Problems of space and time: learning from the experience of studying ubicomp *use* in the wild

Scott Sherwood, Julie Maitland and Matthew Chalmers[1]

Department of Computing Science, University of Glasgow,
Glasgow G12 8QQ, UK
{sherwood, jules, matthew }@dcs.gla.ac.uk

Abstract. In this paper we describe, how the evaluation of ubiquitous computing systems can be problematic, how we augmented existing evaluation techniques to aid in our analysis, and how we created new techniques for evaluation of such systems within everyday settings. We will discuss four systems, briefly highlighting the problems faced and how the experiences from each changed our approach to evaluating the following systems.

1 Introduction

When working with ubiquitous computing (ubicomp) systems, challenges and rewards arise from moving from the relative safety of the usability lab into the uncontrolled environment of everyday life. For example, unpredicted contexts of use and environmental features such as intermittent network connectivity may challenge traditional evaluation methods, and yet we gain the mobility, contextuality and appropriation that lets users take full advantage of new mobile devices. Many of these challenges have already been faced by researchers studying the *use* (rather than *usability*) of ubicomp technologies in the wild. In this paper, we reflect on studies of two mobile games (Treasure [1] and Feeding Yoshi [2]) and two everyday awareness applications (Shakra [4] and FriendPhone). We will discuss the strategies employed by the analysts to discover the users' reactions towards and their experiences of the technology. The systems are presented chronologically, as the challenges faced in one study often influenced the design and study of the next system. In particular, we will consider the temporal and geographic scale of each study as contributing factors to the complexity of gathering evaluation data.

2 Games and Awareness Applications

A key issue to emerge from early ubicomp studies has been the impact of variation or uncertainty with regard to positioning accuracy and network connectivity. Consequently a number of strategies have been proposed for dealing with this issue: one being to utilise it as a resource for design. Seamful design is one such approach that exploits "the limits, gaps and variability of technologies we often assume to be uniform and seamless" [3].

Treasure is a seamful game that exploits the variation in WiFi network coverage. Each player uses a handheld PDA equipped with GPS and 802.11. Players collect

[1] Acknowledgement is given to Louise Barkhuus, Marek Bell, Barry Brown, John Ferguson, Malcolm Hall, Alistair Morrison and Paul Tennent for their work on the systems and evaluations described in this paper.

virtual coins that are placed in areas not covered by the game's wireless network. In order to gain points for collecting the treasure they have to take the coins back to an area covered by the network to upload them to game's server. Each game involved two teams of two people, and was played outdoors. Although the game was run in several cities, in our main study it was played within a game 'arena' of ~7000m^2 on the edge of the University of Glasgow. Despite the constrained area, it was not always possible to observe all of the players at any one time. In order to observe the behaviour of players, two members of our research team were needed to follow them around with video cameras. Immediately after each game, players were interviewed and asked to reflect upon the experience and how they played the game. Each player was allowed to play multiple times, revealing how players' strategies developed over time.

Although positioning and coverage were no longer problematic from a user-experience perspective, since they had become resources within the game, they did prove to be very problematic when it came to collecting data for evaluation. Unlike lab based experimentation, the log data gathered was often unreliable in the sense that the recorded position may not represent the actual location of the player when the log was recorded. We believe this kind of issue to be one that will persist for some time in ubicomp evaluation. Inaccurate positioning and intermittent connectivity meant that it was possible for there to be several versions of the game state at any one time—one for each player and one for the server. Synchronising the multiple logs generated by each player's use of the system was a labour intensive task; further complicated by the need to explore circumstances of play and interaction by marrying-up events captured by multiple video cameras with the multiple states of the system at that time. Such challenges inspired the design of Replayer [5], an analytic tool that integrates and synchronises log data from multiple sources to allow quantative, qualitative and exploratory data analysis.

Feeding Yoshi is a second seamful game that runs on PDAs, and is based on the distribution of secure and unsecure WiFi Access Points (APs), wherever they are found. Teams of players must find plantations (unsecure APs), in which they grow fruit. They gain points by feeding this fruit to Yoshis (secure APs), and the team with the most points wins. Unlike Treasure, Feeding Yoshi is designed to be played long-term and fit into the routine of everyday life. Rather than being confined to a small game arena; our main study lasted one week and involved four teams of players based in three UK cities.

Due to the situated nature of the game and our interest in discovering how players responded to the contingencies of the technology and of the everyday world in which they were playing, the approach that we took to evaluate Treasure was infeasible. Rather than 'staging' a game within a semi-controlled environment, game play was intentionally conditional on players' existing circumstances of work, leisure and home life. Players were free to play the game whenever and wherever they liked, and in fact they ranged across the entire country. We were only occasionally able to observe players, e.g. video their activity, and could not guarantee that the game would be played at all.

Game state was simpler than in Treasure, and we had too few types of data to make the use of Replayer worthwhile. We could keep in touch with players via mo-

bile phones and email but this could be seen as intrusive and bothersome to players. We relied on post-trial analysis based on simple system logs and interviews in which players reflected on their play. Analysis was challenged by the lapsed time between game play and the interviews themselves. The led the team to utilise usage diaries in future evaluations, and to consider ways to make diaries richer and more relevant to players.

