what are major barriers for user adoption.

Over the course of this chapter there are several terms, some already featured in the literature up to this point, which we feel should be clearly defined in scope and nature. This terminology is strongly anchored in user study terminology and although sometimes evident, still needs to be fully elaborated upon so as to avoid misunderstandings. From this point onwards, the following terms will carry the described value:

- **Protocol** refers to the series of steps a user of the platform (researcher or participant) must perform in order to fulfil his respective role. This involves registration, study management, study subscription and participation management.

- **Participant** is a user of the platform that undertakes participation in studies and is therefore subject to, interaction with the platform under participant protocol and, if agreed upon, to study constraints imposed by the researchers. This does not mean however, that a participant cannot withdraw from study whenever he so choses.

- **Researcher** is a user of the platform that interacts with the platform under researcher protocols and whose role it is to create and manage studies. This user would be the main beneficiary of the platform's services, as most studies of this type are of no immediate benefit for participants and are conducted with the intent of sampling phenomena of interest in clarifying research questions.

- **Sample** is a portion of data that constitutes the context and information that we wish to collect regarding the events we wish to study.

In the interest of clarity, the person known in the following literature as the administrator is solely responsible for curating the platform and has no bearing on any other type of proceedings,

this means that said person does not interfere in other affairs or actions that fall to either the participants' or researchers' domain of operation. The same is valid among participants and researchers, naturally.

4.1 Features for researchers

In this section we propose our model for researcher registration protocol and study management and, we present all platform functionality available to researchers, which is entirely web-based.

4.1.1 Registration

Researchers gain access to the platform through registration which happens in the following manner:

- The administrator creates an invitation;
- An e-mail invitation with a link is issued by the platform;
- The researcher follows the link to a registration form where he is asked to fill in the following data:
  - Name;
  - E-mail;
  - Occupation.

After completing these steps a unique identifier (researcher key) is generated. A password is sent and, along with the e-mail, this will allow the researcher to login in to create and/or manage studies.

4.1.2 Study management

Study management is an important part of the researcher's role as a platform user, whether it is creating new studies or managing existing ones. Creating a study will be done by filling a form in the website. This form will collect all the necessary data for a study in accordance to the model presented in 4.1.3. The study is a transient entity and as such, it will be target of various changes during its lifetime. To better convey this idea we present a typical study lifecycle, pictured below (Figure 1). Through it, it is possible to grasp what are the various intervening agents involved in the process of conducting a study and what are the various stages of the process.

A study is created and managed by a researcher. From the moment upon which it is created, participants can receive an invitation and sign up for participation. This will require them to register to the website if they haven't done so already. Having signed up participants are given a key, which will be used conjointly with the study key for login on the mobile application, doing so will result in an activation of sorts for the mobile application. Logging acts as a trigger for downloading study configuration and from that moment on, the application will run in background prompting users to answer questionnaires at times relevant to the object of study. Data gathered during the extent of the study is synchronised via a wireless connection whenever possible (connection type chosen by user).
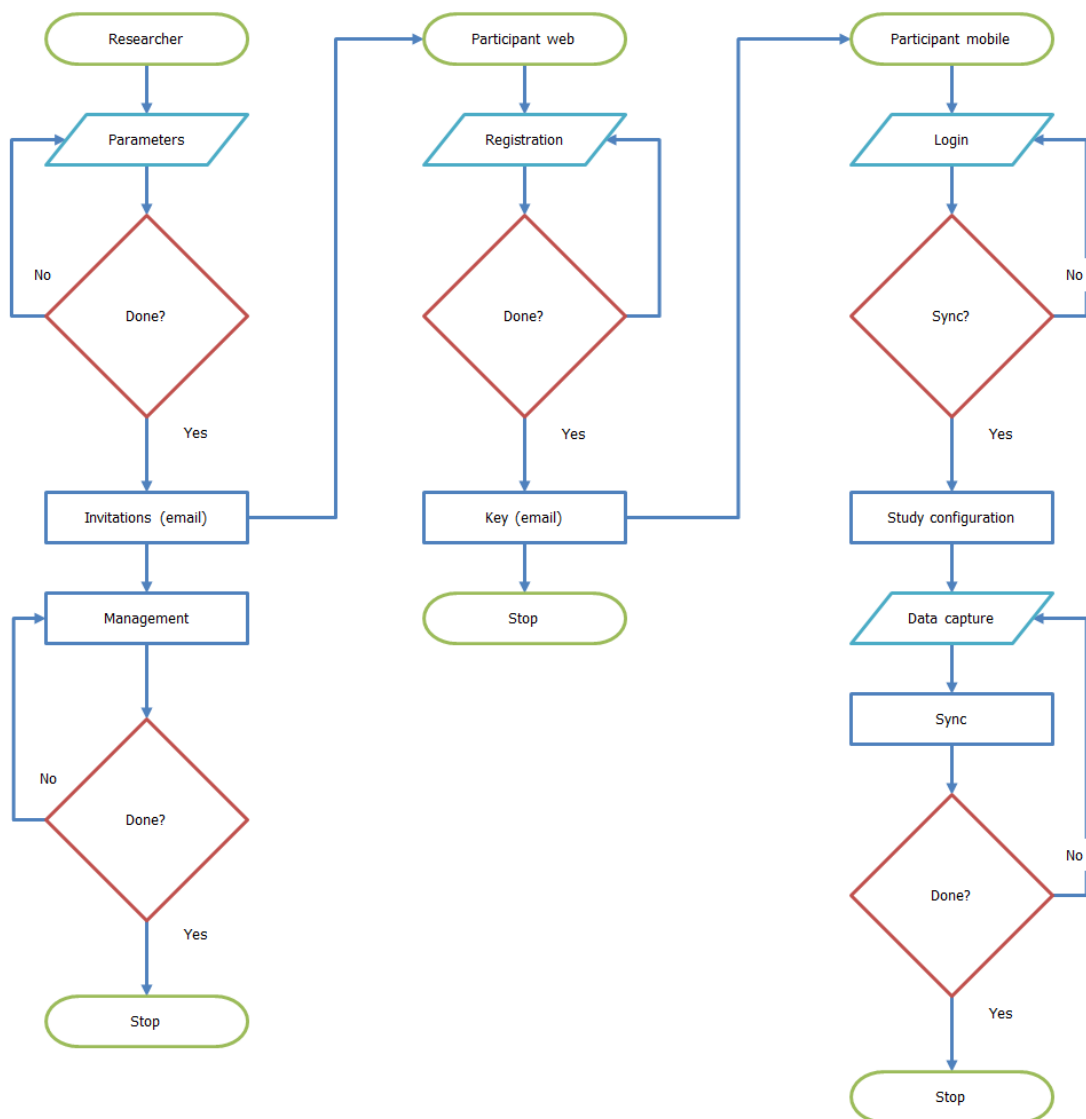
Figure 1 - Study lifecycle

4.1.3   Data model for the studies

When designing a data model for the studies conducted using this type of tool, many considerations have to be factored in, like for instance:

- How do you specify triggering events and what are the supported possibilities?
  - Physical proximity to a service or screen.
- The participants themselves can take initiative and register an experience (e.g., collecting pictures of screens).
- The fact that it is desirable that to have the ability to:
  - Specify the time period during which alerts may go off (e.g., you do not necessarily want alerts to be going off while the participant is sleeping);
  - Specify the number of alerts per time period and overall (e.g., 10 alerts per day, 70 alerts for the duration of the study).
- What type of alerts will the tool support?
  - Random alerts (e.g., from the hours of 9am – 9pm, deliver 10 randomly spaced alerts).
  - Timed alerts (e.g., from the hours of 9am – 9pm, deliver 10 evenly timed alerts).
  - Event-based/user-driven alerts (e.g., when a certain event occurs for the participant, (s)he triggers the questionnaire).
- What to do when there is an alert?
  - Questionnaire (participants fill out a questionnaire whenever they are alerted).
  - Pro-active collecting of photos, sound, text, etc.
  - Automatic sensing.
- Finally, since questionnaires take place during the activities of daily life and participants are alerted several times per day, there is only so much you can ask, a good rule of thumb is <= 2 min./questionnaire.

With these considerations in mind we devised the following data model:

- Study key (unique and automatically generated);
- Study name (user recognisable);
- Researcher key (reference for a registered researcher);
- Duration/Extension (often 1-3 weeks);

- Data collection settings:
    - o Alerts (types, quantity and time periods);
    - o Capture (what type of data will be captured e.g., GPS, audio, video, textual/questionnaire);
    - o Publishing.
- Participants list (no names, participant key only).

Once a study is created, its identification key can then be distributed to participants so that they can sign up. From this point forward, researchers will be able to manage the study from the website, where the following functionality will be available to them:

- Inviting and removing participants from a study;
- Uploading of questionnaires (in XML format) which will be forwarded to participants at an appropriate time;
- Closing the study;
- Changing study data collection settings;
- Consulting information regarding participant activity:
    - o Types of data a user produced during study;
    - o Number of alerts addressed or ignored;
    - o Time periods with greater activity;
    - o Number of data samples generated pro-actively;
    - o View or download answered questionnaires;
    - o View or download other submitted data (photos, videos, audio, etc.).
- Contacting participants (this will allow researchers to encourage people to participate, if necessary).

We feel it is important to refer that even though researchers can view participant data and contact them, a measure of anonymity will be maintained. Being able to contact a participant does not translate into being able to identify who it is.

4.2 Features for study participants

4.2.1 Registration

Participants, akin to researchers, will have to register in the website. Registration proceedings undergo the following sequence:

- The researcher creates invitations for participants;
- E-mail invitations with a study key and a link are issued by the platform;
- Upon following the link, the participant will be presented with a form requesting the following information:
  - Name;
  - E-mail;
  - Age;
  - Sex;
  - Location;
  - Occupation;
- If a user is already registered he can just sign up for the study.

Alternatively the users can register into the platform of their own volition and sign up to participate in a study. When registration is complete, the participant will be issued a unique identifier (participant key) which in conjunction with the study key will allow him to login, both via the website, as well as via the mobile application.

4.2.2 Participation management

As previously mentioned participants can also visit the website. There they will be presented with the following functionality:

- Ending participation in the study (this is to be used in cases that a user wishes to remove himself from the study);
- Consult a summary of the data shared up to that point in the study, this includes:
  - Number of alerts the participant responded to and ignored;
  - Activity periods;
  - Types of data produced.

In addition to this, the user will have a more enhanced control over his participation in the study via the mobile application, where he is allowed to customize his preferences regarding the study. This is discussed in the following section.

4.3   Mobile client features

The mobile client will be the enabler for study conduction as it is the means through which participants receive alerts and, produce data that will be later analysed by the researchers. Ideally, this client should be generic enough that it is able to be used by participants on their mobile phones. The purpose of this is:

- To minimize costs by avoiding issues like having to acquire several devices for users to carry;
- To maximize the possible participant base and increase the likelihood that people actually participate in studies. Mobile phones are something we carry every day and this avoids having to force users to carry an extra device, which is cumbersome.

However, mobile phones nowadays are ever more powerful, capable machines with many sensors built-in (camera, GPS, accelerometer, compass, Bluetooth, etc.). As such, there is something to be said about the possibility of opting for a native client as opposed to a more generic approach or even, combining both. A native client will allow researchers to leverage this computational power and vast array of sensors, in order to collect data that is more rich and comprehensive and, this leads us to believe that we should follow this approach. In doing so, we believe the following to be the list of features the client should have:

- Adaptability (researchers will have the ability to add new questionnaires to the study and those will be forwarded to users whenever possible);
- Archiving the results of captures for synchronisation purposes.
- Audio, photo, video and text input;
- Background location sensing (GPS);
- Background Bluetooth sensing;
- Background Wi-Fi detection;
- Configurable synchronisation options allowing user to control costs (via 3G+Wi-Fi synchronisation happens when needed, via Wi-Fi only whenever possible);
- Pro-active user experience sampling.

