Understanding and recognizing usage situations using context data available in mobile phones

Pekka Ala-Siuru
VTT Technical Research Center of Finland, Mobile Interaction
P.O.Box 1100,
FI-90571 Oulu, Finland
+358 20 722 2461
firstname.lastname@vtt.fi

Tapani Rantakokko
VTT Technical Research Center of Finland, Mobile Interaction
P.O.Box 1100,
FI-90571 Oulu, Finland
+358 20 722 2222
firstname.lastname@vtt.fi

ABSTRACT

In this paper we introduce a hybrid method for personalizing different interface features in a mobile phone. We give a description how to use combined data from GSM base stations and Bluetooth sources to determine user location and change the user profile automatically according to the location context. We give detailed information about our data logging experiments and the learning algorithms which were developed and tested.

Categories and Subject Descriptors

H.5.2 [**User Interfaces**]: Interaction styles. H.1.2 [**User / Machine Systems**]: Human factors. I.2.6 [**Learning**]: Knowledge acquisition.

General Terms

Algorithms, Management, Measurement, Experimentation, Human Factors.

Keywords

Context-awareness, Personalization, Mobile Computing, Location-Based Services, Case-Based Reasoning, Bluetooth, GSM Cell Id.

1. INTRODUCTION

One of the challenging areas of mobile computing device and software personalization is to use sensory data to obtain context information for personalization. This has been one of the main research issues in the pervasive computing community. Several research groups have been working to get the relevant sensory information to be used to build services for the users. Already in 1999 [10] Schmidt et al introduced experimental use of low sensory data in a mobile phone to be used to define different kind of contexts.

UbiPCMM06 September 18, 2006, California, USA

One of the interesting and perhaps most wanted contextual feature is to get personal devices to change the user profile automatically according to relevant context and different kind of methods have been used [10]. In this paper we will describe our hybrid approach to use radio beacon (GSM and Bluetooth) data to build context information and derive from that data cases for personalization. We will try get answers how Case-Based reasoning (CBR) is suitable for the task. This has already studied [7] and modeled in experimental service applications. The challenging questions still exist: is CBR suitable for learning from sensory data and can we get enough data for reasoning and understanding the user situation? From earlier research we know that CBR is suitable for applications which have changing situations. In order to derive context information one must search for behavioural patterns from the scanned data.

Our approach is related to the PlaceLab initiative [6, 9]. The main difference is that we are not using WLAN data and no fixed Bluetooth beacons (if PCs and Laptops in work place are not counted). Furthermore we are not using GPS data as in [1]. The group detection part is also related to that of Järkvik et al [5]. One of the largest projects where BT data is used has been carried out in the MIT project concerning social relationships in the campus area [4].

2. LOGGING CONTEXT DATA

Data logging was made by logging software which was running in a Bluetooth (BT) enabled mobile phone. The software for collecting information from smartphone usage and usage situations was developed in the Adamos project[1]. The software starts automatically when the phone is turned on, runs in the background without disturbing normal use of the device, and logs various data types into a file using consistent file format. The software is simply installed to the test user's phone, and after the test period the log file is copied from the phone to a computer for offline analysis. Consistent format for different data types allow automatic statistical analysis and easy comparison of data gathered from different users or usage situations.

Currently our logger collects various usage data such as keypad lock status, user activity, call status, foreground application, battery strength and charger status. Also usage situation data is

[1] ADAMOS : Adaptive Mobile Services – http://www.msh-alpes.prd.fr/ADAMOS/

logged, for example Bluetooth neighborhood as well as network area code and cell id, which provide rough location information.

Collected data has been used for usability and user experience studies, automatic learning of users' home and work locations, and as inputs for context-based actions. In the latter case, a well-defined context framework [8] provided means for personalizing various phone operations by combining context input sets to actions via simple XML scripts. The scripts can be generated with a simple user interface by the user or operator, or even automatically by learning algorithms.

The goal of our experimentation was, by determining user's location automatically, to change the user's phone profile to appropriate setting (e.g. in the work profile there could be different kind of applications available than in home or other place).

In earlier experiments we used only the GSM cell ids (CID) for location determination, but generally the CIDs give too rough coordinates. For instance in town areas where there are several base stations the cell data usually overlaps each other. Thus the id can change from one to another in same location several times in minutes. We used also logged time data together with the CID information. That gave a little bit more precision to the location determination.

