Analysis of Location Estimation Algorithms for Wifi Fingerprint-based Indoor Localization

Joo-Yub Lee, Cheal-Hwan Yoon, Hyunjae Park, and Jungmin So

Dept. of Computer Engineering,
Hallym University, Gangwon-do, Korea
jso1@hallym.ac.kr

Abstract. In Wifi fingerprint-based indoor localization, a well-known method of estimating user's location is to find the nearest reference point using Euclidean distance in signal space. However, this paper shows that Euclidean distance is prone to error, and propose a new algorithm for selecting the nearest neighbor which penalizes signals from unstable access points and compensates for RSSI shifts due to various reasons. Experiments with real measurements show that the new algorithm reduces mean error distance compared to the Euclidean distance method.

Keywords: indoor localization, Wifi fingerprint, location estimation.

1 Introduction

Wifi-based indoor localization has gained interest recently, because building are getting equipped with Wifi access points for connectivity [3]. Using these access points as location indicators removes the need for additional infrastructure cost. A widely used technique is called the fingerprint-based localization [1, 2]. When using the fingerprint-based method, a commonly used method of estimating user location is to find the nearest reference point, using the Euclidean distance in signal space. The problem with Euclidean distance method is that it is prone to error, especially when Wifi access points are unstable: Some access points may be active when generating radio map and not active when the user estimates her location. The situation could be the other way around. Even if an access point is active, the user may not receive signal due to problems such as collisions. When the signal strengths from unstable access points are fed into the standard estimation algorithm, they can lead to serious estimation errors. Also, an observation is made that RSSIs from access points can be shifted due to various reasons such as user height and way of holding the smartphones. Thus, we need a location estimation algorithm that penalizes unstable access points, and compensates for RSSI shifts. In this paper, a new estimation algorithm is designed under these criteria, and is shown to provide higher accuracy compared to the original Euclidean distance-based algorithm.

2.1 Initial experiment

The experiments took place in the Engineering building at Hallym University. 36 reference points (marked as the black dots in the figure) were selected along the corridor, approximately 1.5 meters apart. We have used an HTC Magic smartphone for radio map generation as well as location estimation. In the experiment region, signals were received from a total of 26 access points.

To generate a radio map, we have recorded RSSI values from each access point, at every reference point. If no signal is received from a particular AP, it is treated as -99dBm, which is a value lower than minimum observed RSSI. When estimating user location, the Euclidean signal distance from the test location is calculated for each reference point as follows:

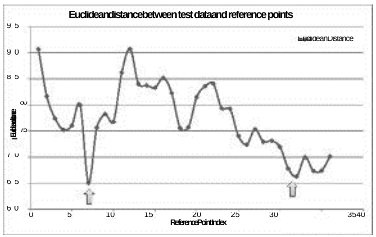
$$l(i) = \sqrt{(4i) - t_1)^2 + (s^{(i)}_2 - t_2)^2 + \ldots + (s^{(i)}} \quad (1)$$

where $s^{(i)}_j$ is the RSSI of access point j at reference point i, and t_j is the RSSI of access point j at the test location. After that, the reference point which has the minimum Euclidean signal distance is selected as the estimated location.

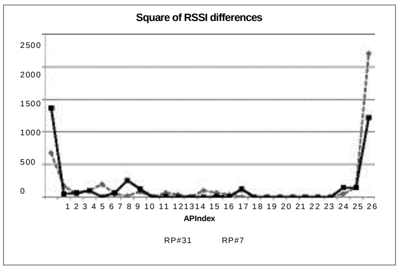
2.2 Analysis of the result

The result of the initial experiment is the line shown in Figure 2a, which is labeled “Euclidean”. The average error distance is 5.50m, meaning that for half of the test locations the estimation algorithm returned with an error larger than 3 meters. The 80% error was 10.6 meters, and 90% error was 13.6 meters. The error amount is unacceptable for most indoor location based service applications.

To find if there is any problem in the estimation algorithm, we analyze a particular test location which showed a large estimation error. Figure 1a shows the Euclidean distance between test data and 36 reference points. The X-axis is the index of reference points. The test location was somewhere between reference point #31 and #32, but the estimated location is #7, which results in a location error of 18.97 meters. While other reference points near RP #7 shows much longer distance from the test location, RP #7 strangely shows short distance. Why is that? Figure 1b reveals a clue to answering the question. In the figure, RSSI differences of AP #1 and AP #26 dominates the result. At test location, the RSSI of AP #26 is -52 dBm which is very high. The RSSI of AP #26 is -87dBm amd -99dBm (no signal), for RP #31 and RP #7, respectively. Because of their large difference, their square value dominates the computation of the Euclidean distance. While the RSSI of AP #26 is not a good evidence that the test location is near RP #7, it is exaggerated in the Euclidean distance computation, resulting in large errors.



