User data selection scheme to reduce database update errors for fingerprint localization

Yusuke Miyamoto¹, Satoru Aikawa¹, a), and Shinichiro Yamamoto¹
¹ Graduate School of Engineering, University of Hyogo, 2167 Shosha, Himeji-city, Hyogo 671–2280, Japan
a) aikawa@eng.u-hyogo.ac.jp

Abstract: We are investigating the fingerprint method using wireless LAN access points (APs) to estimate user location. We focused on database update as a way to prevent the loss of localization accuracy. Updating a database using user-measured data presents the issue of update errors. In this paper, we have proposed and verified this issue.

Keywords: fingerprint, indoor localization, MSE, database update

Classification: Navigation, Guidance and Control Systems

References

[1] C.M. Takenga and K. Kyamakya, “Robust positioning system based on fingerprint approach,” Proceedings of the 5th ACM International Workshop on Mobility Management and Wireless Access, pp. 1–8, 2007. DOI: 10.1145/1298091.1298093
[2] S. Aikawa, Y. Shigematsu, R. Baba, and T. Kominami, “Localization estimation system using wireless LAN,” General Assembly and Scientific Symposium of the International Union of Radio Science, Beijing, Aug. 2014. DOI: 10.1109/URSIGASS.2014.6929349
[3] C. Song and J. Wang, “WLAN fingerprint indoor positioning strategy based on implicit crowdsourcing and semi-supervised learning,” ISPRS Int. J. Geo-Inf., vol. 6, no. 11, 356, 2017. DOI: 10.3390/ijgi6110356
[4] S. He, W. Lin, and S.-H.G. Chan, “Indoor localization and automatic fingerprint update with altered AP signals,” IEEE Trans. Mobile Comput., vol. 16, no. 7, pp. 1897–1910, July 2017. DOI: 10.1109/TMC.2016.2608946
[5] X. Liu, J. Cen, Y. Zhan, and C. Tang, “An adaptive fingerprint database updating method for room localization,” IEEE Access, vol. 7, pp. 42626–42638, 2019. DOI: 10.1109/access.2019.2908000
[6] S. He, B. Ji, and S.-H.G. Chan, “Chameleon: survey-free updating of a fingerprint database for indoor localization,” IEEE Pervasive Comput., vol. 15, no. 4, pp. 66–75, Oct.–Dec. 2016. DOI: 10.1109/mprv.2016.69
[7] C. Wu, Z. Yang, and C. Xiao, “Automatic radio map adaptation for indoor localization using smartphones,” IEEE Trans. Mobile Comput., vol. 17, no. 3, pp. 517–528, March 2018. DOI: 10.1109/TMC.2017.2737004
[8] K. Li, J. Chen, B. Yu, Z. Shen, C. Li, and S. He, “Supreme: fine-grained radio map reconstruction via spatial-temporal fusion network,” ACM/IEEE IPSN, April
1 Introduction

We are studying indoor localization services using wireless LAN access points (APs) by using fingerprinting method [1, 2, 3]. The fingerprint method is able to estimate user location even if the location of the AP is unknown. Therefore, this service can be operated in offices and shopping malls with existing APs.

In order to improve the accuracy of localization, this service requires the database update [4, 5]. The method of updating using user-measured data is less expensive. However, the user location is unclear, and update errors are a major issue. S. He et al. use estimated location for a subset sampled from the DB, detecting altered APs [6]. C. Wu et al. automatically collect data with mobile devices whose location is tracked by dead-reckoning, and update the database [7]. These methods make the user location clear, but the computational cost is high. K. Li et al. use the data collected by GPS-located mobile devices to update the database [8]. However, GPS is not suitable for indoor environments. In this study, we propose and validate the methods of the difference of MSE (DMSE). The DMSE can be obtained from the computational process of localization at low computational cost. This method selects correctly estimated user-measured data and uses only that data for updating. This is discussed in detail in Chapter 3.

2 Indoor fingerprint localization

2.1 Fingerprint localization

In general, the fingerprint method is used to estimate the location of each coordinate. In this study, on the other hand, we perform localization for each area. We first divide the facilities to be estimated into multiple areas. The area is a room in the case of a building and a store in the case of a shopping mall. Next, the received signal strength indicator (RSSI) is measured in each area, a database (DB) is created, and it is compared with the user’s measurement data (UD). The area with the smallest difference is used as the localization result. To compare UD and DB, the Mean Square Error (MSE) of RSSI is used. The MSE is calculated by Eq. (1).

\[
MSE_a = \frac{1}{|UD \cap DB_a|} \sum_{i=1}^{\min(|UD|,|DB_a|)} (UD(AP_i) - DB_a(AP_i))^2
\]

In this formula, \( UD, DB_a \) is the AP set of UD, DB in area \( a \). The \( i \) is an integer with \( 1 \leq i \leq |UD \cap DB_a| \), and \( UD(\cdot) \) and \( DB_a(\cdot) \) are the RSSI (or its arithmetic mean if there are multiple RSSI). \( MSE_a \) represents MSE, which is the comparison result of \( UD, DB_a \).