Because of the use of Bluetooth for close area data communication in mobile phones we decided to log the BT address information and use it together with the CIDs for even better location accuracy.

2.1 Bluetooth

Bluetooth (BT) can be used to build a Personal Area Network (PAN). It is defined as an open standard for short-range transmission of digital voice and data between mobile devices (laptops, PDAs, phones) and desktop devices. Bluetooth devices are generally divided to two categories: Class 2 devices operate in the short distance (10-30 meters) and class 1 devices up to 100 meters. The main bandwidth for class 2 BT devices is 2.1 Mbit/s with speed of 2.4 GHz. This bandwidth is generally known as the Industry/Science/Medical (ISM) free band.

A BT device is identified by its MAC (Media Access Control) address and a possible user name as seen in the example given in the Table 1. From the data we can see that two BT devices has come into the range of the logger and one (D400) is not anymore available.

Table 1. A clip of a mobile phone data log.

Connections: BTDevice: 0010c64ed053:D800 InRange 2005/09/14 14:11:42 phone://TerminalEvents/
Connections: BTDevice: 0010c63a5cc8:Unknown InRange 2005/09/14 14:11:46 phone://TerminalEvents/
Connections:BTDevice:0010c63a5caf:D400 NotAvailable

Further we can observe the Bluetooth device type, its MAC address and name (laptop PC D800), data whether the device is in range or not, time of logging, and the running application in the logging phone.

2.2 Bluetooth logging (scanning)

To get information about BT data we walked thru different places. One route (figure 1) went from home to workplace and from there to university (distance ~ 500m), by car to downtown (distance 5 Kms) and back home. In the university we walked from library to cafeteria, bookstore and places where people were gathering together.

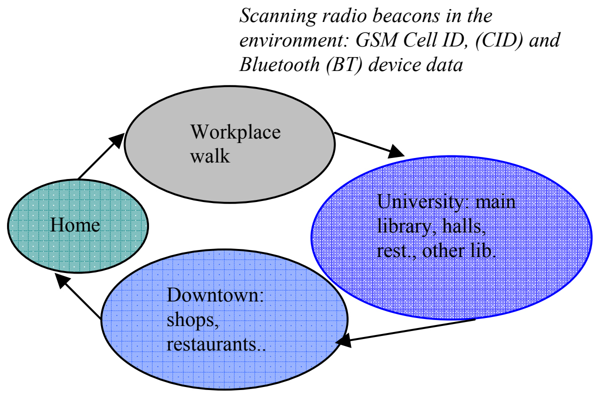

Figure 1. Data logging path (partially predefined).

The scanning software was on all the time and the logged data was downloaded to PC afterwards. To help the analyzing phase we used notebook markings (place, time and some situations) all the time (figure 2). The BT enabled devices which were found active were PCs, Laptops, PDAs and mobile phones. We didn't use any fixed BT or WLAN beacons.

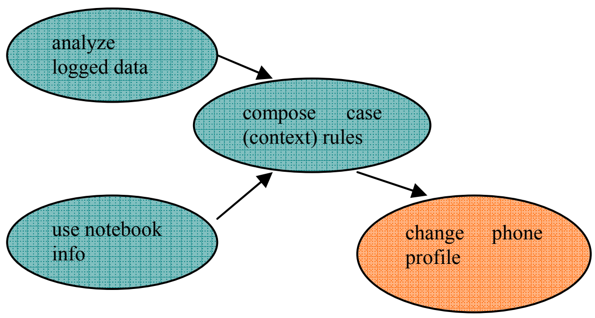

Figure 2. The overall flow of using context information to change the mobile phone profile.

2.3 GSM Cell ID (CID) observations

In general the GSM CID can give location information, but as we mentioned earlier it needs some supporting data to be useful.

Table 2. A clip of the logged data showing different CIDs

2005/09/14 14:30:55 phone://TerminalEvents/ Location:Network:CellID 12526

2005/09/14 14:32:21 phone://TerminalEvents/ Location:Network:CellID 12679

2005/09/14 14:35:38 phone://TerminalEvents/ Location:Network:CellID 12526

2005/09/14 14:35:54 phone://TerminalEvents/ Location:Network:CellID 12679

The data in table 2 shows that in the same physical location the CID have changed back and forth between two cells in five minutes. These CIDs shown in this example were picked up from earlier experiments to indicate work situation.

