An Automatic User Study Demo in Indoor Environments and its Privacy Implications

Reza Farivar, Mirko Montanari, Ellick Chan and Roy H. Campbell
University of Illinois at Urbana Champaign
Emails: {farivar2, mmontan2, emchan, rhc}@illinois.edu

***Abstract*—User studies usually involve much organization and manual labor. Even when performed correctly, a common problem of the studies is user bias. This occurs when participating users' knowledge of the study influences their actions. Another problem is the willingness of the users to participate at all. Finally, participants will always have privacy concerns. We have developed a framework to help with anonymized user studies, trying to solve the aforementioned problems. We have prepared a demo to show the effectiveness of our system.**

I. Introduction

User studies have always been challenging to perform. There are multiple issues involved in performing a user study, which include (but are not limited to) organization, user bias, willingness to participate and privacy concerns. As such, they are notoriously hard to organize. Another big concern is the awareness of the users participating in the study which biases their behavior. Furthermore, it is challenging to find enough users willing to participate in the study in the first place.

We have developed a system to help remedy these issues. To help better organize the study, we have tied it to a location-aware system. This way, user studies that require location information can be highly automated. To help alleviate user bias, we hide the mechanisms and parameters of the user study from the participants, while only informing them of the presence of the user study for ethical reasons. The users have the choice of participating in the study, but are unaware of the mechanisms of user study to ensure that it does not bias their behavior. To help attract people to the study, we provide them with locationbased services, providing them with a strong incentive to participate. In other words, we ask the user to carry around a device to help her in her daily routine, while she can ignore that she is in a user study since it is completely automated and inobtrusive. In any case, there will definitely be privacy concerns to be dealt with to ensure the experiment conforms to ethics standards. We aim to convince the users to participate by guaranteeing anonymity in our system to such an extent that it is impossible for an intruder to gain any knowledge about any individual user.

In this demo, we have implemented an indoor tour guide system using 802.11 localization, which is described in Section 2. The users will receive handheld devices that guide them with a tour of an area, in this case the area will be the demo exhibition of the Percom 2009 conference. The tour guide program will present a multimedia story of each Percom demo to the users automatically as they approach each table. The program will also provide them with some simple to use annotation capabilities to help them take notes during their visit to the demo room. These features are presented in detail in Section 3.

The tour guide program will also perform a separate data collection task in parallel, without the users' knowledge about its details. It will track and record the locations that the user has been to in the demo exhibition. This way, it can determine which demo tables were generally more interesting for users. Moreover, two of the internal sensors of the handheld device will be activated. The first is the microphone, which measures the ambient noise around the demo table, thereby trying to infer the overall popularity of each demo. The other sensor is the accelerometer. The hypothesis is that if a user engages a demo presenter, and becomes disinterested in the middle of the discussion, he or she will probably stay around for the discussion to complete, but may exhibit signs of disinterest and restlessness (to fidget around in the same spot). We aim to record these behaviors from multiple situations, and correlate them with the general level of interest or boredom each demo arouses in its visitors. These features will be discussed in Section 4.

While this approach could potentially remedy the problem of user bias (if the users knew about the features, they would make that sure they do not fidget in front of a demo table to avoid lowering the score of that demo table), one should be very careful implementing it. Generally, recording a person's behavior without his knowledge could be a potential violation of the user's privacy. To make sure that we do not cross the line, we have devised two levels of anonymization and log sanitization to guarantee non-retention of data. The first level is implemented in the handheld device, and the second level is built into the data-mining server. This way, we can guarantee that no one can pinpoint an individual user or associate a trace behavior to a user from the results of the experiment. The anonymization of identities is described in Section 5.

II. Indoor 802.11 Localization

The localization system we have developed is based on 802.11 WiFi fingerprinting. This approach has been described in previous literature [1], [2], [3], [4], [5]. We have implemented it in a campus building using the existing wireless internet infrastructure. We have also developed a toolset to help us in fingerprinting a new building in a short time, which will be part of the demo during the conference [7].

978-1-4244-3304-9/09/$25.00 ©2009 IEEE

We have also developed a fingerprinting client on Apple iPhone and iPod touch devices. We plan to provide such devices with the fingerprinting program already installed and lend them to demo exhibition visitors. We also intend to have the program ready for download on Apple's AppStore, so that interested parties can use their own iPhone/iPod touch. (For simplicity, both will be called iPhone in this paper unless otherwise noted.)

III. The Tour Guide

There have been many previous attempts to develop tourist guide programs in location aware systems, see [6] as an example. However, gathering feedback from the users of those devices for the purpose of user studies has not been extensively explored. In the tour guide portion of our system, the program running on an iPhone will detect its location. Therefore in this demo, we show that by granting users with location based services, we can provide them with a compelling reason to participate in a user study which is very unobtrusive to their experience; all they need to do is carry the device around.

