Time-shifted positioning for location-based logging in mobile devices

André Coelho, Filipe Meneses and Rui José
Universidade do Minho/Centro Algoritmi, Guimarães, Portugal.
acoelho@dsi.uminho.pt, meneses@dsi.uminho.pt, rui@dsi.uminho.pt

***Abstract*—In this paper, we explore the viability and possible limitations of a time-shifted positioning process for location-based logging in which we want to annotate certain events registered on a mobile phone with location data. Instead of estimating the position when the event occurs, the system stores the radio and Wi-Fi data. At some later stage, this data is passed to a server that transforms it into the respective position. This shifting of the moment of positioning avoids frequent communications, saving energy and possibly communication costs. We have conducted a controlled experiment to assess the accuracy of the positions obtained with this process and interestingly, they seem to suggest that the approach is not only viable, but also more accurate than the positions obtained in real-time.**

***Index Terms*—Localization; Mobile computing**

I. Introduction

Location information is critically important in mobile computing [1] and mobile phones are becoming increasingly effective in determining their own location. Considering their widespread use and their continuous presence in people's lives, this represents a major opportunity for many sorts of location-based logging. The popularity of platforms that offer substantial data capture and connectivity capabilities, such as Android and iPhone, is increasing even further the potential of mobile phones to become powerful sensing devices for uncovering new knowledge about the realities of our world and about Human behaviour [2].

In our research, we have used mobile phones as the basis for two types of location-based data collection: mobility patterns and experience sampling. In the studies about mobility patterns within the city, we need to collect traces of users moving across the city in their daily life. In the experience sampling studies [3], [4] we aim to register, either implicitly or as part of an explicit user action, events of various types as they occur in people's daily lives, e.g. a person changing the mobile phone's ringer mode. These events need to be annotated with location information as location is normally crucial for the interpretation of the events.

A. Positioning in mobile phones

The availability of location information in mobile devices has increased considerably in recent years. The Android platform, in which we conducted this study, includes support for the Google location API [5], [6]. This API allows the device to calculate its position using multiple types of data providers, the two most important ones being the GPS provider and the network provider, which makes a combined use of GSM and Wi-Fi data. With a GPS chip present, Android's GPS provider can serve up location information with a high degree of accuracy and no data costs, but with high energy costs. If the GPS information is not available, the network provider can be used instead, but access to the Google service and therefore data connectivity will be needed.

The information for the network provider is gathered through the use of two Android APIs, the TelephonyManager API[1] and the WifiManager API[2]. The first API generates information about the cellular network, such as the MCC (Mobile Country Code), MNC (Mobile Network Code), Cell ID, LAC (Location Area code) and signal level, for both the cell that the device is connected as well as neighbouring cells. The WifiManager API enables the device to perform a scan of the surrounding area and, retrieve the BSSIDs (MAC Addresses) and signal levels of all the detected Wi-Fi APs. When an application calls the *requestLocationUpdates()* method in the LocationManager API, this information is then sent to an undisclosed Google service and the respective position is returned.

B. Positioning requirements in location-based logging

The location requirements for location-based logging are fundamentally shaped by the need to frequently determine the position of the device, while minimising energy or monetary costs. As part of our studies, we need to recruit people and convince them to run our data collection applications on their own mobile phones and as part of their normal daily activities. This is crucial for generating realistic data and enabling larger scale studies. However, they will not have any direct benefit from the application or from the data they are generating. Therefore, it is absolutely essential that the frequent positioning procedures that are inevitable in a location-based logging application, do not result in significant energy or monetary costs as otherwise, they will become a severe obstacle to widespread adoption and volunteer recruitment.

The first implication of these requirements was to completely avoid the use of GPS. We would always need another solution because we need to consider indoor locations and the continuous use of the GPS would necessarily result in excessive power consumption. The second implication was to

[1] http://developer.android.com/reference/android/telephony/TelephonyManager.html

[2] http://developer.android.com/reference/android/net/wifi/WifiManager.html

avoid communications and avoid depending on connectivity altogether. In part this is also important to save energy, but since many people will not have a flat rate data plan, they will not accept the potential costs associated with data communications. The independence from connectivity would also allow us to perform the positioning without having to wait for the availability of a network connection.

