Engaging participants for collaborative sensing of human mobility

Helena Rodrigues, Maria João Nicolau, Rui João José, Adriano Moreira
Centro Algoritmi
Universidade do Minho
{helena.rodrigues, joao.nicolau, rui.jose, adriano.moreira}@algoritmi.uminho.pt

ABSTRACT
Human mobility has been widely studied for a variety of purposes, from urban planning to the study of spread of diseases. These studies depend heavily on large datasets, and recent advances in collaborative sensing and WiFi infrastructures have created new opportunities for generating that data. However, these methods and procedures require the participation of a significant community of users through extended periods of time. In this paper, we address the problem of how to engage people to participate in the data collection process. We have conducted a user study on the utilisation of a mobile collaborative sensing application. We have found that users react positively to campaigns, but it is difficult to keep them participating for long periods of time. We also hypothesise that one must close the loop, rewarding the participants with services based on the collected data, eventually showing that there is added value obtainable from crowd sourcing.

Author Keywords
Human mobility, WiFi networks, collaborative sensing, user engagement

ACM Classification Keywords
Sensors, Human factors, Data sharing, Collaborative computing, Privacy, Public networks

General Terms
Experimentation, Human factors, Design

INTRODUCTION
Understanding human mobility at a wide range of scales – small (building) to global (world) – has been proven to be important for a variety of application areas, such as the study of spread of diseases, optimization of telecommunication networks, urban planning or tourism management, just to name a few. As such, understanding and modelling human mobility has become an active topic of research in the past few years [1-2]. In particular, the increasingly ubiquitous presence of mobile devices has provided an important new tool for generating data and uncovering new knowledge about the realities of our world and our mobility behaviours [2]. Personal devices have the great advantage of being carried in our daily lives and being able to collect, process and communicate information. The most obvious example is the mobile phone, the most ubiquitous of all personal devices, but many other devices can help in collecting personal data, such as a GPS receiver, either individually or as part of a car navigation system, a digital pedometer, a portable computer, or dedicated sensors. The development of procedures and platforms to capture and share this type of data is a research area that has been particularly active in recent years, and several mechanisms have been explored both in terms of collecting individual data and of collaborative collection.

UbiComp'12, September 5-8, 2012, Pittsburgh, USA.
ACM 978-1-4503-1224-0/12/09.

In our research we seek to define methods and procedures for collecting data for human mobility analysis. In the context of the SUM project (Sensing and understanding human motion dynamics), we adopted two complementary approaches to collect data about the mobility of people. One is to exploit the existing infrastructure of the eduroam network, which is a global access and roaming wireless infrastructure used in most of the universities across Europe. The other approach, which we will describe in this paper, is to involve users of WiFi networks in the data collection process. These two approaches can complement each other very well, but the main advantage of the second one is the ability to support a much wider observation space when compared to the eduroam solution, as we gain the possibility to track users behind the spatial scope of the universities premises. There are, however, considerable challenges, the most relevant being the difficulty in creating a significant community of participants and keep them participating for a long period of time. A larger user base is essential to cover a larger spatial context and generate sufficiently significant data, but participants must recognize the usefulness of the system or/and be reward for its use. One approach is to highlight the benefits of the process for the society or community. Examples include the contribution for a common social goal, such as collecting data about air quality, or reporting problems in public facilities. Other approaches might be based on monetary rewarding for the data contributed to the system, where the amount of data, the quality, the uniqueness or the timely uploading of each contribution is valued over other contributions [3]. Payments based on these criteria or based on reverse auctions can be used [4]. Yet another alternative is to offer a service based on the data itself. In this last case, participants accept to collect and contribute the sensed data as a condition for using the service. These services might use the individually collected data only, or make use of the aggregate data collected by the entire group of participants.

Our approach combines opportunistic and participatory sensing [5] and is based on a social application for laptop computers that allows users to disseminate text messages to nearby users. As part of its operation, the application collects data about the nearby Access Points (APs) making it possible to register the locations visited by persons and the movement between those locations. In this paper, we describe the design and implementation of this mobile sensing application and present our preliminary results concerning the engagement of participants.