But actually this CID data can give misinformation if used alone. In the experiment environment the university area is close to our workplace and the CIDs can change although no mobility is observed. At least the other CID seen in the table 2 example appeared also in the university area.

To avoid this overlapping cell data, logged time data (time of day, weekend/weekday) was also used and gave better results. Based on these observations we decided to examine the BT data together with CID information.

3. EXPERIMENTATION FINDINGS

From the logged data we concentrated to examine four basic parameters. The amount of BT devices in range during given time slice, the difference of the seen BT device amount with different CIDs, the intensity of BT device occurrences in the scan (log) and the occurrence of one BT device with same BT devices (to define BT groups).

We analyzed the most CID occurrences known to appear in the work location and used them with the average amount of observed BT devices in the work location (Figure 3). Based on that study we derived these case rules:

Table 3. Workplace case definition rules for profile changing

IF CurrentBTCount => WorkBTmax

ANDF (CID = 12519 ORF CID = 12526 ORF CellID = 12679)

THEN CurrentContext= WORK ; CHANGE CurrentProfile

IF CurrentBTCount <= FreetimeBTmax

ANDF (CellID <> 12519 ORF CellID <> 12526 ORF CellID <> 12679)

THEN CurrentContext= OTHER; CHANGE CurrentProfile

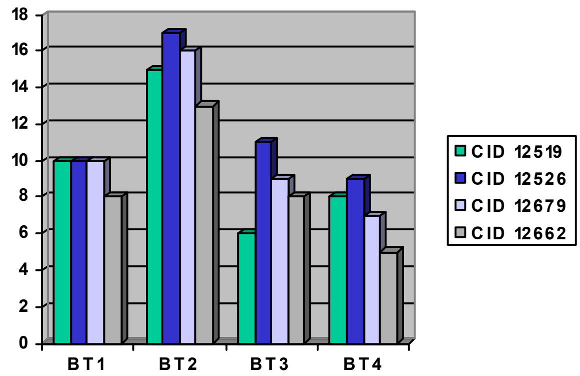

Figure 3. Most occurring BT MAC addresses in workplace and their CIDs. BT1 and 2 are laptops in the same room.

The max parameters were explicitly counted from the logged data and in this situation (figure 1) the parameters FreetimeBTMax was 4 and WorkBTmax was 16. The logged data showed also such amounts of BT devices which could indicate WORK situation but the CID didn't support that. In the case rule example nBTmax is the count of BT devices (work/leisure). The parameter CurrentProfile was changed accordingly.

When checking the reoccurrence of same BT devices it must be realized that the only significant BT -parameter is the BT MAC address. One must match the addresses from the log data one by one, because they don't show always in the same order.

The following table gives one example of the BT device occurrences in a workplace.

Table 4. Example log of the BT occurrences in work

2005/09/14 14:11:42phone://TerminalEvents/ Connections:BTDevice:0010c63a5cc8:UnknownInRange

2005/09/14 14:11:46 phone://TerminalEvents/ Connections:BTDevice:0010c63a5caf:D400 NotAvailable

2005/09/14 14:11:50 phone://TerminalEvents/ Connections:BTDevice:0010c64d9f85:oulkv1k149 NotAvailable

…

2005/09/14 14:23:07 phone://TerminalEvents/ Connections:BTDevice:0010c63a5cc8:Unknown NotAvailable

2005/09/14 14:30:55 phone://TerminalEvents/ Location:Network:CellID 12526

2005/09/14 14:32:21 phone://TerminalEvents/ Location:Network:CellID 12679

2005/09/14 14:35:38 phone://TerminalEvents/ Location:Network:CellID 12526

2005/09/14 14:35:54 phone://TerminalEvents/ Location:Network:CellID 12679

When studying the scanned BT devices and their variation and intensity with time and place we noticed that:

1 typically in our observations there was variation of scanned BT devices between work and other places
2 the amount of BT devices in the work time/place was in one test double than elsewhere
3 the smallest amount of BT devices was at home (1-2)
4 in restaurant and shop environments there was some increase (4-5) of BT devices as well as in the university area

