



## Introduction

With the rapid development of mobile phone and applications, the context-aware technology based on mobile devices is paid more and more attention. Most researches on context-aware based on mobile phone have been proposed, which attempt to detect location or mobility state [1, 2, 3]. However, they couldn't directly infer current surrounding contextual behavior, yet usually need additional information. As Bluetooth technology could discover other Bluetooth devices nearby (if users set as discoverable) to initiate communications, we could easily inquire ambient Bluetooth devices to infer more valuable information. In contrast to previous methods based on Bluetooth sensors [4, 5, 6], which didn't specifically indicate or recognize contextual behavior, the aim of our research is intuitively obtain social contextual behavior in the surroundings only using dynamic Bluetooth information.

In this paper, our main idea is that a significant distinction between different context behaviors is the surrounding dynamic information, which can be inferred from the features of ambient Bluetooth devices. Specifically, building a contextual behavior classification model based on dynamic Bluetooth information of the environment, several representative contextual behaviors can be distinguished according to the Bluetooth features in terms of the quantity and dynamic changes of inquired Bluetooth devices, then we present an inference approach based on sliding window and context restraint to continuously discover the surrounding change and infer current social contextual behavior directly through the real-life Bluetooth traces.

## Methodology

### *Feature-based contextual behavior classification*

Building a contextual behavior classification model involves two steps: extracting features based on our collected Bluetooth data log and designing an appropriate machine learning model for classification. Since the surrounding Bluetooth information obviously changes over time, we should extract features from several successive samples instead of one isolated sample. We define a training timeslot  $t$  consisting of  $w$  samples as the smallest time unit. Based on the time unit, the algorithm can determine the contextual behavior in the surroundings.

In this work, we choose six representative contextual behaviors for training: *working indoors*, *walking outdoors*, *taking subway*, *go shopping in the mall*, *dining in the restaurant*, *watching movie in the cinema*, which are all named *situation*. Considering that if we can successfully infer them based on Bluetooth features alone, the results would contribute a large proportion of our daily lives. We extract the following six features from a particular timeslot  $t_i$ : *The number of Bluetooth devices*, *The rate of fixed devices*, *The mean and standard deviation for change rate of devices*, *The time and rate of durative appeared devices*. For these extracted features, C4.5 Decision Tree (DT) is trained to set up the contextual behavior classifier.

### *Identifying the situation of one minute*

Each time the sliding widow is updated, the classification model will be applied to determine the current timeslot. Therefore, each minute will be involved in sliding window  $w/2$  times as a consequence obtain  $w/2$  classification results. For example, assuming  $w = 6$ , the minute  $m_i$  will be involved in sliding window  $SW_1$ ,  $SW_2$ ,  $SW_3$  in turn. We determine the situation of one minute base on a simple "Majority Vote Rule": labeling a certain minute

with the classification that appears most frequently on the view. While if each situation appears the same time, the minute is determined by the last sliding window.

For the empty timeslot that contains no data, the situation is defined as *NoBluetooth*. And for a certain minute if every sliding window is *NoBluetooth*, then the minute is identified as *NoBluetooth*.

#### *Inference approach*

After identifying the situation of one minute, we will further do the inference depended on the type of situation and the previous one. More specifically, if the situation of current minute is the same as the last one, the situation remains stable to this minute. In contrast, if the situation of current minute is different from the last one, the ambient situation may change and this minute might be the transition time. Yet to avoid a single minute determining a change, a tolerance interval is used where the time length is as the same as the shortest situation length (5 minutes). When that happens, the detailed inference algorithm is described as follows:

Case 1: Among the tolerance interval, if one certain situation appears more than or equal to 3 minutes, the situation of the entire tolerance interval is described by this situation (It may be the same as the previous situation that specifies there is no change of situation). In addition, only if the whole tolerance interval is *NoBluetooth* that the situation is assigned as *NoBluetooth*, otherwise it is determined by other majority situation.

Case 2: If no situation is in the majority, the situation of the tolerance interval is inferred by the last situation.

- Case 2.1: If the last situation belongs to relative static situation(*working indoors, watching movie in*

*the cinema, dining in the restaurant*) or *NoBluetooth*, the presented chaos gave an obvious indicator that the situation probably have shifted. However, the reason we can't determine the contextual behavior directly is mainly because it is in the transition time between different contextual behaviors. We infer this situation is "walking outdoors" empirically because we observe it matches to the practical situation most closely.