THE EPI APPLICATION

The first major design decision has been the use of everyday devices for data collection, thus avoiding the need for specific infrastructures or the distribution of devices specifically for data collection, e.g. GPS receivers [6]. Another design decision has been to explore the use of laptops instead of mobile phones. Mobile phones have been extensively explored for this purpose because their continuous presence in people's lives, coupled with their already substantial data capture and connectivity capabilities, have made them powerful sensing devices for human behaviour. However, the use of laptop computers instead of smartphones also has its own advantages. Firstly, the execution of our application is much less likely to impact the computer performance and energy efficiency than it would on a mobile phone. Secondly, the larger display can make it easier to develop more attractive and friendly user interfaces. Finally, and even though laptop computers are less used on the move when compared with smartphones, people spend around 87% of their time indoors [7], and an increasing number of places is served by WiFi networks where people are more likely to switch on their devices. Therefore, laptop computers should be able to at least provide a complimentary approach to the process of mobility data collection.

To explore this approach, we have developed the Epi application (Figure 1) that enables users to share text messages in the neighbourhood through existing WiFi networks. We collect data about nearby APs because we need to define the scope of neighbourhood where messages can be shared. In this case, the provided functionality does not benefit directly from the collected data, but collecting data is a requirement for the application to work.

The sharing of text messages among nearby users provides an additional benefit for the study of human mobility. By tracking the messages exchanged between nearby users, it is easier to detect encounters between persons and also to study the spread of messages over time and space. This capability is particularly useful in studying the spread of diseases or even computer virus.

User privacy is another important issue in the context of collaborative mobile sensing as private data is collected and stored on global locations. An entity collecting data may not be able to guarantee that personal data will not be inadvertently used for different purposes. In the particular case of human mobility analysis, it may be possible to dissociate data from the sources without loosing important context, but how to guarantee that data will not be used to other purposes? In our approach we have defined a privacy policy that clearly describes which data is collected and how it will be used. In our system, apart from the information about the WiFi APs in the vicinity of the participant, we just collect the identifiers of the laptops where the Epi application is running. Moreover, participants are not required to register or provide any personal information. However, we are aware that, with appropriate data mining tools, the collected data reveals important information about the participants, e.g. where they live, and we will be studying the use of location blurring techniques that apply to certain locations.

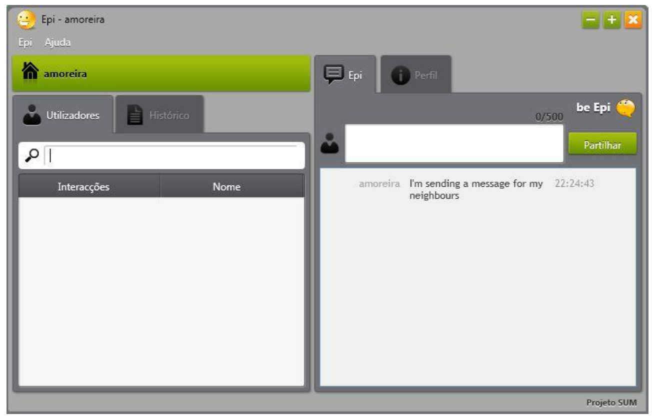

Figure 1. Main user interface of the Epi application.

The impact of the process on the reliability of data is also an important issue on collecting data for the study of human mobility. As our goal is to observe natural human mobility, we avoided any functionality that could induce participants to change their normal spatio-temporal behaviour.

Application technical details

Whenever a user of the Epi application creates a text message, the text string is broadcasted in the local network using IP multicast. Each multicast packet includes the text string, the nickname of the user, a timestamp, a validity period, and a WiFi fingerprint (list of nearby APs and corresponding signal strength) collected in the moment the message is sent. Since the local IP network might extend over a large spatial context, messages might reach a large number of other users. In order to limit the spatial scope of the neighbourhood context, when the message is received, the receiving application collects another WiFi fingerprint and compares it with the fingerprint in the received message. If they are similar enough, the receiving application decides that the message originated from a nearby application and shows it to the user. Otherwise, the message is discarded. Furthermore, the sent and received messages are kept in memory while within the validity period, and broadcasted again at predefined intervals.

Besides the functionality for editing, sending and reading received messages, the user interface includes a list of the users, identified by their nicknames, with which a larger number of interactions occurred (larger number of messages exchanged). This feature is similar to the concept of friends in Facebook, but with a different semantics: these friends do not have to know each other or invite each other for a group; they just have to share a common space and exchange some messages. These are “familiar strangers”. This feature tries to stimulate users to exploit the basic spread of text messages for new forms of social interaction.

