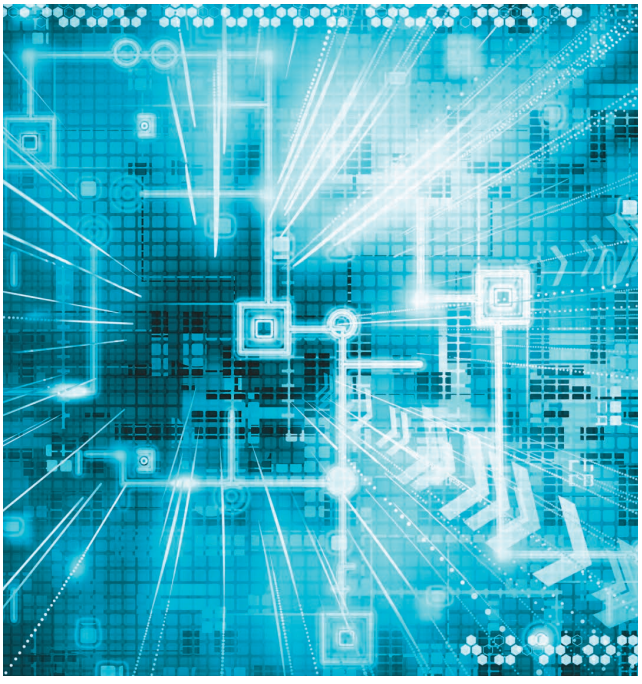

Empirical Research through Ubiquitous Data Collection

Niels Henze, Alireza Sahami Shirazi, and Albrecht Schmidt, *University of Stuttgart*

Martin Pielot, *Telefónica Research*

Florian Michahelles, *ETH Zurich*

Ubiquitous computing technologies enable new in situ data-collection opportunities that can reveal insights into mobile users' behaviors and preferences.

During the past few decades, research in psychology—and more recently in human-computer interaction (HCI)—has relied on highly controlled test settings. Widely accepted findings are often based on studies involving a handful of participants drawn from a specific population.

Like psychologists, HCI researchers commonly conduct studies with their own students. The laboratory provides a reproducible, albeit artificial, environment where the researchers can control all relevant factors, ensuring high internal validity.

The explosive growth in mobile computing, however, is undermining the utility of traditional lab studies. Smartphones, tablets, and notebooks have supplanted the desktop, and people use these devices in many different environments and situations while on the move.

Mobile systems must be studied in varying contexts, and even "living labs" can't capture the enormous diversity of real-world settings. Consequently, HCI researchers are beginning to use ubiquitous computing technologies to collect data for large-scale studies. Internal validity might be lower, but the results are likely to be more meaningful.

HCI RESEARCH IN THE LARGE

Ideally, a mobile system study should be conducted in all situations in which the system might be used and include numerous types of users. However, it's challenging even for a well-funded project to consider device usage in every possible context: in a café, on a train, in bed, on the beach, and so on. Furthermore, it's impractical to recruit representatives of the entire user base.

HCI researchers have explored various ways to conduct large-scale studies. Online experiments that mimic controlled studies present test subjects with simple questionnaires. However, researchers often must rely on convenience sampling when recruiting subjects, who often have little incentive to participate. An approach that taps crowdsourcing marketplaces such as Amazon Mechanical Turk isn't viable for mobile HCI studies because crowd members still use PCs at home.

An emerging solution to the difficulty of conducting mobile HCI studies with high external validity is researching HCI "in the large," which exploits the fact that today's computing devices offer uniform ways of distributing applications and services. The sidebar "Ten Steps to

 Published by the IEEE Computer Society

Conduct a Large-Scale Study" outlines one approach.

Around 2010, mobile HCI researchers began exploring mobile application stores as a way to study device usage. These researchers are developing mobile apps that, in contrast to commercial apps and games, record users' behavior to investigate specific questions with numerous degrees of freedom in parallel. In addition, thousands or even hundreds of thousands of users from all over the world could install a freely available mobile app, ensuring a diverse sample.