Shakra is a mobile-phone based awareness application that uses patterns of fluctuation in GSM cell signal strength to determine whether or not a user is stationary, walking or travelling in a car. Real-time feedback of the currently detected mode of motion is presented via an animation on the home screen of the application. A user can then view their daily activity level (the accumulated time spent walking during the day) in isolation, within the context of previous days' levels, or in comparison with the accumulated activity totals of their friends also using Shakra. Shakra was piloted with three groups of friends over the course of one week, to detail its use and to determine whether it increased users' awareness of their activity level. The focus was on the users' experiences with both the activity tracking and the sharing feature. As with Feeding Yoshi, trialling the system within a controlled environment would not have allowed us to fully explore the experience and consequences of using the system. To overcome the issue of delayed reflection that was found during the Yoshi trial, usage diaries were provided. Although the usage diaries provided valuable insight, serving as a reminder for the user and a prompt for the researcher during interviews, it was limited in its usefulness due to the static and predetermined nature of the questions within it. To address this problem we developed an application for gathering data and making web–based interview forms that we called *FlexiFill*, which we piloted during the trial of the next described system, *FriendPhone.*

FriendPhone is a mobile phone application that displays context information about friends and contacts. It currently supports three main items of information: location detection by the phone itself and sharing of this, measuring of how long the person has been at that location or when they left a known location and finally, sharing of the phone profile. All three pieces of information are shown in the contact list screen. Like many other ubiquitous computing applications, FriendPhone was difficult to study in situ throughout the users' daily lives. In our studies of Shakra and Feeding Yoshi we relied on using a combination of fill-in-diaries and interviews which as we have seen was problematic. The challenge faced was how could we gain insight into the participants' actual use and perception of the application, without intruding or making constant inquiries that would detract from the user experience. Also studying ubicomp user experiences lasting much longer than is usually the case, how could we make sure interesting outcomes were not lost or forgotten before the analysts had time to engage with the users?

For the purpose of daily inquiries, we developed a dynamic web-based diary tool, Flexifill, which adapts its questions to that particular day's phone communication. Since all information about location and communication is logged and sent to a central server via GPRS, it was simple to use this information in the questions. From previous experiences we were aware that participants are reluctant to write thorough diaries by hand, in particular if they have to do it daily and the questions are repetitive. The diary was therefore designed to be flexible; participants could fill it in at a

time and place that suited them. In the diary, the logged information was used to remind the participants whom they communicated with; making it easier for them to recall their motivation and actions around the communication. In each entry we asked about a single, randomly selected phone call as well as a text message sent that day. Apart from questions asking about communication in relation to the FriendPhone application, we also inquired into the participants' level of comfort with having their location tracked throughout the day. We also enquired, if they used it, why they decided to use the manual location mode rather than continuing to use automatic tracking. One of the main advantages we found with an online diary was that both midway and final interviews could be more tailored to the particular participant.

3 Discussion

Several issues combined to shape our methods and tools for evaluation of ubicomp systems used at a far larger geographic and temporal scale than traditional lab studies. Technical issues included inaccurate positioning, intermittent connectivity and multiple versions of system state, and social issues included the inconvenience and intrusiveness of interviews and the relative irrelevance or genericity of fill-in diaries. In order to alleviate the technical issues, we created the Replayer tool that lets evaluators combine system logs and multiple streams of audio/video recordings, and uses information visualisation techniques to support richer interaction and analysis of the resulting heterogeneous data sets. With regard to the social issues, we developed the FlexiFill system that uses sampled data from the day's activity to enrich the forms used for fill-in diaries, making them more relevant and acceptable to trial participants.

We continue to develop these tools, along with new ubicomp user trials and experiences. In effect, we see strong benefits for evaluators in taking advantage of the same technologies that we are developing for users, in terms of using wireless networks and distributed sensors and camera as tools for maintaining awareness, finding relevant information, and in making evaluation more of a synchronous engaged experience despite the vagaries of geographic and temporal scale, shifting context, and the work of fitting ubicomp evaluation into the routine of our own everyday lives.

References

1. Barkhuus, L., Chalmers, M., Tennent, P., Hall, M., Bell, M., Sherwood, S. and Brown, B., Picking Pockets on the Lawn: The Development of Tactics and Strategies in a Mobile Game. in *Ubicomp*, (Tokyo, 2005), 358-374.
2. Bell, M., Chalmers, M., Barkhuus, L., Hall, M., Sherwood, S., Tennent, P., Brown, B., Rowland, D., Benford, S. and Hampshire, A., Interweaving Mobile Games with Everyday Life. in *CHI*, (Montreal, 2006).
3. Chalmers, M., Bell, M., Brown, B., Hall, M., Sherwood, S. and Tennent, P., Gaming on the Edge: Using Seams in Pervasive Games. in *PerGames 2nd International Workshop on Pervasive Gaming Applications*, (2005).
4. Maitland, J., Sherwood, S., Barkhuus, L., Anderson, I., Hall, M., Brown, B., Chalmers, M. and Muller, H., Increasing the Awareness of Daily Activity Levels with Pervasive Computing. in *Pervasive Computing Technologies for Healthcare*, (Innsbruck, Austria, 2006), IEEE.
5. Morrison, A., Tennent, P. and Chalmers, M., Coordinated Visualisation of Video and System Log Data. in *International Conference on Coordinated and Multiple Views in Exploratory Visualization (CMV)*, (2006), 91-102.