Each message sent or received generates an internal record that includes the device id, a timestamp, the collected WiFi fingerprint, and a summary of the text message (computed using MD5). In order to extend the data collection capabilities of the application, additional fingerprints are collected at regular time intervals (typically 15 minutes) and stored as additional records.

Up to this point, the applications communicate in peer-to-peer mode, without the support of any network service. At regular intervals, the application checks the availability of Internet access and, if available, uploads the stored records to a central server.

PRELIMINARY RESULTS

We promoted two main campaigns corresponding to two different versions of the Epi application. The participants involved in our experiment were invited using non-personalized advertising. For attracting them, we set up a web site describing the project, the privacy policy, and how the application could be used. It also provided the access to the application through downloading of an installation package. We also set up a Facebook and a Twitter account to allow participants to send us feedback or request assistance. The web site, the Facebook and the Twitter pages were advertised in two time periods, July and October 2010, using email messages sent to all students and staff members of the university. Two physical banners were also displayed at the two university campus for about two weeks. An additional campaign was performed in September 2011 but just by sending email messages to a shorter group of students.

The following graphs illustrate some of the results we achieved with these campaigns. Figure 2 shows the number of persons we were able to attract to visit the web site. It is clear that the visits are highly time correlated with the advertising initiatives. Although the number of unique visitors per day is considerable, taking into account that our university has a population of around 20.000 people, including students and staff, the success in attracting people to the web site was limited. Figure 3 shows the number of persons that downloaded the application. Again, the number of downloads per day is highly time correlated with the advertising actions.

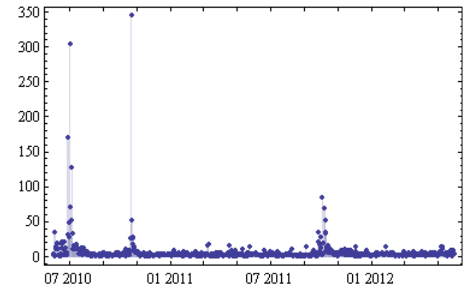

Figure 2. Web site: unique visitors per day.

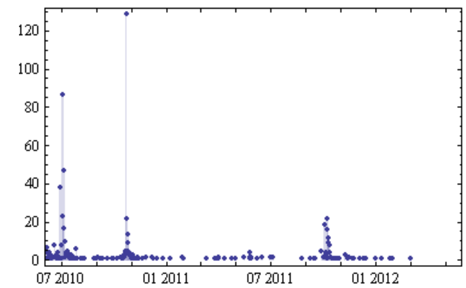

Figure 3. Web site: number of downloads per day.

Figures 4 and 5 illustrate the use of the application. Figure 4 accounts for the number of active users as those that were active in explicitly using the application to create and send text messages. These results show that most users tried the application for some time, but then gradually stopped using it. Through some comments on the Facebook page, we identified one of the possible causes for this behavior: a small number of users makes the probability of encountering another user nearby, with whom exchanging messages would be possible, quite low. Figure 5 shows the number of records uploaded to the server. It suggests that a considerable number of users kept the application running in their computers, as confirmed by the number of data records uploaded to the server per week, even after the users stopped creating and sending text messages.

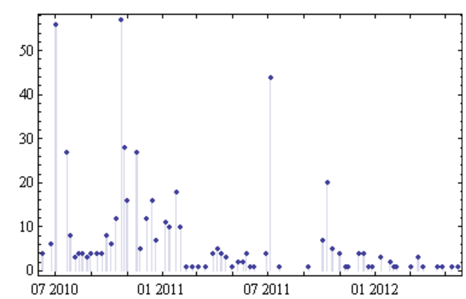

Figure 4. Application: number of active users per week.

The total data collected by these participants allowed us to identify more than 7.000 different APs, giving us a measure of the spatial extent of the collection campaign.

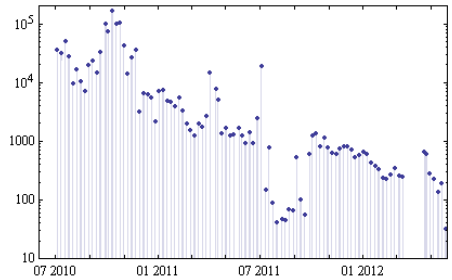

Figure 5. Collected data: number of records uploaded per week.