4.3.1   Mobile device API and behaviour configuration

The platform's API for the mobile device should rely on a RESTful model, this allows for an easier implementation and it also facilitates communication between client and server. It is a well-known approach and it is generally easier to support then other approaches on mobile devices. This API will allow the client to download the study specification file (XML) that

determines the application's behaviour for the duration of the study. This behaviour is determined by:

- The study ID;
- The start and end dates for the study;
- The number of notifications to deliver;
- The frequency with which they are delivered;
- The content of the questionnaires (including choice answers);
- Whether or not they are associated with a system event that acts as a trigger.

4.4   System architecture

The system architecture, seen below (Figure 2), is comprised of four key components, a website, a database, a web service and a mobile client. Next we will explain what the role of each of these components is, in creating a general purpose user experience study platform.

The website will act as a client for researchers and participants alike, although providing different functionalities (already discussed). As for the database, it will obviously be used as a container for all relevant data. The web service will be the mediator between mobile client and server communications, enabling us to create a point of access to the platform that can be easily shared across different mobile operating systems, especially considering that its API is RESTful. The mobile client will be the agent responsible for capturing data related to the phenomenon under study, delivering questionnaires, uploading data and retrieving study configurations from the server. Server-side, there will be a Perl script responsible for parsing some of the data captured on mobile devices, that data will be relevant to the wireless network landscape (GSM+Wi-Fi) and it will be used to provide locations with which to annotate data.
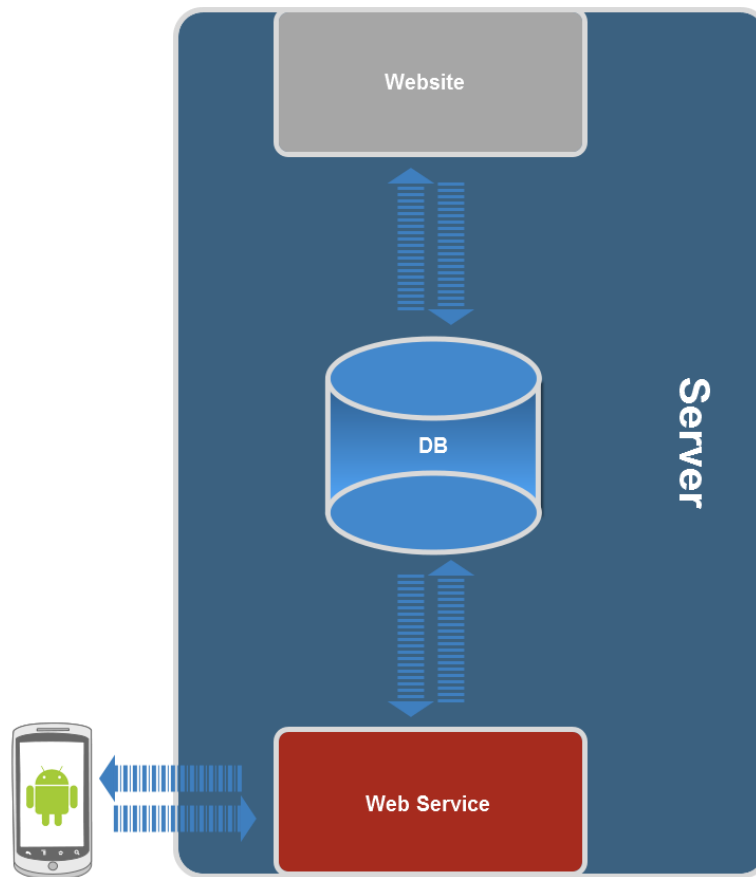
Figure 2 - System architecture

4.5   Implementation

With the platform outlined, we devised a system composed of:

- A database using Microsoft SQL Server 2008;

- A web site written in ASP.NET MVC (Model View Controller) running on an IIS (Internet Information Services) web server ;

- A WCF (Windows Communication Foundation) REST web service;

- A mobile client running on Android.

The database stores all information related to researchers, participants and studies. The website interfaces with the database to provide participants and researchers with their respective information and intended functionality. The android client acts as the means for conveying study data from researchers to participants and vice-versa. Finally, the web service constitutes the bridge between android clients and the rest of the platform, making this, effectively, the link through which data flows.

31

4.5.1 Implementation of web server

The website is implemented using ASP.NET MVC (Model View Controller) which is a part of the ASP.NET Web application framework. An MVC Application is designed and implemented using the following three attributes:

- Model: The model contains the core information for an application. This includes the data and validation rules as well as data access and aggregation logic.

- View: The view encapsulates the presentation of the application, and in ASP.NET this is typically the HTML markup.

- Controller: The controller contains the control-flow logic. It interacts with the Model and Views to control the flow of information and execution of the application.

As for the RESTful web service, it is implemented in WCF (Windows Communication Foundation) which is a framework for building service-oriented applications. Using WCF, you can send data as asynchronous messages from one service endpoint to another, so this will be the link between mobile clients and the server part of the platform. The web service will enable clients to login and data to be exchanged and synchronised.

Web architecture

Figure 3 (below) outlines the architecture of the web portion of the platform and, as is clearly visible, the main components are the web site, web service, database and the Windows system service. The website, which acts as a portal for researchers and participants alike, interacts with the database solely to store any and all information regarding users of the platform, as well as information regarding any studies. As for the web service, it will act as an intermediary between the mobile client and the server, allowing the platform to authenticate clients, before they download study specification files and store study data on the server (questionnaires annotated with network landscape data). Study data uploaded by mobile clients is, as we know, in an intermediate state because it is not yet annotated with location but with network landscape data that will allow us to estimate a location, this is the fulcrum of our delay-tolerant approach. Here, at the server, the windows system service running in the background will scan study directories (hourly) for study data that has not yet been processed, when such data is found, the service runs a Perl script that will parse and treat all the data, extracting all information relevant for estimating device location and producing files with questionnaire data and location coordinates. This is the data that is then used to perform analysis with the intent of answering a study's research questions.
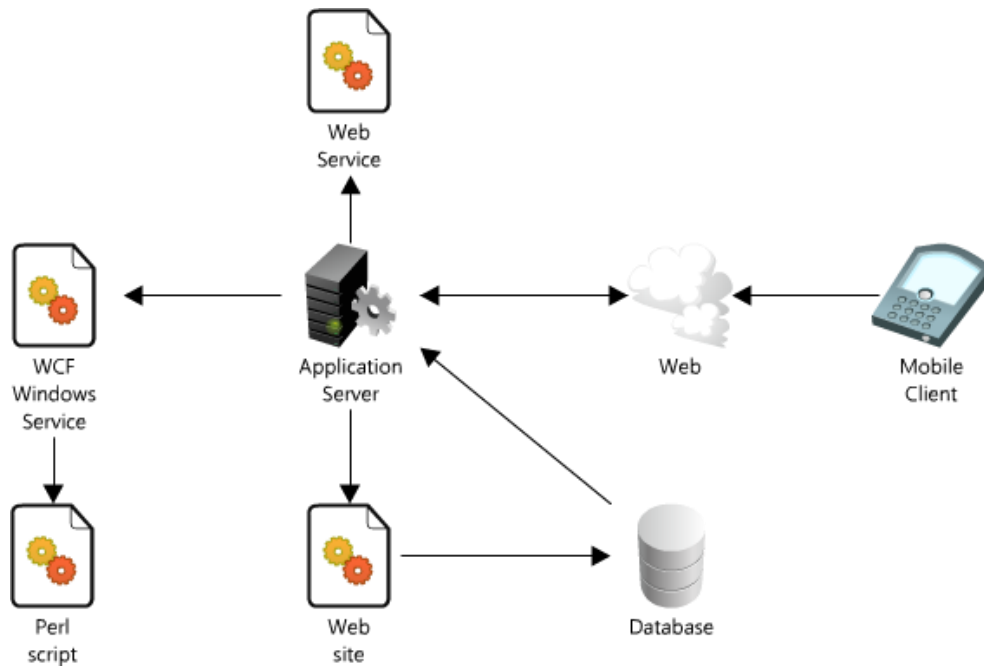
Figure 3 – Web system architecture

Data model

In Figure 4 (below), we can see the entity-relationship model of the database supporting the platform's website.
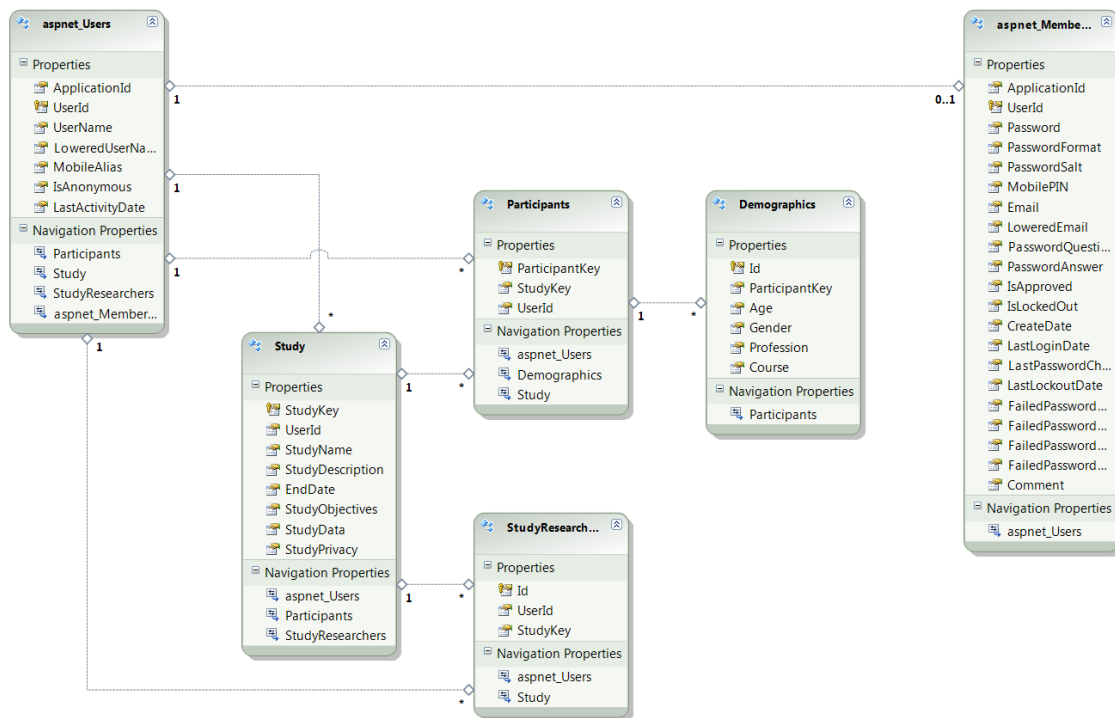

Figure 4 - Database entity-relationship model

The database, as can be seen above, is composed of tables that support user and study information, but also of tables that are automatically generated by the ASP.NET Membership Framework. This framework was used as a means to handle registration, login/logout and role assignment (researchers, participants) procedures, as it provides an already developed and solid basis for these necessary mechanisms and, it features robust security mechanisms for storing sensitive data such as passwords, which are hashed and salted [39], [40] to ensure that they are not accessible by undesired third parties.

Notable development issues

During the development stages of the web-based side of the platform there were no technically challenging issues that impeded progress. The only two issues that surfaced only withheld progress for minute extensions of time and were quickly and easily solved. The first issue was related to the reception and processing of files, the content observable within a file's stream on the server side, namely the delimiters differed slightly from what was observable on the mobile client and this made us have to adjust the code to compensate for that fact. The second issue was related with how an ASP.NET application pool is managed by IIS, given the policies applied on the machining running IIS. The web service was supposed to be responsible for running the Perl script hourly but this was not happening because IIS was not configured to maintain inactive applications running in the application pool. Even when policies were changed to avert this, the changes did not seem to take effect, as such, we were forced to develop an unplanned component, the WCF service[1,2] that came to run in the background and act as the entity responsible for running the Perl script on an hourly basis. Fortunately, this was very easy to accomplish.