For example, one study of text entry behavior investigated the effects of three independent variables—shifting the touch position, shifting key labels, and visualizing touched positions—using *Type It!*, a typing game published on Google Play (N. Henze, E. Rukzio, and S. Boll, "Observational and Experimental Investigation of Typing Behaviour Using Virtual Keyboards on Mobile Devices," *Proc. SIGCHI Conf. Human Factors in Computing Systems* [CHI 12], ACM, 2012, pp. 2659-2668).

In total, the researchers examined keystroke data collected from more than 85,000 installations of the game (see Figure 1), enabling researchers to study 12 conditions for 50 different smartphone models. The experiment—the largest examination of text entry behavior ever conducted—wouldn't have been possible in a traditional controlled study due to the high number of variables considered. In addition, it had very high external validity compared to lab studies because it attracted participants of all ages around the world who played the game in a wide range places and situations.

Besides mobile games, HCI researchers have experimented with the use of widgets. One study obtained data using *appazaar*, a widget that recommends new apps based on previously used apps and the user's location (M. Böhmer et al., "Falling Asleep with Angry Birds, Facebook, and Kindle: A Large Scale Study on Mobile Application Usage," *Proc. 13th Int'l Conf. Human Computer Interaction with Mobile Devices and Services* [MobileHCI 11], ACM, 2011, pp. 47-56).

Also deployed on Google Play, the widget recorded which apps were used, where, and when by more than 4,000 people around the globe. With relatively little effort, the researchers collected a rich dataset that would be impossible to obtain using standard approaches.

HCI research in the large does have limitations. It's impossible to develop successful apps and services for all tasks of interest. Also, relying on users themselves to select apps can result in a skewed population sample. For example, a study of a mobile navigation system might

TEN STEPS TO CONDUCT A LARGE-SCALE STUDY

1. *Clearly identify the research goals*. Having concrete objectives will help focus the study. Collected data could later lead to investigating other questions.

Example: compare two virtual keyboards for use in smartphones.

2. *Select a study method*. In general, two approaches can be used to conduct a large-scale study: correlational and experimental. While correlational research can help identify and describe phenomena, experiments explore the underlying causes.

Example: experimentally assess various keyboard improvements, such as moving the location of functions.

3. *Devise an incentive mechanism*. Rewards are needed to motivate widespread participation in a study. These can come in many forms, including recreation, access to information, reputation, or financial compensation such as money, a voucher, or a discount.

Example: fun!

4. *Choose the target platform(s)*. There are many platforms for which applications can be created that act as experimental data collectors—various types of smartphones, the iPad, widgets, and so on. The choice of platform determines the user population and context.

Example: Android phone app.

5. *Design and develop the app*. The app's design must be engaging or compelling on one hand and an effective data-gathering tool on the other.

Example: text entry game.

6. *Prepare data collection*. Design a system to record all relevant data from the app and transmit it to the study's servers; implement privacy-protection schemes and policies—for example, anonymize Global Positioning System data to not reveal users' location.

Example: record the screen position of typing events and acceleration values in *x*, *y*, and *z*.

7. *Implement scheme to obtain informed consent from users*. Include explicit mechanisms to inform app users that they're participants in a study and obtain consent to collection of their data for research purposes.

Example: include information about the study in the app's description, dialogues, and supporting documentation; have users click a box at start-up signifying their consent to data collection for research purposes before they can run the app.

8. *Distribute and promote the app*. To attract an initial audience, publish the app to a well-known and easily accessible site, such as an app store, and promote it through compelling language in app descriptions, advertisements, blogs, and social networks.

Example: social media sharing and paid ads in a search engine.

9. *Continuously monitor data collection for a designated time period*. Designate time milestones for the study—a month, three months, six months, and so on—and begin collecting data after publishing the app. Monitor servers closely, as a sudden increase in the size of the user base can require a large amount of additional resources to avoid service slowdowns or disruptions.

Example: collect text input data from 32,000 users over three months.

10. *Filter and analyze data to answer the research question*. Partition data as necessary to focus on the question at hand. Data analysis can occur while the experiment is ongoing—it might be necessary to extend the time window to obtain sufficient data to reach a conclusion.

Example: users type faster on keyboard layout A than on layout B.