Using a predeployment training phase, the system can also associate a bounded geographic area to a certain context (e.g. a demo table). When the user approaches the general area of a table, the program will display prerecorded information to the user.

We have developed an easy to use content authoring toolset to aid in creating multimedia content. The content is created in HTML format, and the client can either download it from a server as it approaches a certain location, or it can have the context stored locally in the device. In either case, an embedded browser in the iPhone program will show the context to the user. The context could potentially include abstracts of each demo, pictures of the demo posters and perhaps an audio track with a short description of that demo.

It should be noted that as mentioned, the client can either get the information from a server or locally stored information. In case the Internet connectivity of the demo exhibition allows us to have the content stored on a server, we can also have dynamic updates to the content associated with each demo table. This way, even if we have limited access to other demos of the Percom conference beforehand, we can continuously add this information and update the contents database.

The tour guide application will be written using interactive web technologies, trying to provide additional services to the user. The user can optionally enter her email address into the tour guide. During the demo session, if she wants to keep track of a certain demo, she can press a button on the device's screen indicating that she wants the information about that specific demo to be forwarded to her specified email address. This way, users can keep track of projects and individuals they meet during the demo session by only pressing one button on the screen. This information is not used as part of the user study demo in any way. The feature is only provided for a better experience to the exhibition visitors.

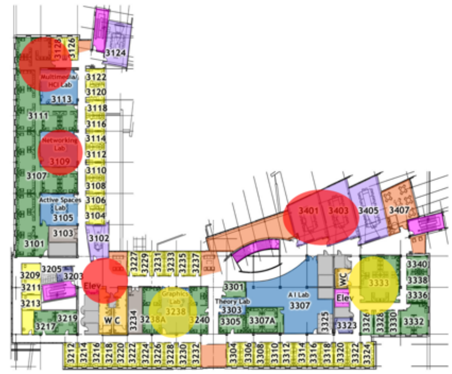

Fig. 1. A sample hot spots map of the building the system currently works in. Different colors show more visitor density in those areas, while the radius of the hot spots are related to the real world.

IV. User Study Sensors

As mentioned in the introduction section, there are three sensors involved in tracking the user's behavior or her surrounding events. The first sensor is the user's location in the time frame of the experiment. Thus, the system records the user's location as frequently as possible. Currently it takes about 4 seconds to perform a scan of the wireless access points, therefore for a one hour demo session at most 900 data points will be saved. The data mining server will then use these data points to identify which demo tables have been visited by the user, and how much time the user spent at each table.

The second sensor used in this demo is the iPhone's microphone. This sensor will be activated only on iPhone devices, as iPod touch devices lack one. The system will activate the microphone and listen in short one second bursts, and then it will compute the average of the waveform during the one second timeframe, and it will store this average power value with the location trace. These values will be used in the data mining engine to deduce the average sound level next to each demo table. The hypothesis is that popular demo tables will have a greater number of people standing close to it and perhaps more discussion.

The auditory data is not transmitted to the server to allay privacy concerns regarding recorded conversations. Also each one second audio clip will be deleted before obtaining the next sample. For more information about the privacy concerns of the demo, please refer to the next section.

The third sensor is the accelerometer in the handheld devices. This time, the system tries to infer body activity the users perform when at each table, and associate that with satisfaction levels. This analysis will be very roughly performed, and may be inaccurate. To increase the accuracy, this sensor only gets activated if the user remains in a small physical area for a minimum period of time, such as 15 to 20 seconds (equalling 4 to 5 WiFi fingerprint samples). This way, we hope to minimize the user movement in the demo hall during our analysis.

We envision three predominant scenarios. First, the user keeps the device steady in her hand while reading a demo poster or listening to the presenter. In this case, the more steady the device is, the more attention the user is paying, therefore it is a sign of her interest in that demo. In the second scenario, the user becomes bored and starts to "fidget" in place, waiting for the presenter's description to finish. In this case we expect to see small fluctuations in the sensor reading. Finally, the user might engage in a discussion with the presenter, in which case she might move her hands, and thus the device will record more pronounced movements of the device.

The movements of the device recorded with the accelerometer could show a wide variance, and therefore we do not include this data in the primary analysis, and will just visualize it as a possible interesting parameter.