3.1 Further experiment results

To further investigate the use of BT data we decided to make a longer scan (48 hours) and in different situations and places. Still the basic idea was to find out the usefulness of the BT data in defining different contexts. We had a partially predefined route as in the earlier experiment. During the test period we found 163 BT "InRange" occurrences (partially same devices). The earlier defined work CIDs could be mapped easily to the Bluetooth devices in the workplace. The same finding applied also to home location: the earlier experienced pair: (Same CIDs & Same BTs = Same Place). In this experiment we noticed also that 10-15 BTs were scanned only once or twice in a short time slice and according to notebook markings they were passers by or devices in a shop or restaurant.

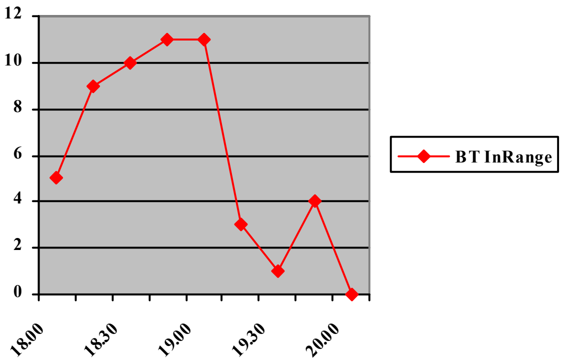

Figure 4. BT in range during a public place situation (school; parents meeting) 4 CIDs observed, 2 available all the time (same CIDs as in home location, 500m).

Figure 4 gives BT device occurrence data during parents meeting in the evening. The peak in the occurrence (11 devices scanned in same time) happened when nearly all teachers and parents were in the school's main hall in 10-20 meters proximity to each other. The use of this kind of information is quite demanding and hard to say if the situation will ever happen again. However, it gives some hints and reasoning advice. If we define this situation as a situation case with the known parameters (BTs + CIDs observed in given time) we can possibly afterwards use it as a basic case for learning same kind of situations.

The location change from home to work can be easily seen in figure 5, where the BT device amount changes strongly when one arrives to work place. So we can make some situation and location type conclusions based on the sudden change of scanned BT device amount together with the known CID data.

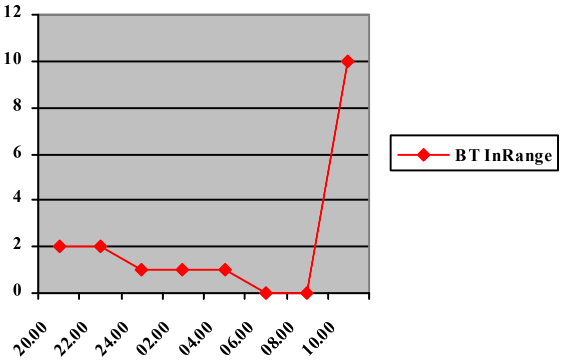

Figure 5. BT in range @ home location. 4 CIDs observed, 2 available all the time. Notice the increase of BT devices in range when coming to work (8 am).

Based partially to the case described in figure 4 we derived a case definition for a basic group (work, home, other). Assume that if a basic group of BT devices occurs several times together in given time slice and there are only couple of other devices seen sporadically we could define this as a basic group case. For this basic group we can define also guest groups from all those sporadically seen BT MACs.

Table 4. Group case definition

CASE BasicGroupA *is* BTMACa, BTMACb, BTMACc...BTMACx *with* CID1...CIDn *with GUESTa, GUESTb .. GUESTx* *use* ProfileGroupA

To define this kind of case we need several occurrences and to finally decide if the group is a real case we must define:

1. The correct time slice or count of occurrence for learning
   - this deals with the situations in (normal for updating a case database)
     - defining a group
     - defining guests
     - defining a new group (- /+ members)
     - deleting a group

4. CONCLUSIONS AND FUTURE WORK

We have described in this paper how to collect primary context data for a mobile phone and specially described our experiments in combining GSM CID data with Bluetooth data collected from mobile and other BT enabled devices (e.g. PCs and Laptops).

We defined some cases based on the logged CID and BT data. For changing the user profile in the mobile (smart) phone we defined case rules which have been tested and implemented first only as rules. This will be further studied. The use of BT data increased the accuracy of work location definition. Generally we found that CBR method seems to be useful: we found data which can be used to build cases.