C. Time-shifted positioning

In most location-based services, location is normally part of an interactive feature that needs to be immediately provided to the user. A particular characteristic of our location-based logging processes is that the position information is merely used to annotate an event and if possible could be determined later, at a more convenient time. Therefore, we sought to address the above requirements by devising a process for time-shifted positioning in which the data needed to determine the position of the device is stored and the position is determined later. Instead of estimating a position straight away, the system stores the radio and Wi-Fi data that is normally used by the location API on the device to determine the position. The calculation of the location is deferred to a more convenient occasion. When a connection is available, a batch of radio and Wi-Fi information is sent to a server that will then use that information to calculate the position. This approach does not make any use of the GPS and may be used without any data communications. Therefore, it is a very promising approach for location-based logging applications in which we need to generate frequent positioning records without forcing the device owner to incur in significant power or network costs.

In this paper, we investigate the viability of this approach and the respective implications. In particular, we seek to understand to what extent a time-shifted approach may achieve the same level of accuracy and precision, as an on-demand approach.

II. Related Work

In [7] the authors detail the creation and deployment of a localization system based on Wi-Fi fingerprinting. The paper focuses on the effect and efficiency of the method used for acquisition of fingerprint data and, the influence that it could have on the accuracy of the localization results produced. Even though this work differs from our own, it enabled us to realize that we needed to devise a protocol for data collection that would ensure the best results possible and that would accommodate for the variation in the network landscape.

In [8] we find a work similar, in nature, to our own but with a different intent. The author used a mobile phone (iPhone 3G) to collect location data (A-GPS, Wi-Fi and Cellular positioning) at several distinct metropolitan locations, much like we did. With this, the author intended to test the accuracy of each of the iPhone's positioning methods against a benchmark location (ground truth), whereas we set out to test the accuracy of a time-shifted location calculation (with Wi-Fi and GSM data), against the location data provided by the device on-site.

TABLE I
Classification and references for data collection locations.

Ref	ID	Type
L1	**Home 1**	Outdoor
L2	**Campus canteen**	Outdoor
L3	**Campus room 1**	Indoor
L4	**Campus cafeteria**	Indoor
L5	**Campus library**	Indoor
L6	**Campus outside**	Outdoor
L7	**Café in the city**	Indoor
L8	**Home 2**	Outdoor
L9	**Street 1**	Outdoor
L10	**Street 2**	Outdoor
L11	**Cinema**	Indoor
L12	**Restaurant**	Indoor

III. Testing procedure

To investigate the viability and limitations of time-shifted positioning in an off-the-shelf mobile phone, we have devised a controlled experiment in which we conducted multiple positioning calculations at known locations using both the real-time and the time-shifted positioning.

We have selected a total of 12 locations for capturing the data. Considering that the nature of locations may affect the quality of the positioning process, we chose a diversified set of locations, including urban and rural areas. In particular, we have selected 6 indoor locations and 6 outdoor locations. Also, to address the fact that different locations may have much stronger and much more frequent variation in their Wi-Fi radio landscape than others, we included 5 location on campus, where we would expect the radio landscape to be more stable, and 7 on other locations across town, where a more dynamic landscape would be expected. The complete location list is described in Table I.

Additionally, we have defined a data collection schedule that ensured that the readings made at each location occurred at different times of the day and, we have defined a data capture protocol that explicitly addressed the possible effects that the time already spent at a location could have on the respective results. More specifically, we defined the following procedures:

- The mobile device had to be active and running the data capture application at least five minutes before approaching the registration point;
- Upon reaching the registration point a record was immediately generated;
- After two minutes at the registration point a second record was generated;
- After four minutes at the registration point a third record was generated.

A. The Android application

To support the data collection process, we have developed an Android application to generate position records. Each record comprises the following information:

- Reading ID: A unique hash generated from the timestamp;

- Timestamp: A timestamp of when the reading was done;
- LocationID: A key identifying the test location where the reading was made;
- Real-time network position: The position of the device as estimated at that moment using the network provider of the Google Location API;
- Network landscape: This is the information about the Wi-Fi and GSM landscape at the moment of the reading.

To execute the data collection protocol, the researcher would activate the application at some distance from the reference location. He would then approach the reference location and immediately, upon arrival, trigger the data collection process. This would immediately generate the first record. The researcher would then wait at the reference location for the application to automatically generate the second record, two minutes later, and the third record, four minutes later. These records were stored on the mobile device to transfer to a server, at a later stage.

