Chapter 16

SGVis:
Analysis of Data from Mass Participation Ubicomp Trials

Alistair Morrison
University of Glasgow, UK

Matthew Chalmers
University of Glasgow, UK

ABSTRACT

The recent rise in popularity of ‘app store’ markets on a number of different mobile platforms has provided a means for researchers to run worldwide trials of ubiquitous computing (ubicomp) applications with very large numbers of users. This opportunity raises challenges, however, as more traditional methods of running trials and gathering data for analysis might be infeasible or fail to scale up to a large, globally-spread user base. SGVis is a data analysis tool designed to aid ubicomp researchers in conducting trials in this manner. This paper discusses the difficulties involved in running large scale trials, explaining how these led to recommendations on what data researchers should log, and to design choices made in SGVis. The authors outline several methods of use and why they help with challenges raised by large scale research. A means of categorising users is also described that could aid in data analysis and management of a trial with very large numbers of participants. SGVis has been used in evaluating several mass-participation trials, involving tens of thousands of users, and several use cases are described that demonstrate its utility.

INTRODUCTION

The rise of ‘app store’ markets is a relatively recent phenomenon. The Apple App Store launched in July 2008 and saw its 10 billionth download in January 2011 (Reuters, 2011). Smartphone usage has also seen a sharp rise in usage in recent years, with market research firm IDC suggesting that 15.4% of the mobile phone market consisted of smartphones at the beginning of 2010 (Llamas, 2010), and several different mobile platforms now offer ‘app store’ software distribution mechanisms.
The combination of this growing potential user base and popular online software repositories provides a relatively simple way to recruit users for worldwide trials of ubiquitous computing (ubicomp) applications and several researchers are beginning to use this ‘mass participation’ approach (McMillan et al., 2010; Cramer et al., 2010). Distributing ubicomp trial applications in this way provides great opportunities for researchers in terms of potentially reaching a very large number of trial participants. For example, Hungry Yoshi, a game discussed below, has had over 40,000 users in the 12 months since release.

The potential advantages of deploying trial software to a wide audience are numerous. These include getting a larger sample size to provide more certainty to quantitative analyses and, additionally, a global release of software provides the opportunity to reach users from vastly different geographic locations so as to help reduce cultural biases stemming from recruitment of only locally-based participants.

The benefits to be gained from this style of deployment do come with some potential drawbacks; however, as running a trial with a large number of users over a vast geographical area can raise significant challenges. Compared to more traditional local deployments of software, researchers can be further removed from the trial, unable to meet participants and perhaps less able to closely observe the use of the software under examination. Additionally, having such a large user base could lead to an overwhelming amount of data being generated and researchers might not have the resources to study every user in detail.

We have designed several tools to aid researchers in conducting such trials. The previously published SGLog framework (Hall et al., 2009) provides a simple means for developers to instrument their mobile application code and stream log data back to researchers’ database. This paper presents SGVis - a complementary desktop analysis tool designed to allow evaluators to study this data. Several modes are available in SGVis, allowing data gathered from users’ devices to be processed in various ways. In one mode an individual’s use of trial software can be studied in detail, while in another overall trends for an application can be analysed, such as download patterns over time or the average number of times participants perform certain actions. A ‘live’ mode allows analysts to watch software use ‘as it happens’, with a collection of maps and summary statistics showing researchers which of their trial applications are being used anywhere in the world in the last few minutes. The final mode in SGVis allows analysts to categorise the many thousands of users they might have, using multidimensional data analysis, clustering and visualisation techniques. In this way, analysts might be able to more successfully manage categories of participants rather than attempting to study each of the thousands of logged users in detail. SGVis has been used in the evaluation of several mobile applications, handling data sets of tens of thousands of users spread worldwide.

The following section surveys related work in the area of mass participation research applications, and of analysis tools that can be used to study data captured from such applications. A description is provided of the Contextual Software project (UK EPSRC EP/F035586/1), and how SGVis forms part of this research. This is followed by a discussion of the challenges arising from conducting research in this manner and recommendations for the types of data other researchers working in this area might want to capture. Thereafter, the SGVis tools for analysing large scale ubiquitous computing trials are introduced, including a rationale for design and illustrations of their utility with specific use cases, before finally a discussion is presented of the use of SGVis and ongoing work in this area.

**RELATED WORK**

Researchers in ubiquitous computing have recently begun to take advantage of ‘mass participation’