4.5.2 Implementation of the mobile client

The Android client application is installed through the Android Market (now Google Play Store), once this is done, the user can start interacting with the application. On start-up the user is prompted to enter his participant key and the study key, which will be validated in the server, via the web service. At this point the application will download the study's configuration data (XML file) and configures itself to run the study. This application will leverage the device's sensing

---

[1] http://msdn.microsoft.com/en-us/library/ms729746.aspx

[2] http://blog.themobilebrand.com/technology/tutorial-building-a-windows-service-application/

capabilities to gather information that is pertinent towards the study, or even to trigger an alert that will prompt the user to take some action relative to the study. All text data collected by the application is stored in XML format and, in conjunction with the rest of the captured data, sent to the server at an appropriate time according to user preferences. The client application will feature the functionalities listed in section 4.3.

Below, in Figure 5 we can see a diagram of the architecture of the Android mobile client application, split into its various core components. As with any Android application there is an Activity (Main Activity) that is launched when the application is started by a user. Once started the application will check Android's account manager framework for credentials, if these are not present then the user will be prompted for the credentials he was assigned upon enrolling in a study. Said prompt comes courtesy of the Authenticator Activity which is bound to an Authenticator Service that validates the credentials on our server and then stores them as an account in the system's account manager. This step is skipped if there are authentication credentials already present, in which case the user is presented with a tabbed activity that will show him a graph of the application's synchronisation status, or if he switches tab, the Preferences Activity. Here the user can customize his preferences regarding the type of data connection used to synchronise data (this activity is also shown upon successfully validating authentication credentials, or it can be brought up from the Menu available via the Menu key).
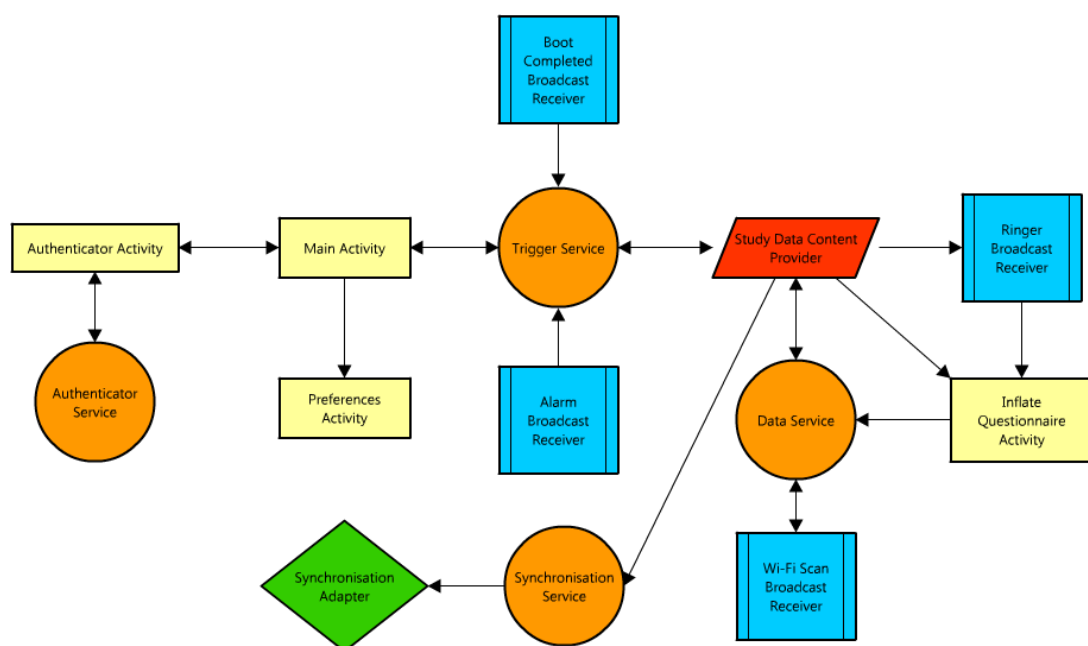
Figure 5 – Mobile client architecture

Successful validation of a user's authentication credentials triggers background processes in the application that will define the application's behaviour throughout the study and brings users to the settings screen where they can configure their connectivity preferences.

Notification management

After validation, the Main Activity requests that the Trigger Service retrieve the study's specification, which is an XML file that stores information on several aspects of the study (questions, number of notifications, type of notifications, study duration, study key). Having retrieved this file it is then parsed and its information is stored in the device's SQLite database via the Study Data Content Provider. If the information parsed dictates that the application will perform questionnaires that are not linked to a device event, but rather are distributed (evenly or randomly) across time, then the Trigger Service will run an algorithm that will generate and arm system alarms to trigger these questionnaires. The information regarding these alarms is also stored in the device's database because in the event of a system reboot, these alarms are no longer armed, as such the application was coded to capture the Boot Completed Intent and when it does, it will access the database via the Study Data Content Provider and re-arm any alarms that have not already been triggered. Any triggered alarm will be captured by the Alarm Broadcast Receiver and that will cause the Trigger Service to issue a system notification (Figure 6) that will prompt the user to answer a questionnaire, the same is valid for the Ringer Broadcast receiver.

Figure 6 – System bar and notification pane

If left unattended this notification (Figure 6) will disappear when 5 minutes have elapsed since its time of origin, this is a mechanism designed to accommodate for the fact that users may not always wish to be interrupted, which as previously discussed is a very important concept to enforce. In conjunction with this, the notification also abides by system settings that the user has defined, so it the mobile phone is set to vibrate, there will be no audible notification and if it is set to silence, there will be no audio or vibration feedback. Tapping the aforementioned notification will cause the application to access the relevant questionnaire's information via the Study Data Content Provider and, this information is then used by the Inflate Questionnaire Activity to create a dynamic layout that will adapt the view to the type of question that it is presenting to the user, in this case multiple-answer type questions with variable number of choice-answers. Completing the questionnaire takes the application to a state in which it will request that the Data Service gather information on the radio network landscape (GSM+Wi-Fi), once it has done so the Wi-Fi Scan Receiver's broadcast will be caught by Data Service, and that in turn, will access the Study Data Content Provider and store the combination of this information (questionnaire answers and network landscape data) that later gets sent to the server, when a connection is available.

Data synchronisation

The synchronisation of data is carried out by the system's synchronisation framework when an active connection is detected and faults in synchronisation are handled appropriately by the system (exponential back-off and retry). When synchronisation is triggered the Synchronisation Adapter checks the account manager for valid credentials and, if it they are present, it then works in tandem with the Synchronisation Service and the Study Data Content Provider to upload to the server any relevant data that is stored locally. When participating in a study, users have in their main screen a pie chart that lets them have a sense of the synchronisations status of the data they have produced.

Delay-tolerant positioning for the mobile client

Delay-tolerant positioning works by storing on the device's local storage all the GSM radio and Wi-Fi data generated, that is used by a location API to determine position and, when a connection becomes available, a batch of GSM and Wi-Fi information is sent over to a server that will then use that information to calculate the positions. By gathering information about GSM cells and Wi-Fi access points locally, it can later be used server-side to query a location API and obtain device positioning at the moment the information was captured. This approach does not make any use of the GPS and works very well with only occasional connectivity, both indoors and outdoors, even having the potential to circumvent data costs, depending on whether or not the user configures the application to only use Wi-Fi (which is the default behaviour) or any type of data communications.

Notable development issues

In this respect, the mobile side of the platform was more challenging than the web side; there were some technical issues that hampered steady progress. First of which, came into light when initially researching the development of features that did not come to take part in the prototype. The possibility of adding media (audio, pictures or video) to a response on a questionnaire, or having it be the response, is something that was initially pursued and will certainly be pursued in the future. Nonetheless, it is important to mention that this might present several issues and not all of them of technical nature. Managing media size and media upload to a server, whilst trying to keep data costs to a minimum (if not on Wi-Fi) is a challenge in and of itself but what we intend to bring to light are technical issues in Android. Not only were the APIs that handle the various media types, at the time of development, very distinct, poorly documented and non-

organic, as there were issues with how things were handled by the platform itself. For instance, on Android 2.1 which at that time was the version on our development device, the API that allowed us to interact with the camera and take a picture from within our application resulted in a picture that was limited in resolution, in spite of us enforcing parameters to maintain system settings on resolution. This issue was resolved when our device received an upgrade to version 2.2. The API that allowed us to perform audio recordings within an application would simply not work. These are issues that are most likely resolved given that Google's documentation on Android APIs is nowadays much improved and also, taking into consideration the fact that present day Android versions have feature-rich media APIs.

A second issue that arose during development was tied with the uploading of data captured on the device to the server, we needed for the mechanism to be reliable, to take advantage of connectivity whenever available and according to user settings and, we needed it to guarantee retries in case of failure but without having to implement complex logic so as to avoid constantly hammering the system with retry requests and impacting battery life. The solution came via a Sync Adapter[3,4], an Android application component that interfaces with the system's account manager. This allowed us to use the account manager to store study credentials safely and use them to authenticate requests to the server, particularly data upload requests and it also provided us with guarantees for request retries. When a device has connectivity and there is out-of-sync data or data to upload, the Sync Adapter tries to synchronise/upload data and if it fails, it has an exponential back-off algorithm, meaning this meets our requirements to their fullest extent. Unfortunately, at the time of development this component was very poorly documented to the extent that most good information on it was available from outside Google documentation. This was particularly troublesome as this component was not trivial to develop and could potentially have been even more complex had our needs been different. Yet another issue plagued this component as well, on Android 2.1 any mistake that caused unexpected behaviour in the Sync Adapter would cause the device to reboot itself which is to say the least, troublesome for development. It is too another instance of a problem that was resolved with version 2.2 of Android.

[3] http://ericmiles.wordpress.com/2010/09/22/connecting-the-dots-with-android-syncadapter/

[4] http://www.c99.org/2010/01/23/writing-an-android-sync-provider-part-2/

5 Evaluation

In this chapter we explain that which was our evaluation process for two distinct steps in our investigation work, one is logically the system we intended to build and, the other, is our delay-tolerant positioning method (a piece of the final system) which also needed validation, in order to guarantee that it could serve our intended scenarios. In what concerns the evaluation of the system as a whole we start by discussing our data collection and analysis processes, were we approach the issues of data validity and data privacy. We describe our target group for the study, as well as the study setting in which they would perform their roles. The study workflow is outlined, from participant invitation and registration to application distribution and download, additionally we explain what the lifecycle of a study is. Next we present study rollout data throughout the time period in which the study was under effect, this data features things such as active device installations, application uninstalls and other metrics, these are used as a means by which we can measure adoptions among the target population, as well as other phenomena, like early quitters. When the study ended we performed some follow-up interviews regarding user experience and we also present the results of these in this section. Furthermore, we discuss the results of the deployment study as a whole. In regards to the delay-tolerant positioning process we present our process for evaluating the validity of this positioning technique, which naturally involves describing the application created for data collection, the collection process itself, our data processing methodology and, finally, our results.

5.1 Running a study on ringer mode changes

With the platform ready the need arose for system validation. For this purpose we first conducted some internal testing with fellow volunteers that were meant to be a system-check of sorts designed to test the workflow of studies from the user's perspective, as well as uncovering any underlying bugs that had not yet been caught. After completing this, we conducted a study test in real-world conditions and set out to answer, not only research questions pertaining to this work, but also research questions regarding the study we created for testing.

We devised a study that aimed to establish if there was a definite need for a context-aware application that would manage mobile phones' ringer mode. The study was supported by our Android application that collects data (GSM+Wi-Fi, for location estimation) and delivers questionnaires to the users/participants. The principle behind the study is that, if we could identify geographical hotspots for ringer mode changes and, tie them to a particular sound profile

and social environment, we would have established the existence of contextual parameters that can be used to trigger this change automatically.

The study will be conducted through the delivery of experience sampling questionnaires and the collecting of data for location estimation, through our delay-tolerant process. These questionnaires will prompt the users with one of two questions, with multiple choice answers.

1. Porque mudou o perfil de som do telemóvel?

    o   Reunião

    o   Descanso

    o   Local de silêncio

2. Este é um local onde gostaria sempre de ter o telemóvel em silêncio?

    o   Sim

    o   Não

The results of the study could lead us to conclude that there is a definite need for a context-aware application that manages mobile phones' ringer mode. This could point towards future research needs/possibilities in key aspects, such as, the inference and definition of an appropriate context for a ringer mode change, the use of location sensing techniques based on multiple sensor inputs and their adequacy in terms of accuracy, precision, monetary and energetic costs, as well as other relevant aspects. Another line of research would be to investigate how calendar data could potentially be of use as a source for contextual information that, in conjunction with the information already used, could better aid in deciding when to trigger an automatic ringer mode change and even in providing automatic feedback via SMS, about why the user has not answered a call.

