Location Determination based on WiFi Infrastructure

Yiming Lin, Daokun Jiang, Roberto Yus, Sharad Mehrotra, Nalini Venkatasubramanian
Department of Computer Science
University of California, Irvine
{yiminl18,daokunj,ryuspeir}@uci.edu, {sharad,nalini}@ics.uci.edu

ABSTRACT

PVLDB Reference Format:
Ben Trovato, G. K. M. Tobin, Lars Thørväld, Lawrence P. Leipuner, Sean Fogarty, Charles Palmer, John Smith, Julius P. Kumquat, and Ahmet Sacan. A Sample Proceedings of the VLDB Endowment Paper in LaTeX Format. *PVLDB*, 12(xxx): xxxx-yyyy, 2019.
DOI: https://doi.org/10.14778/xxxxxxx.xxxxxxx

1. INTRODUCTION

The pervasiveness of WiFi infrastructure enables a large number of services based on it. In this paper we focus on Location Based Service (LBS). LBSs use the location of a device (e.g., a smart-phone) and hence of the owner to offer personalized information. Thousands of interesting applications have been built based on such localization techniques. [1] However, WiFi data is far from enough to provide location information with high quality if we use it directly.

We conclude WiFi connectivity data as dirty for LBS in the following two aspects. First, there exist some time interval where no connectivities are captured by WiFi APs. We need to identify if the device is out of building or disconnected but in building, which is formulated as xxx. [1]; Second, it is hard to generate fine-grained location information, i.e., room-level location. Considering a typical scenario, when a device is connected to some WiFi access point (AP), we can conclude that the device is in the region covered by that AP, but no idea of which exact room as a region often covers multiple rooms. The goal is find the exact room among candidates, which could be viewed as a *value disambiguation* problem. As an illustrating example, WiFi connectivity data comes in the format of {*Mac Address, WiFi AP, Timestamp*}. As shown in Table 1, for device with Mac Address *5f0b*, it is captured by AP_1 at "14:16:12" and "14:25:36 ", but we do not know the location information between them. Another device with Mac address *258e* connected to WiFi AP *2* at time "*2018-03-14 12:03:35*". Now we can draw that this device must be in one of the rooms covered by AP_2, which are {2002, 2004, 2014, 2012, 2008, 2066, 2051, 2059, 2052, 2054, 2056, 2058} (From Fig 1) Our goal is to decide the exact room the device is in.

The cleaning of dirty WiFi data poses several challenges. First, [2]; Second, value disambiguation for room-level location is very tricky only based on WiFi connectivity data as a-priori knowledge is limited. Third, since WiFi data comes on streams, the technique has to be very efficient to support online location determination; Finally, the number of devices captured by WiFi APs within a builiding could be large, so that we need to address scalability issue very carefully to be applicable in large systems.

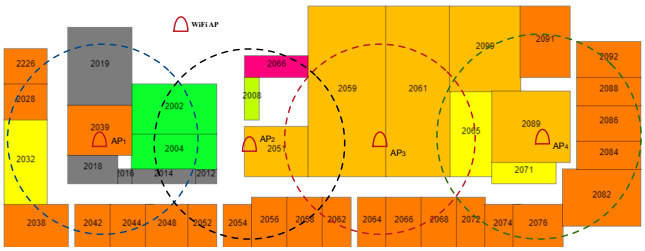

Figure 1: Example of a space with different types of rooms and four WiFi APs.

Table 1: Sample connectivity data from our experiments.

MAC Address	WiFi AP	Timestamp
258e...	2	2018-03-14 12:03:35
258e...	3	2018-03-14 14:15:53
5f0b...	1	2018-03-14 14:16:12
5f0b...	1	2018-03-14 14:25:36

The rest of paper is structured as follows. Section 2 defines problems formally. Section 3 and 4 describes the techniques to clean WiFi data in building, region and room levels. Scalability issues are discussed in Section 5. Section 6 evaluates our methods, Section 7 presents the related work and Section 8 concludes paper.

2. PROBLEM DEFINITION

3. ROOM DETERMINATION

[1]daokun: to summary dirty data from the aspect of gaps

[2]daokun: describe challenges in your part

Proceedings of the VLDB Endowment, Vol. 12, No. xxx
ISSN 2150-8097.
DOI: https://doi.org/10.14778/xxxxxxx.xxxxxxx

4. BUILDING DETERMINATION

5. SCALABILITY

6. EVALUATION

7. RELATED WORK

8. CONCLUSION

9. REFERENCES

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE communications magazine*, vol. 40, no. 8, pp. 102–114, 2002.