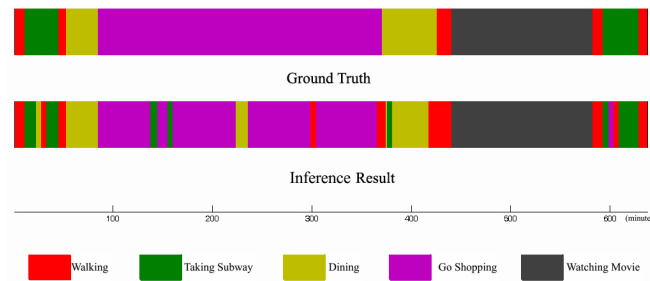
- Case 2.2: If the last situation belongs to relative dynamic situation (*walking outdoors, taking subway, go shopping in the mall*), it is difficult to infer whether it has changed or what current situation is. Thus the tolerance interval is still regarded as the last situation in this case.

### Experiment results

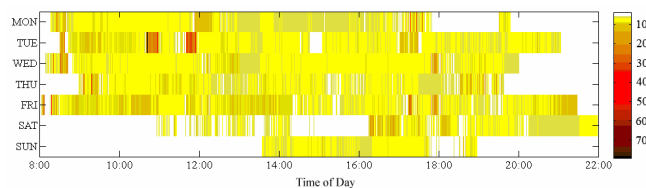
In the data collection, we recruit three volunteers who have different jobs and daily lives to collect Bluetooth traces for 1-2 weeks along with their normal lives, they are asked to recall and label their contextual behaviors for recording the ground truth at the end of each day.

Running on the real-life Bluetooth traces, we validate the inference algorithm based on the classifier determined by 10-fold cross validation, where plenty of Bluetooth devices could be inquired in most public places. Figure 1 shows the ground truth and inference result of one day trace (unit: minute), which illustrates the performance of the proposed algorithm, and experimental results show the inference accuracy of social contextual behavior is 85.8% averagely (not counting that none of Bluetooth devices is inquired in one timeslot, referred as *NoBluetooth*). According to the error distribution, we find that our approach based on Bluetooth trace is particularly sensitive

to the fluctuation of surroundings, causing misjudgments on the contextual behavior change may occur easily. Besides, the average percentage of time-length of *NoBluetooth* is 24.2% in our experiments. Figure 2 illustrates the density map of Bluetooth devices collected by one participant who mostly works in the office during a week continuously, from which also can be seen that in most public places, plenty of Bluetooth devices could be inquired. In fact, most *NoBluetooth* cases occur at home and some cases are in the outdoor environment. Specifically, when users are in the open outdoor regions or few people around and the changing rate is low, it is probably hard to inquire any Bluetooth device.



**Figure 1:** Inference performance with proposed algorithm. The Bluetooth trace is collected by one day.



**Figure 2:** The density Bluetooth device map of one week for a data collection participant.

## Conclusion and future work

This paper presents a novel approach for inferring social contextual behavior from Bluetooth traces. A Bluetooth feature based classification model of contextual behavior is constructed firstly, then context restraint based majority vote rule is further applied into the analysis of the real-life Bluetooth traces. The experimental results prove the feasibility of inferring social contextual behaviors through Bluetooth traces adequately. To future study this topic, we might focus on two aspects. The one is to make the inference result more accurate particularly when Bluetooth devices are detected sparsely. The other is to conduct some kinds of fusion researches combining Bluetooth with other sensor technologies like GPS, GSM, WiFi, etc.

## Acknowledgements

This work is supported by Natural Science Foundation of China (No.61173066, No.61070110), Beijing Natural Science Foundation (No.4112056), Co-building Program of Beijing Municipal Education Commission and Graduate Fellowship program from China Scholarship Council (CSC).

## References

- [1] Donnie H. Kim, Y. Kim, D. Estrin, M. B. Srivastava. SensLoc: Sensing Everyday Places and Paths using Less Energy. In *Proc. SenSys'10*, pages 43–56. ACM, 2010.
- [2] Martin Azizyan, I. Constandache, R. Choudhury. SurroundSense: Mobile Phone Localization via Ambience Fingerprinting. In *Proc. MobiCom'09*, pages 261–272. ACM, 2009.
- [3] Sasank Reddy, Min Mun, J. Burke, D. Estrin, etc. Using mobile phones to determine transportation modes. *ACM Transactions on Sensor Networks*, 6(2), 2010.
- [4] Donnie H. Kim, Dae-Ki Cho. BlueSense: Detect individuals, location and regular activities from bluetooth signals. <http://urban.cens.ucla.edu/cs219/images/0/0b/BlueSense.pdf>.
- [5] Nathan Eagle, Alex (Sandy) Pentland. Reality mining: sensing complex social systems. *Personal and Ubiquitous Computing*, 10(4):255–268, 2006.
- [6] Ryo Nishide, Tatsuya Ushikoshi, etc. Extracting Social Contexts from Bluetooth Device Logs. In *Adjunct Proc. UbiComp09*. ACM, 2009.