5.1.1   Data collection

The first data to be collected was of demographical nature, such as **name**, **gender**, **age**, **occupation**, **degree** during the on-line registration procedure. With this part of the process concluded, a user could then install the tool (mobile application) that would be the basis for further data acquisition. Said tool would then gather network landscape data from GSM cells (Cell ID, LAC - Location Area Code, MNC – Mobile Network Code, MCC – Mobile Country Code, Signal level) and Wi-Fi networks (BSSIDs – MAC Address, Signal level), as this information would be paramount in estimating the device's geolocation. Additionally, the tool would also gather data

about the ringer mode selected by the user, the time of day of the occurrence and the participant's answer to the question served up by the application (if any).

The data collection process is triggered on ringer mode changes, which means that there will be no other data collected during the study at any time, other than the already mentioned data. The collection process involves accessing system resources, through available APIs, to gather data on the device's ringer mode, the GSM and Wi-Fi landscapes. Furthermore, it should be noted that this process does not occur every time a ringer mode change is performed. The collection process only occurs while the study is on-going and is only triggered every 5 times a change is performed or if more than 48h have occurred since the last change, whichever condition is met first.

5.1.2 Processes for analysing the data

Data regarding the study will be anonymous, this means, that it cannot be traced to a particular user registered to the website, it can only be tied to a user's participation profile, which is composed of demographic data that cannot be used to identify him. Also, the mobile application will only collect network landscape information at the time of an ESM event occurrence and not continuously, so a user's location is never continuously monitored and can never really be known in real-time, as this data will live on the device until it is uploaded to the server and eventually processed to obtain an actual location. What the study sets out to accomplish is not a location tracking tool it is merely to study what was already detailed. It is also important to note that, data collected during the study will be destroyed after the study has been completed and the data fully analysed and, participants will know at all steps what it is that we are asking them to volunteer.

5.1.3 Issues of validity

Result validity is always an issue in any live experiment, however, the use of experience sampling questionnaires aims to ensure that the obtained results are valid, as the very process of inquiring as to what are the participant's intentions at the time of the event was founded with the very purpose of improving data validity. It not only circumvents the need to have a researcher present, to conduct the inquiry, as it insures that the data gathered is more relevant and does not suffer from cognitive bias, seeing as it is collected in a period of time of close proximity to the event we wish to sample.

In order to ensure that our data will be minimally representative of the population at campus we will be targeting a participant base of about twenty (20) people and, we will be collecting data for about 1 month. Location data will be analysed and compiled into hotspots, any location within a proximity of 100m of any other will be considered to belong to the same hotspot.

5.1.4 Data privacy

Information pertaining to the network landscape will be stored in a file on the device. This file will be transferred to a server for processing, at a convenient time (depending on connectivity availability and user preference). Participant information will be stored in a database and will only be accessible to either the participant in question, through the website, or the researcher(s) conducting the study.

Clear information on the study and the tool will be available on the website (http://applications.instantplaces.org/esmsite/Home/About), should the need arise for further clarification, there is an e-mail listed on the website for such that purpose. Also, before installation, the Android market always presents the user with the list of system resources the application requires permission for, the user may choose to allow or deny this request. If denied, the application will not install, as per standard Android market practice. The targeted audience does not include any members of vulnerable user groups.

5.1.5 Participants (target group) and recruitment

The research's subject pool will be composed of members of the academic community at Universidade do Minho. The target group will be the student community, in particular a specific student community (LTSI – Licenciatura em Tecnologias e Sistemas de Informação). Using a specific sample of the academic community, with the same or similar schedules and obligations, would allow us to better extrapolate patterns related to the behaviour we aimed to sample. Recruitment will be made via e-mail and a blog (Human Digital Systems) posting. These communications will contain a brief explanation of the study, links to the website for registration and to Android market for application download. Participants in the study will be rewarded with a €10 voucher from FNAC.

5.1.6 Study setting

The physical setting can be any point in the world, so long as the application is installed, given that it triggers data collection at any point a ringer mode change occurs.

The social environment can be described as a typical academic environment. The space is composed by many buildings (departments, schools, administrative buildings), and it is inhabited by teachers, students and staff. The nature of the interrelationships between all of those agents, during the course of their activities at campus, is the catalyst that triggers ringer mode changes (e.g., attending class, going to a meeting, etc.). So the social environment can be described as one with a mix of formal and informal happenings, depending on the relationship between agents and the context of their interactions.

5.1.7   Study workflow

Participant invitation and registration

Participants will be invited via e-mail. An invitation e-mail, with an explanation of the study, will be forwarded to University of Minho's academic community. The aforementioned e-mail will contain a link to the platform's web page, where users can register if they have no account yet.

Figure 7 − Web site homepage

Figure 8 – Web site registration form

Figure 9 – Successful completion of registration

Figure 10 – Registration e-mail notification

Upon registration they will be prompted to choose their username and password (Figure 8) which will be e-mailed to them (Figure 10). We are naturally well aware that sending plain-text passwords via e-mail is an unsafe practice and this is something that would never feature in a production system, it was only meant to facilitate proceedings for the prototype. When deploying the system into production, this practice would no longer take place and instead users would have the option to reset their password which is something that the ASP.NET Membership framework also contemplates. After completing registration (Figure 9), they can enrol as participants in a study by choosing from the list of available studies (Figure 11). Choosing to do so, users will be presented with a study description (Figure 12) and will be prompted to supply the demographic data previously detailed (Figure 13).

Figure 11 – List of studies

Figure 12 – Study description

Figure 13 – Study subscription form

Figure 14 – Successful study subscription

Figure 15 – List of subscribed studies

TESS Web: Dados de subscrição

TESS Web <acoelho@dsi.uminho.pt>

Enviado: seg 17-09-2012 15:33
Para: André

Olá caro(a) André,

Obrigado por se registar e participar no estudo Utilização dos perfis de som nos telemóveis. De seguida pode encontrar transcrita a sua informação de acesso:

Chave de participação: c4c1c946
Chave de estudo: 6106929c

Atentamente,
TESS Web

Departamento de Sistemas de Informação,
Universidade do Minho
17-09-2012 15:33:09

Figure 16 – Study subscription e-mail notification

Upon completion of this process (Figure 14), an automatic e-mail will be sent to the users (as in Figure 16) with the study key and their participant key, so they can perform authentication on the mobile client. Should, for whatever reason, a user decide to terminate his participation ahead of schedule, they can do so from the management options available in the list of subscribed studies (Figure 15). However, they should be aware that any data submitted up to that point will not be erased.

Distribution and download

The application will be distributed to participants via Android Market download. The application page can be found at: https://market.android.com/details?id=com.esmringer.app, this URL is accessible from any browser. To download the application to your device you can use one of the following methods:

- Use the search function in the Android Market and type TESS;

- Navigate to market://search?q=pname:com.esmringer.app on your Android device;

- Or use a Barcode scanner application to download via the following QR code.

Figure 17 - QR code for TESS on Android market

- After having installed the application, users can introduce their authentication data (study key and participant key) in the application. From then on, the application will prompt the participants to answer a questionnaire whenever a ringer mode change is triggered (under the previously specified conditions), additionally, the application will collect network landscape data at the time of the occurrence. This data will be stored on the mobile device and uploaded to the server for processing, at a convenient time (according to connectivity availability and options).

Ringer mode study lifecycle

Below, in Figure 18, we can observe a flowchart for the lifecycle of a study. Through it, it is possible to grasp what are the various intervening agents involved in the process of conducting a study and what are the various stages of the process. A study is created and managed by a researcher. From the moment upon which it is created, users can sign up for participation. This will require them to register to the website if they haven't already done so. Having signed up to a study, participants are given a participant key, which will be used conjointly with the study key for login on the mobile application, doing so will result in an activation of sorts for the mobile application.
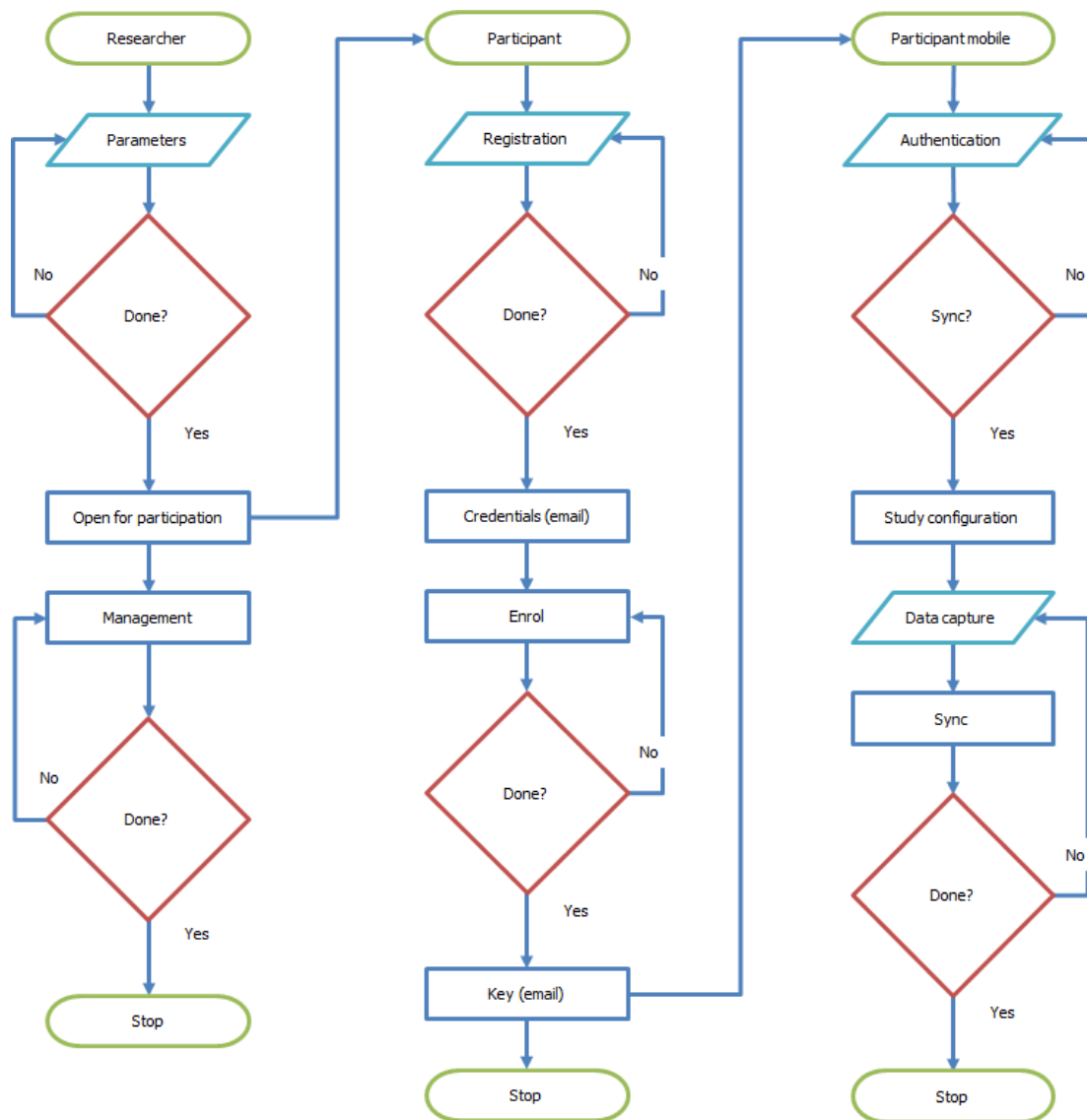
Figure 18 – Ringer mode study lifecycle

Logging on acts as a trigger for downloading study configuration and from that moment on, the application will run in the background prompting users to answer questionnaires at times relevant to the object of study. Data gathered during the extent of the study is synchronised via a wireless connection whenever possible (connection type chosen by user).