We have identified a shortcoming in the Android API that limited the type of radio information we have used. When the phone is operating in 3G mode the API returns -1 as the default value for cell IDs and LAC of neighbouring cells. The only valid information is for the cell to which the device is currently connected. As a result, we decided to change the device configuration to operate strictly in 2G mode. In this mode, we were able to include all the information from neighbouring cells, but we have ignored the possibly valuable contribution that 3G cells could have for the positioning process.

B. Time-shifted positioning

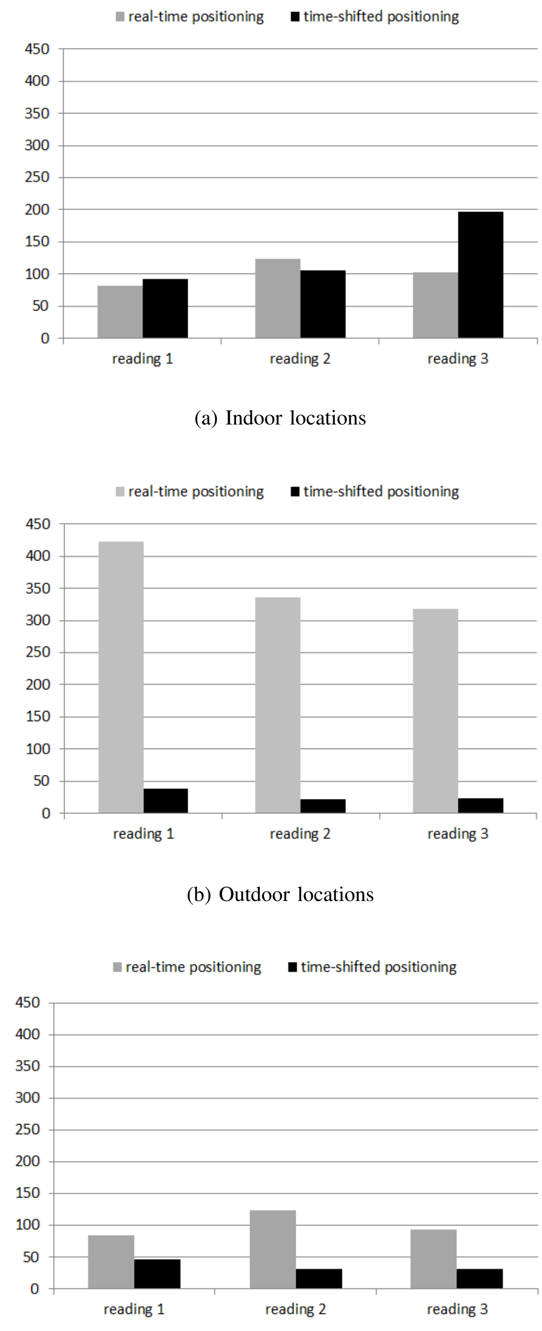
The position records generated by the Android application were uploaded to a server. Once in the server, a set of Perl scripts were used to estimate the position using the network landscape information collected at the location. The scripts parsed each record and generated a JSON request with the GSM and Wi-Fi information. This request was then posted through HTTP to the Google geolocation service[3] that returned a JSON reply containing the estimated position. The script decoded the JSON response and stored the respective position as the time-shifted position for that record.

IV. RESULTS

Using the Android application we created position records at the 12 locations during a three week period, with three weekly observations at each of the reference locations, two times a day. This resulted in a total 54 readings per location, or 648 readings in total.

To compare the accuracies of the various sets of position records, we have defined our ground truth by setting a reference position for each of the locations in our study using the coordinates obtained from Google maps. Then, for each location, we have calculated the distance from this reference position to each of the position records, including the time-shifted positions.

[3]http://www.google.com/loc/json



(a) Indoor locations

(b) Outdoor locations

(c) Campus locations

Fig. 1. Average distance in meters between locations and ground truth for 1(a) indoor, 1(b) outdoor and 1(c) campus

The graphics in Fig. 1 display the errors obtained for particular groups of locations. In each graphic, we have the average error for the first, second, and third reading, i.e. upon arrival, 2 minutes later and 4 minutes later. For each of these 3 readings we show the average error obtained for the positions that were calculated in real-time at the moment of the reading and, the average error for the positions that were estimated later on the server using the radio and Wi-Fi information collected at the moment of the reading.