(a)



(b)

Fig. 1: (a) Analysis of a particular case. The actual location is somewhere between RP #31 and #32, but the estimation algorithm returns RP #7 as the closest neighbor. (b) Score of each AP for RP #31 and #7.

3 Proposed Scheme and Evaluation

We propose new schemes to find the nearest reference point using Wifi signal strengths. The idea is to make the algorithm so that the impact of large differences in RSSIs is reduced. There are two approaches that serve this goal. The first scheme (Scheme-1) is to filter out large differences so that if RSSI difference is larger than a certain threshold, the distance measure is no longer increased. The distance measure using this scheme can be written as follows.

$$l^{(i)} = \sqrt{c_1^{(i)2} + c_2^{(i)2} + \ldots + c_k^{(i)2}}$$

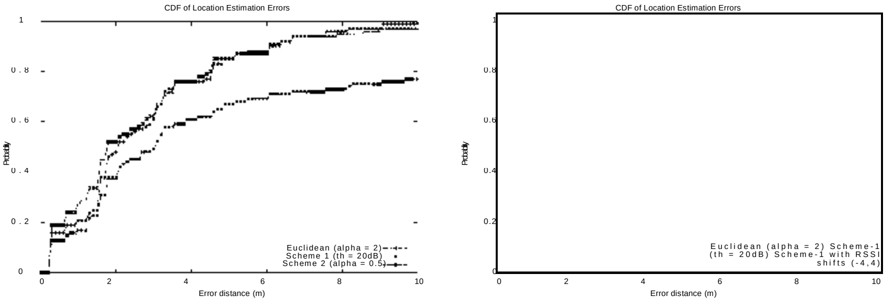
$$c_j^{(i)} = \begin{cases} |s_j^{(i)} - t_j| & \text{if } |s_j^{(i)} - t_j| \le D_{th} \\ D_{th} & \text{if } |s_j^{(i)} - t_j| > D_{th} \end{cases} \quad (2)$$

The second scheme (Scheme-2) is to use a power coefficient that is less than 1, instead of 2 as in Euclidean distance. This will reduce the impact of large RSSI differences, and increase the importance of exact match. The scheme can be written as follows.

$$l^{(i)} = c_1^{(i)\,a} + c_2^{(i)\,a} + \ldots + c_k^{(i)\,a}, \quad 0 < a < 1 \quad (2)$$

The performance results for these two schemes are shown in Figure 2a. As shown in the figure, the two schemes outperform Euclidean distance-based scheme significantly, whereas their performances are comparable.

During the case analysis, we observed that the signal strength pattern is similar between the test location and its nearest reference point, but the RSSIs are shifted. This is expected to be caused from various reasons, such as user orientation, user height, how the user holds the smartphone, etc. The RSSI shift leads to wrong location estimation, and sometimes results in very high error distance. In order to compensate for this effect, we shift the RSSIs of test data between a certain range, and choose the minimum signal distance in this range.



(a) (b)

Fig. 2: (a) CDF of error distance for Scheme-1 and Scheme-2. (b) CDF of error distance for Scheme-1 with RSSI shift mechanism.

Specifically, we compute the signal distance as follows.

$$l_{(i)} = \sqrt{(c_1^{(i)} + d)^2 + (c_2^{(i)} + d)^2 + \ldots + (c_k^{(i)} + d)^2}, \quad d_{min} \leq d \leq d_{max} \quad (4)$$

The performance of RSSI shift mechanism is shown in Figure 2b. The average error distance is reduced by 15% using this mechanism. With Scheme-1 and RSSI shift mechanism combined together, the average error distance is reduced by 57%, and 90th-percentile error distance is reduced by 64%.

4 Conclusion

In indoor localization system, the location estimation algorithm should not only produce low-error answers, but also be resilient to real-life events such as unstable access points. The Euclidean distance-based nearest neighbor selection method is prone to error when unstable access points are present. Also, RSSI shifts can occur due to various reasons such as how users hold the smartphone. We propose a location estimation algorithm that penalizes unstable access points, as well as compensates for RSSI shifts. The performance evaluation from real measurements show that the proposed algorithm reduces localization errors compared to the original Euclidean distance-based algorithm.

References

1. P. Bahl and V.N. Padmanabhan: RADAR: An in-building RF-based user location and tracking system. Proc. of IEEE INFOCOM, pp. 775–784, 2000.
2. M. Youssef and A. Agrawala: The Horus location determination system, Springer Wireless Networks, 2008.
3. H. Liu, H. Darabi, P. Banerjee, and J. Liu: Survey of Wireless Indoor Positioning Techniques and Systems. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, vol. 37, no. 6, pp. 1067–1080, Nov. 2007.