5.1.8   Study rollout data

As part of the Android Market developer console we have access to detailed application related data. This is useful in gauging user adoption and, to some extent, user involvement during the study, although a more faithful metric in this aspect consisted simply of observing per-user synchronised questionnaires server-side. Having said that, as we can see below in figure Figure 19, there were 7 active device installations during the time period where the study took

place (1 month). This number is somewhat misleading, as there were only 2 enrolled participants in the study, all other active installations were from laboratory partners that had tested the application previously in pre-rollout studies and our own device.

Statistics for TESS (com.esmringer.app)   Export as CSV

Active device installs ▾ ⓘ for Feb 29, 2012 - Mar 31, 2012   Show: last month 3m 6m 1y all

7,0
6,5
6,0
5,5
          Mar 8, 2012   Mar 16, 2012   Mar 24, 2012

| Android Version | Device | Country | Language | App Version | Carrier |

Active device installs by Android Version

4
3
2
1
          Mar 8, 2012   Mar 16, 2012   Mar 24, 2012

Active device installs on March 31, 2012

	Your app		All apps in Tools	Top 10 Android Versions for Tools	
☑ Android 2.2	3	50.00 %	23.89 %	Android 2.3.3+	51.57 %
☑ Android 2.3.3+	1	16.67 %	51.57 %	Android 2.2	23.89 %
☑ Android 2.1	1	16.67 %	6.04 %	Android 2.1	6.04 %
Others	1	16.67 %		Android 4.0.3	1.90 %
				Android 3.2	1.29 %
				Android 3.1	0.83 %
				Android 1.6	0.78 %
				Android 4.0	0.70 %
				Android 1.5	0.46 %
				Android 2.3	0.38 %

Figure 19 – Active device install data for the duration of the study

In Figure 20 we can observe data that corroborates our explanation for the number of device installs vis-à-vis the actual number of study participants. This figure shows us that there was 1 installation at the very beginning of the study's time period. At this point one might wonder why there are not 2 installations in this graphic, given that there were 2 participants. The explanation is quite simple, the other participant had an Android device that had not been sanctioned by Google, this meant he had no access to the Android Market place, so we had to provide him with the .apk file for direct installation on his device, therefore, circumventing the marketplace and the possibility for him to generate marketplace data.

Statistics for TESS (com.esmringer.app) Export as CSV

Daily device installs ▾ ⓘ for Feb 29, 2012 - Mar 31, 2012   Show: last month 3m 6m 1y all

1,00
0,75
0,50
0,25
        Mar 8, 2012   Mar 16, 2012   Mar 24, 2012

| Android Version | Device | Country | Language | App Version | Carrier |

Daily device installs by Android Version

1,00
0,75
0,50
0,25
        Mar 8, 2012   Mar 16, 2012   Mar 24, 2012

Daily device installs on March 31, 2012

	Your app		All apps in Tools
☑ Android 2.2	1	100.00 %	13.37 %

Top 10 Android Versions for Tools	
Android 2.3.3+	46.95 %
Android 2.2	13.37 %
Android 4.0.3	7.09 %
Android 2.1	3.17 %
Android 4.0	2.18 %
Android 3.2	1.83 %
Android 3.1	0.61 %
Android 2.3	0.48 %
Android 1.5	0.30 %
Android 1.6	0.27 %

Figure 20 – Daily device install data for the duration of the study

The next figure (Figure 21) reveals that there was 1 uninstallation of the application during the time frame of the study. This could have been a troublesome occurrence, meaning that we could have been left with just 1 participant early-on in the study. Fortunately, that was not so, it was merely the case of us uninstalling the application from our own device, in order to be able to install a newer version so we could conduct preliminary tests regarding future features of the client. An Android market application is digitally signed with a key, reason for which a debug version of the same application cannot override the market installed application, since it is signed with a different key, automatically generated by the ADT (Android Development Tools for Eclipse) plugin.

Figure 21 – Daily device uninstall data for the duration of the study

Below (Figure 22) we can see data for the total number of user installations, one safeguard should be mentioned, this data reflects what happened during the time period of the study but, given that it reports on total user installations the numbers shown pertain to the total lifetime of the application since it became available in the Android market. Naturally, this data reflects many reinstallations of different versions by our lab partners that acted as alpha and beta testers, and some runaway installations that resulted in Android market comments such as: "I don't know if this app is any good, but it is only in Spanish (I think) so as I am English it is no use to me. Nothing wrong with Spanish language apps, just it should be clear to the user before they try to download it. I.e. write the description in Spanish!" The data in Figure 22 shows us what versions of Android our application was installed on, as is observable Android 2.1 was the most common version.

Total user installs   ⊙ for Feb 29, 2012 - Mar 31, 2012   Show: last month 3m 6m 1y all

25
24   Mar 31, 2012
23   ▪ Total user installs: 24
22
        Mar 8, 2012   Mar 16, 2012   Mar 24, 2012

| Android Version | Device | Country | Language | App Version | Carrier |

Total user installs by Android Version

12
9
6
3
        Mar 8, 2012   Mar 16, 2012   Mar 24, 2012
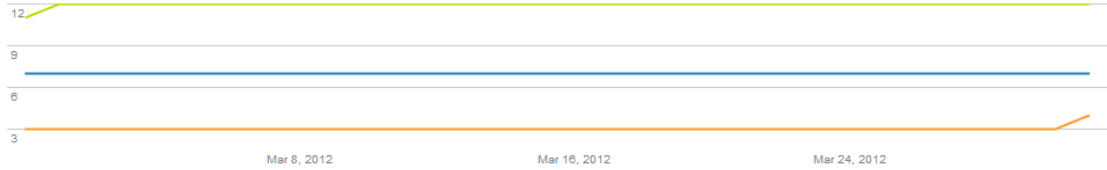
Total user installs on March 31, 2012

	Your app		All apps in Tools
☑ Android 2.1	12	50.00 %	8.74 %
☑ Android 2.3.3+	7	29.17 %	41.98 %
☑ Android 2.2	4	16.67 %	23.58 %
☐ Others	1	4.17 %	

Top 10 Android Versions for Tools	
Android 2.3.3+	41.98 %
Android 2.2	23.58 %
Android 2.1	8.74 %
Android 4.0.3	1.42 %
Android 1.6	1.29 %
Android 1.5	0.97 %
Android 3.2	0.74 %
Android 3.1	0.49 %
Android 4.0	0.44 %
Android 2.3	0.39 %
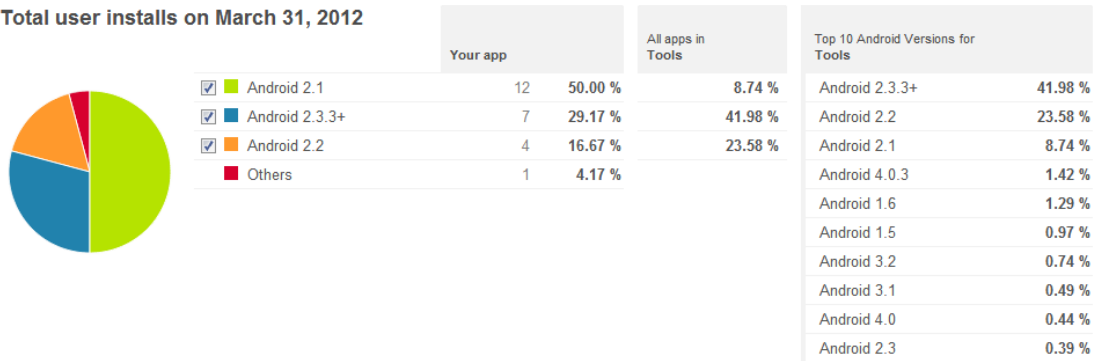
Figure 22 – Total user install data for the duration of the study

5.1.9   Follow-up interviews

As important as it is to validate the platform we created, as a whole, it is also important to follow-up on user experience. Reason for which, at the end of the study participants were to be interviewed in order to garner feedback on their experience and suggestions on how to improve processes. For this purpose we devised two scripts, one for active participants and one for inactive participants (people who would sign-up to the platform and enrol in the study but produce little or no data). These scripts were to be used when conducting the interviews but not to be followed to the letter, the answers provided by the participants would mould the course of the interview. The scripts for these interviews can be consulted in Annexe D, sections 9.4.1 and 9.4.2.

5.2 Results for the ringer mode study

The initial round of testing went without a hitch in what regards to the web side of things (user sign-up and study enrolment). On the mobile side of things, a teething issue was brought to our attention, data from users with Android v2.1 devices was not being synchronised with the server. The application was coded in such a way that it takes advantage of a component in the Android framework called a synchronisation adapter. This component automatically synchronises data in the background if a connection is available, however, it only works from Android v2.2 upwards and under certain conditions, like having background data and automatic synchronisation active (system settings), which we caught in early testing, reason for which we included a manual synchronisation mechanism in the application, that mechanism was then validated in the second round of testing.

As there were no outstanding issues caught in the second round of testing the message calling for participants was passed along to the target academic population during a class and through a posting in their e-learning platform alas, and even though candidates were paid, we only got two subjects to enrol, which is a number far inferior to our initial target. Reason for which there are no major findings to report in regards to the study's own research questions, as data from two users can hardly be used to establish any pattern. Beyond this, we can report that the study ran its course without major faults, although there were some minor issues worth mentioning. The table below provides us with a general overview of the users' feedback on the major aspects that we wanted to evaluate.

Subject	Registration	Installation	Information	Privacy	Configuration	Intrusiveness
A	Easy	Easy	Clear	Unaffected	Default	Negligible
B	Easy	Difficult	Clear	Unaffected	Easy	Minor

Table 1 – Study subject assessment on critical aspects of the experience

On a more detailed level we can report that participant A had an issue related to the synchronisation mechanism and it caused her to spend the length of the study without (apparently) ever producing data, this lead us to believe she had lost interest in participating. As it turns out, for some reason, her device had failed to synchronise data to the server. We could still retrieve her data afterwards but we could not determine why the mechanisms for synchronisation had failed, given that she had uninstalled the application and had no additional feedback to provide on the subject. Outside this technical aspect interviewing participant A revealed that she felt that although questionnaires where clear and concise, she would have

appreciated having more choices for an answer in question 1 (above in section 9.4.1), or the possibility of writing her own answer. Also, she expressed a wish of having more detailed information about when questionnaires would be served, as she had no prior knowledge of this she assumed every ringer mode change would prompt a questionnaire, and since that was not the case, every time she changed her ringer mode and no questionnaire was served, she proceeded to change it multiple times until she was met with a notification to fill-out a questionnaire. This was not the intended behaviour and is also the reason why this participant ended up generating an enormous amount of data. In regards to being aware of the state of her participation she said that she considered that the synchronisation pie-chart graph would have sufficed, had she been able to synchronise data. When questioned about her awareness of how location data gathering was being done by the application the user did not recall any information pertaining to this so she was not able to provide feedback. No further points were raised or suggestions made by this subject.

In what concerns participant B's experience, we can detail that he had a device that was not sanctioned by Google so he had no access to the Android Market therefore, the application had to be installed via the .apk we e-mailed to him, circumventing the default distribution channel. This participant remarked that the application was easy to configure and use but unlike participant A, he felt that some of the notifications occurred in inopportune moments (class, cinema), even though he said that in general the application was minimally intrusive. Much like participant A, participant B also expressed a will to have the possibility of writing his own reply to some questions. Concerning awareness about his participation this user revealed that he felt the pie-chart graph was a good indicator of whether or not things were progressing but perhaps additional information could also be provided. In querying him about the method for location data gathering in the application we were told that it facilitated his participation in the study. As a final suggestion he added that we should perhaps provide a client for iOS devices.

In general terms feedback from participants was positive, there appeared to be no issues raised with the processes of signing-up to the platform and enrolling in a study. Regarding application distribution and installation there were mixed opinions but there was a motive behind that. Information on all fronts (web and mobile) seems to be clear and concise, although perhaps some more detail in the study's description page could perhaps be needed so as to ward off unintended use of the application like participant A performed. Participants considered that their

privacy was not greatly affected by their participation in the study and the application seems to not cause much disturbance in the subjects' daily lives.