As the demo session progresses, these sensor data streams will be sent to the data mining server on our own demo table, and will dynamically visualize a map of the room, showing hot spots in the room (see Figure 1). At the same time, a sorted list of demo tables based on visitor's interest in them will be updated and shown on our demo monitor. To not interfere with the "best demo" competition, this sorted list of demos will be anonymized, i.e. they will not contain the name of each demo, but only a numerical representation of each demo table is used in this sorted list.

V. Anonymization and Privacy

Even though the participants will be informed of their involvement in a user study, they will not know much about the types of information that are gathered during the demo. We have taken multiple steps in our system to ensure that the privacy of each individual is maintained.

As the first mechanism, we have designed the system to avoid using the optionally entered contact information of the user in the user study demo in any way. Therefore, the user study demo is agnostic to the carrier of the mobile device, and only considers where the mobile device is moving.

The audio recordings are also captured in 1 second bursts every 4 seconds. The recording is quickly distilled down to a power level indicator, and the recording itself is deleted as soon as this measurement is taken. The recordings are in no way transmitted to the server.

It has been shown that just removing user identifiers from location samples is insufficient for privacy purposes because a potential adversary could use trajectory information to track paths and follow users' footsteps home [8]. To make sure such privacy violations do not occur, as each user location sample arrives to the data mining server any indication of which device it was received from is removed. In other words, we are not gathering the paths of users. All we are interested in is the number of people standing next to each demo table at each point in time. The same mechanism is used for the audio power levels and accelerometer sensors.

Likewise, the demo monitor will only show aggregate information to the visitors, thus no individual can be tracked on the data mining monitor. In fact, due to previous anonymization techniques, there is not enough information to identify any specific user, even if we wanted to do so.

Demo Requirements

The main requirement for the demo to work correctly is to have the demo room equipped with multiple access points throughout the demo session. For successful localization, we need at least a handful of access-points "visible" through the demo exhibition hall.

It is imaginable that a demo hall will host a considerable number of visitors, if not packed with people. We have not experimented with our system in a packed room, but we expect that the impact of people in the room on wireless signals to be lower if the access-points are mounted on the ceiling of the room, which is customary in many business environments. Nevertheless, it would be beneficial if we can learn of the approximate location of access-points in the demo hall beforehand.

In case the access-point environment is not suitable for our demo, we would like to be able to setup a few wireless access-points of our own in the demo hall for the duration of the demo. Finally we need all the handheld devices to be connected to the Internet.

Acknowledgments

The authors would like to thank the following individuals who helped in different stages of the project: Alejandro Gutierrez, Abdullah Akce, Lokkju Brenner and Daniel Peebles. This project has been partially funded by a grant from Motorola Inc.

References

[1] LaMarca, A., Y. Chawathe, S. Consolvo, J. Hightower, I. Smith, J. Scott, T. Sohn, J. Howard, J. Hughes, F. Potter, J. Tabert, P. Powledge, G. Borriello, and B. Schilit (2005). Place lab: Device positioning using radio beacons in the wild. In Pervasive 2005: Third International Conference on Pervasive Computing, pp. 116133. Springer-Verlag.

[2] P. Bahl and V. N. Padmanabhan, RADAR: An in-building RF-based user location and tracking system, in IEEE INFOCOM, 2000, pp. 775784.

[3] A. M. Ladd, K. E. Bekris, A. Rudys, G. Marceau, L. E. Kavraki, and D. S. Wallach, Robotics-based location sensing using wireless ethernet, in ACM MOBICOM, Atlanta, GA, September 2002.

[4] P. Krishnan, A. S. Krishnakumar, W.-H. Ju, C. Mallows, and S. Ganu, A system for LEASE: System for location estimation assisted by stationary emitters for indoor RF wireless networks, in IEEE Infocom, Hongkong, March 2004.

[5] M. Youssef, A. Agrawala, and U. Shankar, WLAN location determination via clustering and probability distributions, University of Maryland, College Park, MD, Tech. Rep., March 2003.

[6] Keith Cheverst, Nigel Davies, Keith Mitchell, Adrian Friday, Christos Efstratiou. "Developing a context-aware electronic tourist guide: some issues and experiences." CHI 2000:17-24.

[7] Gruteser, M., and D. Grunwald. 2003. Anonymous Usage of Location-Based Services Through Spatial and Temporal Cloaking. In Proceedings of the 1st international conference on Mobile systems, applications and services, ACM Press New York, NY, USA, p. 31-42.

[8] Hoh, B. and Gruteser, M. 2005. Protecting Location Privacy Through Path Confusion. In Proceedings of the First international Conference on Security and Privacy For Emerging Areas in Communications Networks (September 05 - 09, 2005). SECURECOMM. IEEE Computer Society, Washington, DC, 194-205. DOI= http://dx.doi.org/10.1109/SECURECOMM.2005.33