When using MSE, the localization error can be large. Because the MSE becomes small when the number of AP observed in common between DB and UD is small and their RSSI values are close. In order to prevent this, we use APCR (AP Coincidence Rate), which is defined by Eq. (2). And if it is smaller than the predetermined minimum APCR value (APCRth), the MSE is not computed, nor is it estimated to
the corresponding coordinates.

\[ APCR_a = \frac{|UD \cap DB_a|}{|UD \cup DB_a|} \]  \hspace{1cm} (2)

### 2.2 DB update

In the fingerprinting method, the localization accuracy may decrease if old DBs are used. This is because the radio wave environment changes with the disappearance/appearance of AP or the relocation of fixtures with the passage of time. Therefore, it is necessary to update the DB. One way to update the DB is to re-measure the RSSI, but this is costly. The method using UD is less expensive and the data is fresh. An estimate is used for the user location, where the UD was measured. Then, the RSSI is updated as an average value. However, it is difficult to identify the exact user location and there is an issue of update errors. A proposal for the issue of updating the DB is described in detail in Section 3.

### 3 UD selection scheme to reduce DB update errors

#### 3.1 Principle of UD selection

A estimated coordinate is one of the predetermined DB coordinates. However, the user location exists around the estimated coordinate. If this positional deviation is far from the estimated coordinate, the reliability of the localization result is expected to be low.

The “Estimated coordinate” shown in Fig. 1 is one of the DB coordinates. Let’s consider the estimated coordinate in the center. If the results are correct, then the user position should be in the “Zone of correct localization”. Then, in case the user location is far from the estimated coordinate, the localization results are more likely to be wrong. Conversely, in case the user location is in the “Zone of using update”, which is close to the estimated coordinates, the localization results are more reliable.
Therefore, we think that selecting the UD to be used for DB update can reduce the number of DB update errors.

### 3.2 Proposal for UD selection by APCR and DMSE

In this paper, we propose a method for selecting UD to avoid UD update errors. In order to eliminate DB update errors, it is necessary to select UD that is associated with the correct user locations, and APCR and DMSE can be used for this purpose. We proposed a method for combining these two methods. The method is based on two criteria, namely, the APCR-based sorting and then the DMSE-based sorting. The DMSE is the value calculated as the difference between the smallest and the second-lowest MSE among the MSEs calculated for each DB.

First, the APCR selection is done by setting the APCRth to narrow down the estimated area as described in II, and if the results are obtained, we move on to the next step. Next, for DMSE sorting, we set the minimum value of the DMSE (DMSEth) and leave the UD in the case of a DMSE greater than this. As a result, the localization results of the UD is considered as the measurement area of the UD, and is used for updating the DB. The DMSE is obtained by Eq. (3).

\[
DMSE = MSE_{2nd} - MSE_{\text{min}}
\]

![Fig. 2. Verification environment](image-url)
3.3 Verification
The verification environment is shown in Fig. 2. APs were used in the verification facility. A Nexus 9 was used as the measurement terminal, and it was held at

![Graph A](image1)

(a) UD Selection by APCR

![Graph B](image2)

(b) UD Selection by DMSE

![Graph C](image3)

(c) UD Selection by APCR and DMSE

Fig. 3. Principle of UD selection
the height of human chest. Six rooms were used for the area in the verification environment. The number of measurement coordinates per area varies with the size, with a minimum of 11 coordinates and a maximum of 28 coordinates. The spacing between the coordinates was about 1 m, and 20 RSSI measurements were made per coordinate.

We measure the RSSI under the conditions described in Section II and verify the proposal in Section 3.2. Fig. 3 shows the result of UD selection. (a) is APCR only, (b) is DMSE only, and (c) is APCR and DMSE with $APCR_{th}$ is 0.4. There are two curves in each figure. The UD utilization is the percentage of data left over from the UD selection process, and is the percentage of UD that remains in the final 120 UD. The Update accuracy is the percentage of UD that can be correctly estimated for the measurement area of the UD, and is the percentage of UD that is correctly estimated among the remaining UD.

From (a), the Update Accuracy did not reach 1 for APCR only selection, but it was highest when $APCR_{th}$ was set to 0.4. From (b), the Update Accuracy was 1 when the $DMSE_{th}$ was over 30, and the UD Utilization was the highest when the $DMSE_{th}$ was set to 30 for the DMSE only selection. (c) is the result of sorting by DMSE after the $APCR_{th}$ was set to 0.4. From (c), the Update Accuracy was set to 1 when the $DMSE_{th}$ was over 60 and the UD Utilization was highest when the $DMSE_{th}$ was set to 60. Comparing (b) and (c), the results for (b) with $DMSE_{th} = 30$ and (c) with $DMSE_{th} = 60$ are almost identical. However, when the same $DMSE_{th} = 60$ was compared, there was a difference of about 20% in UD Utilization. This shows that APCR and DMSE have a synergistic effect on UD selection. In this verification environment, $APCR_{th} = 0.4$, $DMSE_{th} = 60$ is recommended.

4 Conclusion

Indoor fingerprint localization services requires the database update. We proposed and verified a method to reduce update errors, which is an issue in database updating. This method selects correctly estimated user-measured data and uses only that data for updating.

We identified how APCR and DMSE were designed by conducting experiments. UD selection in APCR was effective in reducing database update errors. UD selection in DMSE was similarly effective in reducing update errors, and we were able to select only those UD that were estimated correctly. By combining APCR and DMSE, we were able to increase the amount of data used for updates while eliminating update errors. Therefore, it can be concluded that APCR and DMSE are synergistic effect on UD selection.