5.3   Evaluating delay-tolerant positioning

Naturally as we had to come up with a different way to perform location estimation, we needed to assess the viability of our delay-tolerant approach for location-based logging. For this purpose, we conducted a specific evaluation process which we will proceed to describe. In particular, we aimed to compare the level of accuracy with what could be obtained directly on the device and also to study the effect of other variables in that accuracy, such as the effect of time spent at the location before the location estimation and the nature of the Wi-Fi landscape at the point where location was estimated.

To investigate the potential and limitations of delay-tolerant positioning in an off-the-shelf mobile phone, we devised a controlled experiment in which we conducted multiple positioning calculations at known locations using our own location data collection application to obtain real-time and delay-tolerant positions.

5.3.1   The data collection application

To support the data collection process, we have developed an Android application to generate position records. The Android platform support the Google location API[5] that allows devices to calculate their position using multiple types of data providers. In devices with a GPS chip, the GPS provider determines location with a high degree of accuracy and no data costs, but with a high energy cost. The network provider makes a combined use of GSM and Wi-Fi data by gathering data from two Android APIs: TelephonyManager API and WifiManager API. The first generates information about the cellular network, e.g. local Cell IDs and respective signal level, for both the cell that the device is connected to, as well as neighbouring cells. The other enables the device to perform scans of the surrounding Wi-Fi radio landscape and retrieve the BSSIDs and signal levels of nearby Wi-Fi access points. When an application calls requestLocationUpdates() in the LocationManager API[6], information from all the active providers is obtained, sent to an undisclosed Google service and the respective position is returned. In our application, we disabled the GSP provider and used only data from the network provider. When a position was to be estimated, the procedure involved the following steps: 1) obtain data

---

[5] http://developer.android.com/guide/topics/location/obtaining-user-location.html

[6] http://developer.android.com/reference/android/location/LocationManager.html

generated by the network provider 2) determine location using the LocationManager API; 3) generate a position record to be stored on the mobile device and uploaded to the server when appropriate. Each record comprises the following information:

- Reading ID: A unique hash generated from the timestamp;

- Timestamp: A timestamp of when the reading was done;

- LocationID: A key identifying the test location where the reading was made;

- Network Provider info: This is the information about the Wi-Fi and 2G GSM land-scape at the moment of the reading as generated by the network provider;

- Real-time network position: The position as estimated on the device using information from the network provider and respective accuracy estimation.

5.3.2 Data collection

We have selected a total of 11 collection locations, as listed in Table 2. Considering that the nature of locations may affect the quality of the positioning process, we chose a diversified set of locations. In particular, we have selected 6 indoor locations and 5 outdoor locations. The outdoor locations were all located in residential areas with high density of Wi-Fi access points, and thus enough radio information for positioning techniques. To address the fact that different locations may have much stronger and much more frequent variation in their Wi-Fi radio landscape than others, we included 4 locations on campus, where we would expect the radio landscape to be more stable, and 7 on other locations across town, where a more dynamic landscape would be expected.

Indoor	Outdoor
Campus room, Campus cafeteria, Campus library, Café in the city, Cinema, Restaurant	Outside rural residence 1 and 2 Outside urban residence 1 and 2 Outside location on Campus

Table 2 - Data collection locations

Additionally, we have defined a data collection protocol that explicitly addressed the possible effect of movement patterns in the data collected. In our target scenario, we should be able to track the position of a person that may be moving in a city, and therefore may only spend a few seconds at any given location.

To explicitly address the possible effects of movement we defined a data collection protocol, in which the researcher would activate the application at some distance from the reference location. He would then approach the reference location and immediately, upon arrival, trigger

the data collection process. This would immediately generate the first record. The researcher would then wait at the reference location for the application to automatically generate the second record, two minutes later, and the third record, four minutes later. This was repeated twice at different times of the day.

5.3.3 Data processing

Once in the server, a Perl script was used to calculate locations. The script parsed each location record and generated a JSON request with the respective GSM and Wi-Fi information to be posted through HTTP to Google's geolocation service that re-turns a JSON reply containing the estimated position. The script then decodes the JSON response and stores the result as the delay-tolerant position for that record. We have made this estimation multiple times over the experience period to accommodate possible changes in accuracy through time.

To compare the accuracy of the position records, we have defined a ground truth by setting a reference position for each of the locations in our study using the coordinates obtained from a high-sensitivity GPS navigator (Garmin eTrex Legend HCx). For indoor locations, we used satellite imagery from Google Maps, always oriented towards north, and overlaid perfect squares over the images where one of the vertexes would overlay the point whose coordinates we wished to determine. Having done so, we then proceeded to take GPS coordinates of 2 of the vertexes (from the same edge) and then we used the Haversine formula (1) to calculate the distance between those 2 points relative to the Earth's surface.

$$haversin\left(\frac{d}{r}\right) = haversin\left(\phi_2 - \phi_1\right) \times \cos(\phi_1) \times \cos(\phi_2) \times haversin(\Psi_2 - \Psi_1) \quad (1)$$

In formula (1) d is distance, r is radius (the radius of the Earth in our particular case), finally $\phi1$ and $\phi2$ and $\Psi1$ and $\Psi2$ are respectively, the X and Y coordinates for the relevant points. We solve the haversine for d and obtain the distance. Afterwards, having performed the distance calculations, we chose a vertex from each square, along which we would be travelling the previously determined distance, along a fixed bearing according to the following formulas for latitude and longitude:

$$lat_{dest} = asin\left(sin(lat_{orig}) \times \cos\left(\frac{d}{r}\right) + \cos(lat_{orig}) \times \sin\left(\frac{d}{r}\right) \times \cos(\Theta)\right) \quad (2)$$

$$lon_{dest} = lon_{orig} + atan2\left(\sin(\Theta) \times \sin\left(\frac{d}{r}\right) \times \cos(lat_{orig}), \cos\left(\frac{d}{r}\right) - \sin(lat_{orig}) \times \right.$$
$$\left. \sin(lat_{dest})\right) \quad (3)$$

For formulas (2) and (3), d and r are again distance and radius (Earth's radius) respectively, Θ is the bearing (in radians, clockwise from north, i.e., North = 0, East = 90, South = 180 and

West = 270, given a conversion from degrees to radians), the rest of the variables are self-explanatory in name. In Formula (3) we use atan2(Y, X) which is a variant of the arctangent function which returns the arctangent of Y/X in the range $-\Pi$ to Π (mathematical PI).

5.4   Results for delay-tolerant positioning

Using our data collection application, we generated position records at the 11 locations during a three week period, with three weekly observations at each of the reference locations, two times a day. This resulted in a total 54 readings per location, or 594 readings in total. These were processed at the server at different moments, generating a total of 2816 delay-tolerant positions.

To assess the accuracy of the delay-tolerant positions, we have used two different types of reference data: the estimated error (reported) given by the Google location API at the server and at the device; and our own estimation of that error (real) based on the ground truth positioning data calculated for each of the locations. Having determined all the ground truth points, we then calculated for each location the distance from this reference position to each of the position records, including the delay-tolerant positions.
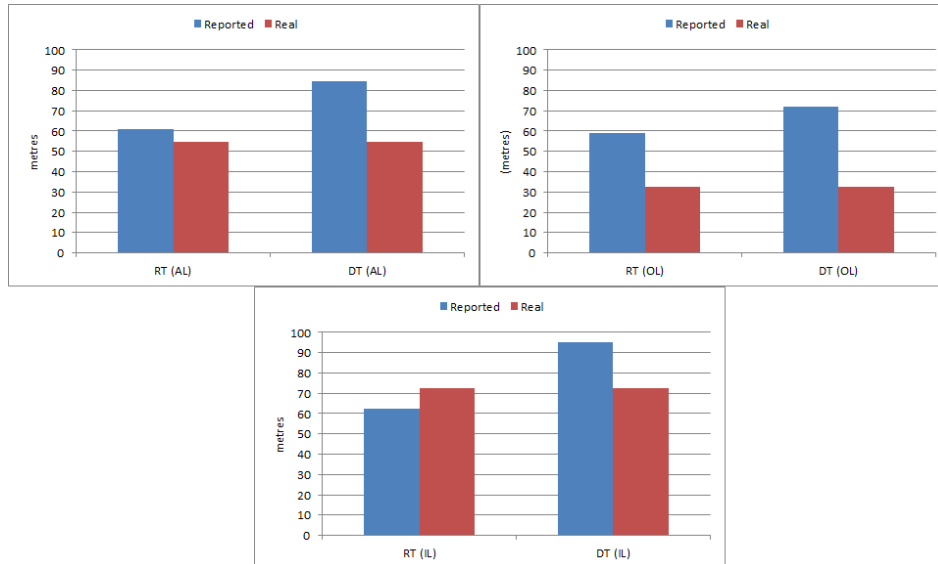
Figure 23 - Real-time (RT) vs delay-tolerant (DT) error results in reading 1 (left – all locations (AL), right – outdoor locations (OL) and bottom – indoor locations (IL))

The main conclusion from Figure 23 is that, when considering the error in relation to the ground truth, there are no observed differences between the location determined in real-time at the device and the location determined latter at the server. This is the most fundamental observation of this study in the sense that it backs up our initial hypotheses regarding the viability

of delay-tolerant positioning. It is also relevant to note that despite the differences in accuracy, the similarity between real-time and delay-tolerant locations existed for both indoor and outdoor environments. We can observe, however, that the error estimated by the Google Location API is shown to be lower in real-time when compared to the error reported by the delay-tolerant estimation process, for all cases but one. We have no explanation for this behaviour of the Google API, but since the effective accuracy is not affected, it should not be any sort of problem for location data logging.
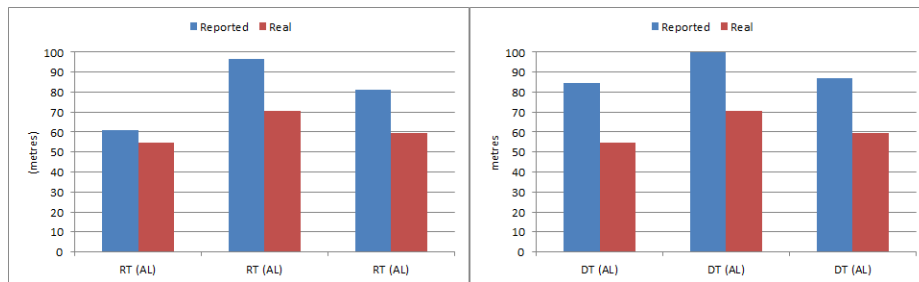

Figure 24 - Left – Real-time (RT) error results for all locations (AL) in readings 1, 2 and 3. Right – Delay-tolerant (DT) error results for all locations (AL) in readings 1, 2, and 3

Figure 24 reveals the errors obtained for all locations in all three readings, for the real-time and the delay-tolerant processes. For all scenarios the first reading is always the best, with the third reading coming a close second and the second reading always being the worst. In regard to the time when the location was estimated on the server, we have not observed any meaningful effect.

5.4.1   Cautionary considerations

A limitation of this study is our lack of knowledge about the internals of the Google's location API. There are no public details about how it uses the radio and Wi-Fi information to calculate position, and whatever the current approach might be, it may suddenly change without any prior announcement. Such changes could possibly affect the results obtained in this study and lead to potentially different conclusions. Also, the internals of specific devices and particularly their support for capturing radio and Wi-Fi information may also vary and lead to potentially different results in specific types of mobile phones. As such, one possible future direction of research would be to conduct the experiment with several different devices and analyse the results obtained to get a grasp on what sort of variability can be expected from different devices in terms of positioning error. Another line of research would be to understand the maximum validity of the network information stored by the mobile device and sent latter to the server. As network

information associated with a position evolves, observations made in the past will eventually become unsuitable for an adequate location determination. Understanding the timescale in which this effect may become relevant would help defining upload policies for location logging tools.

6 Conclusions

In this chapter we present the conclusions of our work in light of the objectives we defined for our platform and the process of validating it. Our initial objective was to establish a solid understanding of the evaluation processes based on experience sampling and to uncover the key abstractions that may be used for designing generic tools. In this respect we feel like we successfully gathered the necessary knowledge surrounding the critical aspects of the experience sampling technique and how to design a tool around it. We analysed several study techniques, we delved into experience sampling techniques and tools, and we were able to extract solid design guidelines to inform our choices. These guidelines moulded the content and length of questionnaires, they lead us to design for simplicity of interaction and interruption minimisation, as well as serving other requirements of mobile ubicomp settings and experience sampling studies.

