Smartphone Sensor Calibration for Indoor Localization (PD1-15)

TSUI, Ka Hing, YAN, Shing Ho
Prof. Palomar, Daniel

Project Overview

1. Background

What is indoor localization?
- The key of indoor localization is to provide accurate and fast localization performance.
- The common technique is using fingerprinting.

2. Problem statement

Why we need Clustering?
- To reduce a large number of access points in radio map (fingerprint database), they consume much time in computing the whole map during the location determination phase.
- Clustering is a method to group the locations that have common features.
- Clustering is used to reduce the size of the search space and the computation time of estimating current area.

Why we need Calibration?
- Due to the heterogeneity devices, the RSS fingerprint database is applicable to the testing device only.
- It is impossible because of the cost and the large number of smartphones. Therefore, Calibration is essential.
- Calibration is used to reduce the effects of “heterogeneity device” during the indoor calibration.

3. System Blocks Diagram

<table>
<thead>
<tr>
<th>Offline</th>
<th>Online</th>
</tr>
</thead>
</table>
| Clustering
  (1) Joint Clustering technique
  (2) K-means Clustering |
| Testing Sample |
| Clustering Simulation |
| Calibration
  (1) Signal Tendency Index
  (2) Linear Transformation
  (3) Calibration Free |
| Result |

Methodology

1. Clustering

(a) K-means Clustering Technique

K-means Clustering Algorithm is used the joint probability distribution that the number of access points occurred in one location. There is a parameter k as a control variable: (1) $k$ is the largest probability value of access points in one location and with the range from 1 to the size of the whole set of access points in training. For each location contains the k, 1 vector and K-means use these vector of all locations to form the clusters. In the online location determination, we used Bayes theorem to estimate the testing sample in which cluster belong.

(b) Joint Clustering Technique

This approach used the joint probability distributions of the mean of the signal strength of different access points to form clusters. There are two parameters used in this algorithm: (1) $k$ is the number of location means of signal strength value of access points varies in one location, and (2) $q$ is the number of access points that all the locations are sharing a access points in one cluster, while $q < k$.

2. Calibration

We implemented some common Calibration algorithms to deal with heterogeneous devices. They can be categorized into Probabilistic approach and Normalizing approach.

(a) Probabilistic Approach

We found that although the problem leads to signal difference, it has a linear relationship between the devices. Hence, linear least square is applied to compute the numerical relationship and transform the RSS to the testing device. Then we use naïve Bayes classifier to determine the most likely reference point.

(b) Normalizing Approach

We also normalized the RSS vector to minimize the effect. By applying different strategies of normalization: Standardized Signal Tendency Index (STI), Difference of Signal Strength (DSS), and Hyperbolic (HL). We can apply correlation calculation to estimate the most similar reference point.

Simulation Results

1. Clustering

1.1 K-means Clustering technique (Left)

We tested the k from 1 to the size of whole set of access points in training. The result shows that the probability of error that the testing sample belongs to wrong cluster is smaller when k is increasing. The best result is 4.34% when k = 13.

1.2 Joint Clustering technique (Right)

We changed the k from 3 to 7 and q from 3 to 6. The result shows that the probability of error will decrease when the parameter $q$ and $k$ are increased. The best result is 7.9% when $k = 6$ and $q = 6$.

2. Calibration

By testing all the algorithms that we implemented, the maximum accuracy can be computed easily. As the IDDF and KLF shows a poor performance. Hence, we do not show the result.

The testing device: Samsung MEGA

The training devices: Samsung Galaxy Tab, Samsung MEGA, LG

Conclusion

From the empirical experiment, we achieved our final goal that the Clustering and Calibration are able to significantly reduce the search space and the effect of heterogeneous devices.