1. CID with BT MAC address can be used to define *profile cases*
2. with help of learning case data CID and BT data can be used to build *group cases*

We were also thinking about the social and privacy issues when using the scanning software. We didn't mention anybody that we were scanning because according to our knowledge Bluetooth, WLAN, etc. access points are legal to scan in EU. Furthermore users of Bluetooth enabled devices can explicitly define if they release the data or not (BT enable/disable) and Bluetooth name can be left as unknown (table 4). Also according to [2, 3] social denial is not fear, because people tend to disclose location and other context data if they think it will be useful for them.

The data logging method that we used (walking and scanning) reminds the WarDriving[2] method:

- Using WarDriven (-Walked) data from Web could give more location information
- Oulu not yet WarDriven

WarDrive = Driving in public areas with WLAN & BT -devices scanning the public network with GPS device and mapping the areas with explicit coordinate data.

WarWalking = same but by walking one can get more accurate data from for instance fixed BT beacons.

5. ACKNOWLEDGMENTS

This work has been a part of the Finnish-French project ADAMOS (Adaptive Mobile Services). The work was supported by the Academy of Finland and French Science funding institutions. The partners were University of Oulu, VTT Technical Research Centre of Finland, CEA Leti, University of J. Fourier, CNRS, MSH Alpes, ST Microelectronics and France Telecom R&D. Thanks to all the active partners giving advice and commenting on this work.

6. REFERENCES

[1] Borriello Gaetano, Chalmers Matthew, LaMarca Anthony and Nixon Paddy, Delivering REAL-WORLD Ubiquitous Location Systems, *Communications of the ACM, March 2005, Vol 48, No.3*. 36-41. ACM Press 2005.

[2] Consolvo Sunny, Smith Ian E., Matthews Tara, LaMarca Anthony, Tabert Jason, Powledge Pauline, Location Disclosure to Social Relations: Why, When, & What People Want to Share, *CHI 2005 Conference Proceedings, pp. 81-90*, ACM 2005

[3] Danezis George, Lewis Stephen and Anderson Ross. How Much is Location Privacy Worth?. *Fourth Workshop on the Economics of Information Security (WEIS 2005)*. Harvard University, 2 - 3 June 2005.

[4] Eagle Nathan, Pentland Alex (Sandy), Reality mining: sensing complex social systems, *Personal and Ubiquitous Computing,* Nov 2005, Pages 1 – 14

[5] Järkvik et al . Group Detection Using Bluetooth, Poster presentation at UbiComp 2005.

[6] Kang Jong Hee, Welbourne William, Stuart Benjamin, Borriello Gaetano, Extracting Places from Traces of Locations, *ACM SIGMOBILE Mobile Computing and Communications Review, Vol. 9, Nr. 3*, 2005, pp. 58-68.

[7] Kofod-Petersen Anders, Aamodt Agnar, Case-Based Situation Assessment in a Mobile Context-Aware System, *Proceedings of the Ubicomp 2003 Workshop on Artificial Intelligence in Mobile System 2003*, October 12, Seattle, USA.

[8] Korpipää Panu. *Blackboard-based software framework and tool for mobile device context awareness*. Doctoral thesis. VTT Electronics, Espoo, 2005. 225p. VTT Publications : 579.

[9] Schilit, B. N., LaMarca, A., Borriello, G., Griswold, W. G., McDonald, D., Lazowska, E., Balachandran, A., Hong, J., and Iverson, V. 2003. Challenge: ubiquitous location-aware computing and the "place lab" initiative. *In Proceedings of the 1st ACM international Workshop on Wireless Mobile Applications and Services on WLAN Hotspots* (San Diego, CA, USA, September 19 - 19, 2003). WMASH '03. ACM Press, New York, NY, 29-35. DOI= http://doi.acm.org/10.1145/941326.941331

[10] Schmidt Albrecht, Aidoo Kofi Asante, Takaluoma Antti, Tuomela Urpo, Van Laerhoven Kristof, Van de Velde Walter, Advanced Interaction in Context, *Lecture Notes in Computer Science, Volume 1707*, Jan 1999, Page 89.

[2] http://wigle.net/)