The objective of devising and implementing a system architecture for the creation of a web-based tool that supports general user experience sampling studies, was also accomplished. In realising the first objective we established a solid basis upon which we could forge a vision of what a system of this type should be like, from an architectural standpoint. Implementing it was mostly a matter of choosing what where the critical features that a prototype should have in order to be representative of our vision for the system and, overcoming the emerging technical problems.

Another important objective in validating our work was that of evaluating the tool in regards to its ability to address the needs of researchers wishing to run user studies based on experience sampling. Naturally, no one tool will cover every single scenario. However, we feel our tool is flexible enough to serve a wide range of needs in conducting user studies. The tool is easy to use and allows researchers to do their work, and inform participants of what are the study's objectives and requirements. In spite of this, some processes are still rather crude. For instance, the creation of the study specification file would benefit from us offering an interface to support this on the web site, rather than it being a manual process where you edit an XML file. Data analysis is also something that could be improved by automating additional data processing functionality, and featuring ways for researchers to consult the result of that processing on the web site's interface.

The last objective was tied to our aim of evaluating the platform for mobile data collection in regard to its ability to collect samples and integrate into everyday practices and devices. Our

findings here are two-fold. On the one hand, results point to the tool being an effective means of collecting samples and it seems to be reasonably effective in integrating itself in day-to-day activities and peoples' personal devices. On the other hand, our findings can be said to be of somewhat limited nature, given that the niche academic community we targeted for our particular study did not adhere to it in sufficiently expressive numbers. The reason we chose a niche academic community was to be able to present richer results in regards to what where the research questions of the ringer mode study. When designing the study it was clear that this would present the most valuable dataset, as it made sense to analyse people that where generally in the same physical spaces within the same time frames. This would however mean that we would face the risk of not having a truly representative sample size and that is exactly what happened. This translated into us being unable to answer the research questions of the ringer mode study.

In what concerns the delay-tolerant positioning process, the main conclusion of this work is that it is a perfectly viable alternative approach for our usage scenario, especially considering that in terms of real error (in relation to ground truth) both methods achieve the same performance. This is a very important contribution to inform the design of any sort of location-based logging tools in which, as in our case, the position information is not needed at the moment of logging. The second conclusion is that time spent at the target location does not improve the accuracy of the positioning process. Moreover, there seems to be no gain whatsoever in staying more than 2 minutes at a given location. Together these two observations suggest that this process will perform well for the generation of location traces in high mobility scenarios. The only observed difference is in the estimated error that is associated with the location estimations, but this should not have any impact for most application domains.

7   Future Work

As future work there are several interesting things to discuss and ways in which the platform can evolve going forward. As an example, it would be interesting to implement the invitation system presented in our vision of the architecture. Also, in terms of researcher functionality, it would be valuable to provide continuous access to study data through the interface on the website, as opposed to this having to be done manually. This would pose interesting challenges like ways to present questionnaire data in an easy to consult fashion and perhaps means for presenting location data visualisation, in a way that would be informative and tied to other study data. Researchers would also benefit from an interface that would allow them to create study specification files, in contrast to having to do this manually by editing an XML file.

Other interesting challenges lay in the possibility of including additional methods of input for answering questionnaires, like audio, pictures and video. This would not only present a challenge in terms of how to present this data to researchers, but also on how to manage all of it on the mobile side, as media files have the potential to increase data consumption by a huge margin.

Supporting interaction with external systems is also something we would like to visit. We see this as a potential way of evolving the tool towards meeting its objective of being truly general purpose. A colleague and researcher proposed that we extend the tool in a way that would allow him to present a questionnaire to someone who compiled and experimented software in his research lab, so that he could gather input on that process, and on the software itself. This would be an interesting technical undertaking but also a complex one as it would likely involve creating additional software, customized to that particular case.

A case can also be made for integrating additional sensing functionality into the mobile client, so as to allow it, for instance, to interact with public display networks and trigger questionnaires related to the surroundings of a detected public display.

Finally, an important aspect in expanding the tool's reach would be to develop additional clients for other mobile operating systems, covering platforms such as iOS and Windows Phone, which would allow the tool to serve a larger potential study participant base.

8 References

[1]   J. M. Hektner, J. A. Schmidt, and M. Csikszentmihalyi, *Experience sampling method: Measuring the quality of everyday life*. Sage Publications, Inc, 2006, p. 368.

[2]   S. Consolvo and M. Walker, "Using the experience sampling method to evaluate ubicomp applications," *IEEE Pervasive Computing*, vol. 2, no. 2, pp. 24–31, Apr. 2003.

[3]   I. M. Welpe, M. Reiss, P. Seiffert, C. Flake, R. Lauterbach, A. Tumasjan, M. Strobel, S. Hassold, and K. Freundorfer, "Mobile Experience Sampling (MobXamp)," *Technische Universität München*, 2010. [Online]. Available: http://www.strategie.wi.tum.de/index.php?id=79&L=1. [Accessed: 18-Sep-2010].

[4]   M. Killingworth and V. Pitiyanuvath, "Track your happiness," *Harvard University*, 2009. [Online]. Available: http://www.trackyourhappiness.org/. [Accessed: 18-Sep-2010].

[5]   D. T. Campbell, "Relabeling internal and external validity for applied social scientists," *New Directions for Program Evaluation*, vol. 1986, no. 31, pp. 67–77, Jan. 1986.

[6]   S. Schwartz, "The fallacy of the ecological fallacy: the potential misuse of a concept and the consequences.," *American journal of public health*, vol. 84, no. 5, pp. 819–24, May 1994.

[7]   B. J. Calder, L. W. Phillips, and A. M. Tybout, "The Concept of External Validity," *Journal of Consumer Research*, vol. 9, no. 3, pp. 240–244, 2012.

[8]   M. A. Schmuckler, "What Is Ecological Validity? A Dimensional Analysis," *Infancy*, vol. 2, no. 4, pp. 419–436, Oct. 2001.

[9]   K. L. Kraemer, *The information systems research challenge (vol. III): survey research methods*. Boulder, CO, USA: Harvard University Graduate School of Business Administration, 1991, p. 395.

[10]   L. Palen and M. Salzman, "Voice-mail diary studies for naturalistic data capture under mobile conditions," in *Proceedings of the 2002 ACM conference on Computer supported cooperative work - CSCW '02*, 2002, pp. 87–95.

[11]   J. Chipchase, P. Persson, P. Piippo, M. Aarras, and T. Yamamoto, "Mobile essentials: field study and concepting," in *Designing For User Experiences; Vol. 135*, 2005.

[12]   G. Jacucci and M. Isomursu, "Facilitated and performed 'happenings' as resources in ubiquitous computing design," *Digital Creativity*, vol. 15, no. 4, pp. 223–231, Dec. 2004.

[13]   B. Gaver, T. Dunne, and E. Pacenti, "Design: Cultural probes," *interactions*, vol. 6, no. 1, pp. 21–29, 1999.

[14]   E. Ruspini, *Longitudinal Research and the Analysis of Social Change*, Quality & . Kluwer Academic Publishers, 1999, pp. 219–227.

[15]   K. Cheverst, N. Davies, K. Mitchell, A. Friday, and C. Efstratiou, "Developing a context-aware electronic tourist guide," in *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '00*, 2000, pp. 17–24.

[16]   S. Gabrielli, V. Mirabella, S. Kimani, and T. Catarci, *Supporting cognitive walkthrough with video data*. Salzburg, Austria: ACM Press, 2005, p. 77.

[17]   V. Roto, A. Oulasvirta, T. Haikarainen, J. Kuorelahti, H. Lehmuskallio, and T. Nyyssönen, "Examining mobile phone use in the wild with quasi-experimentation," Helsinky, Finland, 2004.

[18]   R. K. Yin, *Case Study Research: Design and Methods*, 3rd ed. Beverly Hills, CA, USA: Sage Publications, Inc, 2002, p. 200.

[19]   M. C. González, C. A. Hidalgo, and A.-L. Barabási, "Understanding individual human mobility patterns.," *Nature*, vol. 453, no. 7196, pp. 779–82, Jun. 2008.

[20]   I. Eriksson, "Simulations for user training," Åbo akademi Department of Computer Science, Åbo, 1990.

[21]   M. S. Andreasen, H. V. Nielsen, S. O. Schrøder, and J. Stage, "What happened to remote usability testing?," in *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '07*, 2007, pp. 1405–1414.

[22]   M. H. Vastenburg and N. R. Herrera, "Adaptive Experience Sampling : Addressing the Dynamic Nature of In-Situ User Studies," in *Ambient Intelligence and Future Trends - International Symposium on Ambient Intelligence (ISAmI 2010)*, J. M. Corchado, J. C. Augusto, P. Novais, and C. Analide, Eds. Springer-Verlag Berlin Heidelberg, 2010, pp. 197–200.

[23]   M. Cherubini and N. Oliver, "A Refined Experience Sampling Method to Capture Mobile User Experience," in *CHI 2009 Workshop: Mobile User Experience Research: Challenges, Methods & Tools*, 2009, vol. 2009, p. 12.

[24]   G. Iachello, K. N. Truong, G. D. Abowd, G. R. Hayes, and M. Stevens, "Prototyping and sampling experience to evaluate ubiquitous computing privacy in the real world," in *Proceedings of the SIGCHI conference on Human Factors in computing systems - CHI '06*, 2006, pp. 1009–1018.

[25]   J. E. Fischer, "Experience-Sampling Tools : a Critical Review," *Journal of Youth and Adolescence*, pp. 1–3, 2009.

[26]   S. S. Intille, J. Rondoni, C. Kukla, I. Ancona, and L. Bao, "A context-aware experience sampling tool," in *CHI '03 extended abstracts on Human factors in computing systems - CHI '03*, 2003, pp. 972–973.

[27]   A. Kittur, E. H. Chi, and B. Suh, "Crowdsourcing user studies with Mechanical Turk," *Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems - CHI '08*, p. 453, 2008.

[28]   B. Fogg, J. Marshall, T. Kameda, J. Solomon, A. Rangnekar, J. Boyd, and B. Brown, "Web credibility research," in *CHI '01 extended abstracts on Human factors in computing systems - CHI '01*, 2001, p. 295.

[29]   B. Paulson, A. Wolin, J. Johnston, and T. Hammond, "SOUSA : Sketch-based Online User Study Applet," in *SBM'08 Proceedings of the Fifth Eurographics conference on Sketch-Based Interfaces and Modeling*, 2008, pp. 81–88.

[30] G. E. Fitzgerald, T. Hall, and L. P. Semrau, "User Studies : Evaluating the Use of EPSS Tools for Self-Management by Theoretical Base for the Use of EPSS with Children The Two Components of KidTools," in *Proceedings of Ed-Media 2000 World Conference on Educational Multimedia and Hypermedia*, 2000, pp. 1999–2001.

[31] H. Yoshida, S. Ito, and N. Kawaguchi, "Evaluation of pre-acquisition methods for position estimation system using wireless LAN," in *Third International Conference on Mobile Computing and Ubiquitous Networking (ICMU 2006)*, 2006, pp. 148–155.

[32] P. A. Zandbergen, "Accuracy of iPhone locations: A comparison of assisted GPS, WiFi and cellular positioning," *Transactions in GIS*, vol. 13, pp. 5–25, Jun. 2009.

[33] B. N. Schilit, A. LaMarca, G. Borriello, W. G. Griswold, D. McDonald, E. Lazowska, A. Balachandran, J. Hong, and V. Iverson, "Challenge: Ubiquitous Location-Aware Computing and the 'Place Lab' Initiative," in *Proceedings of the 1st ACM international workshop on Wireless mobile applications and services on WLAN hotspots - WMASH '03*, 2003, p. 29.

[34] M. Paciga and H. Lutfiyya, "Herecast: An open infrastructure for location-based services using WiFi," in *WiMob'2005, IEEE International Conference on Wireless And Mobile Computing, Networking And Communications, 2005.*, 2005, vol. 4, pp. 21–28.