Fig. 1(a) shows the results for the set of indoor locations in our study (6 locations). In this case there are no significant differences across the various variables. Fig. 1(b) shows the results for the set of outdoor locations in our study (6 locations). In this case, the positions calculated on the server show significant improvements in accuracy. Fig. 1(c) shows the results for the set of campus locations in our study (5 locations). In this case, the positions calculated on the server also seem to be more accurate, although the differences are not as relevant as in the outdoor scenario.

By using all of the available information (Wi-Fi+GSM), we may have forced the system into producing results with better accuracy than the one obtained on-site and that is probably only based on Wi-Fi information, unless this is very poor or non-existent, in which case, the GSM information is also used. The positioning results obtained by the device on-site could potentially be derived from only part of the available network landscape data, for reasons of speed or efficiency vs. accuracy, some sort of trade-off could be in place on the device. Also, the outdoor locations we designated are all located in areas with high residential density where there are always several Wi-Fi networks, reason for which, the signal would always be sufficient for performing localization even though its strength is not as good as it is in indoor locations.

V. Conclusions and future work

The main conclusion of this work is that a time-shifted positioning process is a perfectly viable alternative approach for our usage scenario and can even improve the accuracy of the positions. This is a very important contribution to inform the design of any sort of location-based logging tools in which, as in our case, the position information is not needed at the moment of logging. The second conclusion is that time spent at the target location can slightly improve the accuracy of the time-shifted positions, but the improvement is not very significant. Moreover, there seems to be no gain whatsoever in staying more than 2 minutes at a given location. Together these two observations seem to indicate that this process will perform well even for the generation of location traces in high mobility scenarios.

A limitation of this study is our lack of knowledge about the internals of the Google location API. There are no public details about how it uses the radio and Wi-Fi information to calculate position, and whatever the current approach might be, it may suddenly change without any prior announcement. Such changes could possibly affect the results obtained in this study and lead to potentially different conclusions. Also, the internals of specific devices and particularly their support for capturing radio and Wi-Fi information may also vary and lead to potentially different results in specific types of mobile phones.

References

[1] S. J. Vaughan-Nichols, "Will Mobile Computing's Future Be Location, Location, Location?" *Computer*, vol. 42, no. 2, pp. 14–17, Feb. 2009. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4781962

[2] T. Reichenbacher, "Geographic relevance in mobile services," in *Proceedings of the 2nd International Workshop on Location and the Web - LOCWEB '09*. New York, New York, USA: ACM Press, 2009, pp. 1–4. [Online]. Available: http://portal.acm.org/citation.cfm?doid=1507136.1507146

[3] J. M. Hektner, J. Schmidt, and M. Csikszentmihalyi, *Experience Sampling Method: Measuring the Quality of Everyday Life*. Sage Publications, Inc, 2006. [Online]. Available: http://www.amazon.com/Experience-Sampling-Method-Measuring-Everyday/dp/1412925576

[4] S. Consolvo and M. Walker, "Using the experience sampling method to evaluate ubicomp applications," *IEEE Pervasive Computing*, vol. 2, no. 2, pp. 24–31, Apr. 2003. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1203750

[5] Google, "LocationManager @ developer.android.com," 2011. [Online]. Available: http://developer.android.com/reference/android/location/LocationManager.html

[6] Google., "Obtaining user location @ developer.android.com," 2011. [Online]. Available: http://developer.android.com/guide/topics/location/obtaining-user-location.html

[7] H. Yoshida, S. Ito, and N. Kawaguchi, "Evaluation of Pre-Acquisition Methods for Position Estimation System using Wireless LAN," in *Third International Conference on Mobile Computing and Ubiquitous Networking (ICMU 2006)*, 2006, pp. 148–155. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.115.1414

[8] P. A. Zandbergen, "Accuracy of iPhone Locations: A Comparison of Assisted GPS, WiFi and Cellular Positioning," *Transactions in GIS*, vol. 13, pp. 5–25, Jun. 2009. [Online]. Available: http://doi.wiley.com/10.1111/j.1467-9671.2009.01152.x