OTHER RELATED WORK

Collaborative mobile sensing for human mobility analysis has been widely explored in different projects. Dockerman and his team [8] used the trajectories of different items, such as user tracked US dollars bills around the United States and "travel bugs" in GeoCaching systems to capture human mobility, as there is a correlation between the trajectories of those items and the motion of the persons that carry them.

Martha González [1] and her team explored an infrastructure-based sensing system and used a data set containing positioning records of around 100 000 users of a cellular network collected over a period of six months. Cellular networks are a better proxy to observe the trajectories of humans than non-personal items. However, the majority of the sensing data records correspond to position data collected only when a person initiates/receives a phone call or a SMS message.

In other projects, data about the trajectories of people in urban environments have been explicitly collected through the use of GPS. An example is the Spatial Metro project [6], where GPS trajectories are being used to help urban planners and architects to better adapt several city centres to pedestrians.

Mobile sensing is also used in collaborative sensing contexts [9, 10]. In [9] it is explored the sharing of sensed personal data with social networks. In [10] it is described a general-purpose framework for anonymous opportunistic tasking and reporting of anonymously sensed data.

What mainly distinguishes our work from the previous ones is that we developed an infrastructure and a process specifically for getting access to data about human mobility in extended urban contexts, where special attention has been given to engaging users for long periods of time.

CONCLUSIONS

Collaborative mobile sensing has the potential to enable the collection of large data sets about human mobility.

Calling the attention of participants for such type of projects may be supported by well designed advertise campaigns. However, getting them actively involved for long periods of time is still a big challenge. Integrating social networking functionality into the sensing process might help to keep participants active, but achieving a critical mass on the number of users seems to be essential.

Currently we are studying how to extend the Epi application to improve the users' awareness about the value of the data collected collaboratively. One solution might be to extend the existent social networking functionality through the use of the collected data, thus closing the loop from the data repository to the participant.

Acknowledgements. Research group supported by FEDER Funds through the COMPETE and National Funds through FCT – Fundação para a Ciência e a Tecnologia under the Projects SUM – Sensing and Understanding human Motion dynamics (PTDC/EIA-EIA/113933/2009) and FCOMP-01-FEDER-0124-022674.

REFERENCES

1. Gonzalez, M., Hidalgo, C.A. and Barabasi, A.-L. Understanding individual human mobility patterns. *Nature 453*, 7196 (2008), 779–782.
2. Abdelzaher, T., et al. Mobiscopes for Human Spaces. *IEEE Computing 6*, 2 (2007), 20-29.
3. Lee, J.-S. and Hoh, B. Sell your experiences: a market mechanism based incentive for participatory sensing. In *Proc. Percom 2010*, IEEE Press (2010), 60-68.
4. Reddy, S. et al. Examining Micro-Payments for Participatory Sensing Data Collections. In *Proc. Ubicomp 2010*, ACM Press (2010), 33-36.
5. Lane, N.D., Eisenman, S.B., Musolesi, M., Miluzzo, E. and Campbell, A.T. Urban sensing systems: opportunistic or participatory? In *Proc. Mobile computing systems and applications* 2008, ACM Press (2008), 11-16.
6. van der Spek, S. Mapping Pedestrian Movement: Using Tracking Technologies in Koblenz. *Lecture Notes in Geoinformation and Cartography* (2009), 95-118.
7. Klepeis, N.E., Nelson, W.C., Ott, W.R., Robinson, J.P., Tsang, A.M., Switzer, P., Behar, J.V., Hern, S.C. et al. The national human activity pattern survey (NHAPS): a resource for assessing exposure to environmental pollutants. *Journal of Exposure Analysis and Environmental Epidemiology 11*, 3 (2011), 231–252.
8. Brockmann, D. and Theis, F. Money Circulation, Trackable Items, and the Emergence of Universal Human Mobility Patterns. *IEEE Pervasive 7*, 4 (2008), 28-35.
9. Lane, N.D. et al. A Survey on Mobile Phone Sensing. *IEEE Communications Magazine 48*, 9 (2010), 140-150.
10. Shin, M., Cornelius, C., Peebles, D., Kapadia, A., Kotz, D. and Triandopoulo, N. AnonySense: A system for anonymous opportunistic sensing, *Pervasive Mob. Comput. 7*, 1 (2011), 16-30.