[35] S. B. Eisenman, E. Miluzzo, N. D. Lane, R. A. Peterson, G.-S. Ahn, and A. T. Campbell, "The BikeNet mobile sensing system for cyclist experience mapping," in *Proceedings of the 5th international conference on Embedded networked sensor systems - SenSys '07*, 2007, p. 87.

[36] J. Fogarty, S. E. Hudson, C. G. Atkeson, D. Avrahami, J. Forlizzi, S. Kiesler, J. C. Lee, and J. Yang, "Predicting human interruptibility with sensors," *ACM Transactions on Computer-Human Interaction*, vol. 12, no. 1, pp. 119–146, Mar. 2005.

[37] T. Reichenbacher, "Geographic relevance in mobile services," in *Proceedings of the 2nd International Workshop on Location and the Web - LOCWEB '09*, 2009, pp. 1–4.

[38]   M. B. Kjærgaard, J. Langdal, T. Godsk, and T. Toftkjær, "EnTracked: energy-efficient robust position tracking for mobile devices," in *Proceedings of the 7th international conference on Mobile systems, applications, and services - Mobisys '09*, 2009, p. 221.

[39]   B. Pinkas and T. Sander, "Securing passwords against dictionary attacks," *Proceedings of the 9th ACM conference on Computer and communications security - CCS '02*, p. 161, 2002.

[40]   N. Chou, R. Ledesma, Y. Teraguchi, J. C. Mitchell, and S. Ca, "Client-side defense against web-based identity theft," in *Proceedings of the Network and Distributed System Security Symposium*, 2004.

9   Annexes

In this section we present additional information that readers can consult in order to get further details into some aspects of our work.

9.1   Annex A

In this annex we feature screenshots of the mobile client application.
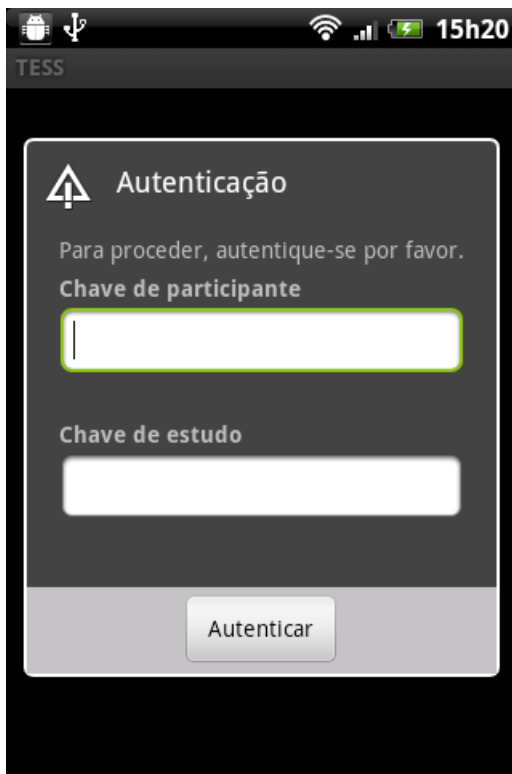

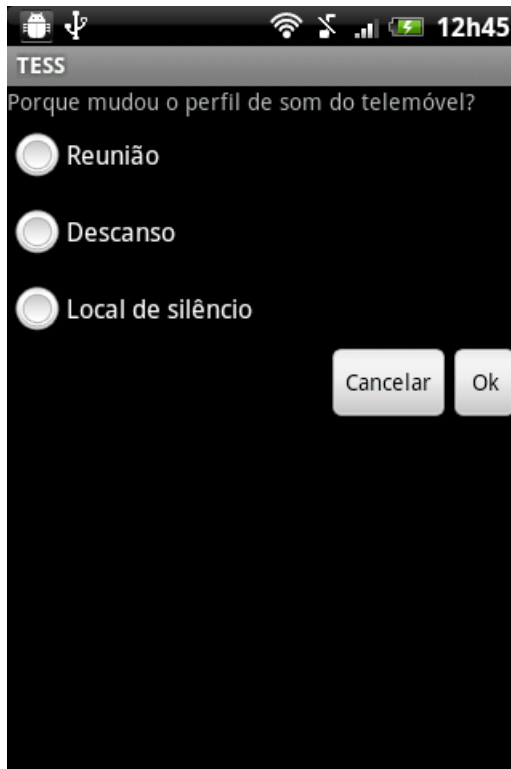Figure 25 – Authentication prompt

Figure 26 – Preferences screen

Figure 27 – Questionnaire dialog

Figure 28 – Synchronisation progress pie chart

9.2 Annex B

The following annex presents a transcription of the help information section of the web site.

9.2.1 Info

Este site destina-se ao registo de participantes para realização de estudos de experiência de utilizador, recorrendo ao método de *experience sampling*.

Ética e privacidade

A participação nos estudos não está directamente ligada à identidade de registo na plataforma TESS. Os seus dados não serão fornecidos a terceiros e quaisquer dados recolhidos durante o estudo serão permanentemente apagados, assim que o estudo for apagado da plataforma (o que sucederá quando os dados forem analisados e conclusões forem estabelecidas). A aplicação TESS recolhe dados das redes Wi-Fi e GSM circundantes na altura do preenchimento dos questionários. Esses dados ficam armazenados no dispositivo e permitiram, mais tarde, quando passados para o servidor, determinar a posição do dispositivo, aquando do preenchimento do questionário. A localização do dispositivo não é monitorizada em tempo real, em altura nenhuma do estudo e destina-se apenas a permitir melhor compreender o contexto do evento sobre o qual foi inquirido.

Para qualquer esclarecimento adicional, é favor contactar: acoelho@dsi.uminho.pt.

9.2.2 TESS Android

Nesta secção será dada uma breve explicação do funcionamento da ferramenta móvel de apoio à realização de estudos na plataforma TESS. Esperamos que sirva para esclarecer quaisquer dúvidas que possam surgir, sobre o funcionamento da ferramenta.

Autenticação

Após criar uma conta no sítio web e efectuar a inscrição para um dos estudos disponíveis, ser-lhe-á enviado um e-mail com os dados necessários para participar através da aplicação móvel (a chave de participação e a chave de estudo). Assim que instalada, a aplicação, quando iniciada, vai ser-lhe pedido que introduza os seus dados de autenticação relativos ao estudo em que está inscrito. Para os introduzir e validar, será necessário algum tipo de ligação à internet.

Opções

Uma vez validados os dados, não será necessário voltar a introduzi-los e a aplicação prosseguirá pedindo-lhe que configure as opções de conectividade. As opções disponíveis são Wi-Fi e 3G/Wi-Fi, sendo que a primeira faz com que apenas sejam sincronizados dados quando estiver disponível uma ligação Wi-Fi e a segunda, faz com que sejam sincronizados dados estando activa uma ligação 3G ou Wi-Fi.

9.2.3   Funcionamento

Alertas

Sempre que ocorrer alguma acção relevante ao estudo em que está inscrito, a aplicação TESS gerará uma notificação no seu telemóvel. Essa notificação ficará visível na barra de sistema e será possível, a partir do menu de notificações, pressionar a aplicação e preencher o questionário relativo ao evento que gerou a notificação.

Questionários

Após pressionar uma notificação, será então disponibilizado um questionário para preencher, consistindo de perguntas de escolha múltipla. O intuito das questões será o de determinar qual o motivo pelo qual ocorreu determinado evento e qual a sua resposta a esse mesmo evento.

Sincronização

Para que lhe seja possível acompanhar o estado da sua participação no estudo e averiguar se tudo está a correr bem, a aplicação TESS disponibiliza um gráfico que lhe mostra qual a porção dos seus dados que se encontra já sincronizada, bem como, a porção não sincronizada.

9.3 Annex C

The following text presents the contents of an XML file for study specification.

```xml
<?xml version="1.0" encoding="UTF-8"?>
    <Study ID="18bb5fc6">
        <Dates>
            <Start>11/02/2012 00:00:00</Start>
            <End>29/02/2012 00:00:00</End>
        </Dates>
        <Notifications>
            <Number>0</Number>
            <Type>Random</Type>
        </Notifications>
        <Questions>
            <Question ID="1">
                Porque    mudou    o    perfil    de    som    do    telemóvel?
                <QType>Ringer</QType>
                <Option ID="1">Reunião</Option>
                <Option ID="2">Descanso</Option>
                <Option ID="3">Local de silêncio</Option>
            </Question>
            <Question ID="2">
                Este é um local onde gostaria sempre de ter o telemóvel em silêncio?
                <QType>Ringer</QType>
                <Option ID="1">Sim</Option>
                <Option ID="2">Não</Option>
            </Question>
        </Questions>
    </Study>
```

9.4 Annex D

The following lists of questions belong to the interview scripts that we elaborated for guiding our follow-up interviews to active and inactive study participants.

9.4.1 Active participants

1. Considera que a informação disponibilizada sobre o estudo era suficiente para permitir uma participação informada? Se não, porquê?

2. Achou esclarecedora a informação fornecida sobre os dados que a aplicação recolhe? Se não, porquê?

3. Ficou suficientemente esclarecido relativamente à política de privacidade da plataforma? Se não, porquê?

4. Sentiu dificuldade em entender os requisitos de participação no estudo? Se sim, porquê?

5. Sentiu dificuldade em perceber qual o momento em que seria calculada a localização do dispositivo e a forma como isso seria feito? Se sim, porquê? Sente que isso é um factor determinante para ditar a sua participação ou não participação no estudo?

6. Teve dificuldades em navegar no sítio web da plataforma? Se sim, porquê?

7. Teve dificuldades no acto de registo na plataforma? Se sim, porquê?

8. Teve dificuldades no acto de inscrição no estudo? Se sim, porquê?

9. Teve dificuldades em configurar ou navegar na aplicação? Se sim, porquê?

10. Teve dificuldades em perceber qual o estado da sua participação no estudo, durante o período em que este decorreu? Se sim, porquê?

11. Considera que a aplicação o interrompeu com demasiada frequência, interferindo nas suas tarefas normais? Se sim, porquê?

12. Considera que as questões colocadas através da aplicação eram suficientemente claras? Se não, porquê?

13. Teve dificuldade em interagir com os questionários da aplicação? Se sim, porquê?

14. A participação no estudo afectou negativamente a autonomia da bateria do telemóvel? Se sim, até que ponto.

15. Sente que a participação no estudo afectou negativamente a sua privacidade? Se sim, porquê?

16. Poderia fornecer uma breve descrição da sua experiência de participação, apontando os aspectos positivos ou negativos que achou relevantes?

17. Apresente sugestões/alterações relativas a qualquer aspecto da plataforma (sítio web e aplicação móvel) que considere serem conducentes a uma melhoria da experiência.

9.4.2 Inactive participants

1. Considera que a informação disponibilizada sobre o estudo era suficiente para permitir uma participação informada? Se não, porquê?

2. Achou esclarecedora a informação fornecida sobre os dados que a aplicação recolhe? Se não, porquê?

3. Ficou suficientemente esclarecido relativamente à política de privacidade da plataforma? Se não, porquê?

4. Sentiu dificuldade em entender os requisitos de participação no estudo? Se sim, porquê?

5. Sentiu dificuldade em perceber qual o momento em que seria calculada a localização do dispositivo e a forma como isso seria feito? Se sim, porquê? Sente que isso é um factor determinante para ditar a sua participação ou não participação no estudo?

6. Teve dificuldades em navegar no sítio web da plataforma? Se sim, porquê?

7. Teve dificuldades no acto de registo na plataforma? Se sim, porquê?

8. Teve dificuldades no acto de inscrição no estudo? Se sim, porquê?

9. Teve dificuldades em configurar ou navegar na aplicação? Se sim, porquê?

10. Explique qual a razão ou razões pela qual não gerou dados durante a sua participação no estudo. O que poderia contribuir para que essa situação fosse diferente?

11. Sente que a participação no estudo afectou negativamente a sua privacidade? Se sim, porquê?

12. Apresente sugestões/alterações relativas a qualquer aspecto da plataforma (sítio web e aplicação móvel) que considere serem conducentes a uma melhoria da experiência.