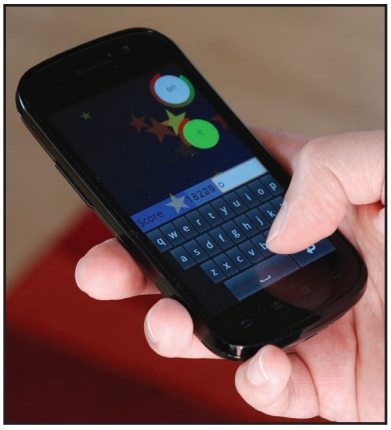

Figure 1. To study the effects of different variables on mobile text entry behavior, researchers developed the game *Type It!*, which to date has been downloaded from Google Play by more than 200,000 users.

obtain a good sample of those familiar with such systems but could exclude potential first-time users.

LOOKING BEYOND HCI

Deploying mobile apps and widgets to answer research questions can provide insights beyond HCI.

For example, transportation researchers spend much effort recording pedestrian walking speeds to time light lengths at street crossings. It's possible to approximate this metric using a mobile phone's accelerometer, and obtaining sampling data is simple using smartphones. A widely deployed widget that collects this data could yield insights into walking behavior on a global scale.

Researchers could use a similar approach to collect distributed data from mobile users on several questions: What is the average distance that people walk from where they park their car to a restaurant? How does sleeping time differ between weekdays and weekends? How often do people go to the hairdresser? With smartphones continually adding more functions, such as electronic payments and near-field communication, researchers can address a broad array of subjects.

Experimental research in software engineering could likewise be scaled up to investigate the impact of programming languages on errors or the design of APIs using appropriate development tools that collect distributed user information.

INFORMED CONSENT

User studies require informed consent from participants, but this can be particularly challenging when conducting research in the wild. In the case of designing apps for ubiquitous data collection, it's a simple matter to inform users that the app is a data-collection tool—for example, in a dialog that runs when the app is activated—and require users to click a box to signify their consent. More details could be provided elsewhere in the app and in the app store. However, this isn't equivalent to obtaining informed consent in a traditional lab study, where participants meet with researchers face to face.

A recent study revealed that simply telling device users what input is being recorded, such as their geolocation, leads to very different responses compared to explaining the data's meaning—for example, that the data reveals their position on a map (A. Morrison et al., "Informed Consent and Users' Attitudes to Logging in Large Scale Trials," *CHI 11 Extended Abstracts on Human Factors in Computing Systems* [CHI EA 11], ACM, 2011, pp. 1501-1506). Even if users are aware that their actions are being monitored, ethics demands that they be fully aware of the implications of their participation.

HCI research in the large isn't limited to academia. Companies that have a large user base can provide various app and website designs to their customers for evaluation. YouTube, for example, conducted a study to evaluate 1,024 small changes to three sections of its homepage (http://youtube-global.blogspot.de/2009/08/look-inside-1024-recipe-multivariate.html). Assessing that many variables is impossible in lab settings, but easy in the case of a site like YouTube with millions of users.

As all types of Internet-connected devices communicate information about themselves back to the webserver, researchers can conduct similar research to optimize content for PCs, mobile phones, TVs, tablets, and so on. Directly integrating the results of a continuous, large scale-study into product development creates many new innovation opportunities.

Niels Henze *is a researcher in the Human-Computer Interaction group in the Institute for Visualization and Interactive Systems at the University of Stuttgart, Germany. Contact him at niels.henze@vis.uni-stuttgart.de.*

Alireza Sahami Shirazi *is a PhD student in the Human-Computer Interaction group in the Institute for Visualization and Interactive Systems at the University of Stuttgart. Contact him at alireza.sahami@vis.uni-stuttgart.de.*

Albrecht Schmidt*, Invisible Computing column editor, heads the Human-Computer Interaction group in the Institute for Visualization and Interactive Systems at the University of Stuttgart. Contact him at albrecht@computer.org.*

Martin Pielot *is a researcher at Telefónica Research in Barcelona, Spain, focusing on HCI and ubiquitous computing. Contact him at pielot@tid.es.*

Florian Michahelles *is a lecturer at ETH Zurich as well as associate director of Auto-ID Labs at the University of St. Gallen, Switzerland. Contact him at fmichahelles@ethz.ch.*