AP Selection Algorithm in WLAN Indoor Localization

Yingjun Zhou, Xihan Chen, Shulei Zeng, Jingzhi Liu and Dong Liang
Education Ministry Key Laboratory of Universal Wireless Communication,
Beijing University of Posts and Telecommunications,
Beijing, 100876, China

Abstract: In recent years, indoor localization which is based on fingerprint has become more and more popular. With time passing by, indoor localization will be used on mobile equipments. However, because of the universal use of WLAN, numerous APs will be detected when doing indoor localization. More APs means more energy, time and storage to finish localization, but there are always limitations on mobile equipments. Besides, non-ideal APs will decrease the accuracy for indoor localization. So, how to choose the best combination of APs has become a crucial problem to solve. There are two classical AP selection algorithms called MaxMean and InfoGain. MaxMean algorithm is the earliest one, but it’s difficult to combine it with localization algorithms. InfoGain algorithm has better performance than MaxMean, which means it cooperates well with localization algorithms and has higher localization accuracy. But it also has disadvantages such as not considering the correlation between two different APs which may use similar APs. This causes more calculation and reduced accuracy. This paper proposed two AP selection algorithms which are based on InfoGain algorithm but considering correlation between different APs. Many experiments have been done to verify that these two algorithms can not only decrease the calculation amount but also increase the localization accuracy.

Key words: Indoor localization, AP selection, CorMin algorithm, jointChoice algorithm

INTRODUCTION

In recent years, fingerprint localization is a burgeoning way in the area of indoor localization. Fingerprint localization consists of two phase, offline data collecting and online localization. The major work is collecting the Received Signal Strength (RSS) from every Access Point (AP) that every reference point can receive in the localization area to establish the wireless signal radio map during the offline data collecting phase. The major work in the online localization phase is saving the RSS values that the sample point can receive and using the localization algorithm to finish localization. In the wireless environment, mobile equipments can usually detect more than 3 APs. It will need a large amount of mobile equipments’ storage and energy to finish indoor localization. However, there are often limitations on mobile equipments’ storage and energy. Thus how to get an efficient subset of all the APs which can be detected by the mobile equipments has become a crucial problem to solve (Wang and Zhang, 2009).

This essay proposed two improved AP selection algorithms based on information entropy which can select the best AP subset from a large number of APs efficiently to improve localization accuracy and decrease the amount of calculation in the offline phase.

PROBLEM STATEMENT

On the one hand, AP selection can filter APs which may influence the accuracy of localization or help a little in improving the accuracy of localization. On the other hand, it can decrease the amount of calculation to decrease the consumption of energy. All in all, the aim of AP selection is to select a best group of APs which can improve localization accuracy and decrease the amount of calculation at the same time. The easiest way to select AP is make permutation and combination for all the APs to select the best combination of AP; however, this is a really complicated way to achieve. Kushki et al. (2007) has proposed the importance of AP selection in Indoor localization, there are also a few AP selection algorithms have been proposed. Nevertheless, with the development of WLAN, hundreds of APs will be detected at any position and indoor localization will be achieved more on mobile equipments. There’re always limitations on mobile equipments such that online AP selection can’t be achieved because of large consumption of time and

Corresponding Author: Yingjun Zhou, Education Ministry Key Laboratory of Universal Wireless Communication, Beijing University of Posts and Telecommunications, Beijing, 100876, China
storage. Besides, online AP selection may not improve much more accuracy than offline AP selection. Thus, offline AP selection may be more suitable for mobile equipments. Above all, this essay made a research on the existing offline AP selection algorithm and proposed improved AP selection algorithm.

**EXISTING OFFLINE AP SELECTION ALGORITHMS**

**Max mean algorithm:** MaxMean algorithm Fang *et al.* (2008) was first proposed in 2003. The number of APs which can be detected at a fixed location is varying with time according to the feature of wireless signal. Thus, the APs which can be detected most frequently should be chosen. The detected frequency for an AP is proportional to its average RSS, that is to say, an AP with higher average RSS can be detected more frequently. So, in the MaxMean algorithm, it sorts all the APs in the decreasing order of average RSS at a fixed position, select the first k APs to be the input of online localization to finish the process of localization.

**InfoGain algorithm:** Fingerprint localization divides the region into many cells and a cell which is the nearest to the sample point will be chosen to be its estimated position. Thus finding the nearest cell using APs’ RSS is important. K APs with strongest discriminability should be selected to be the members of the best AP group. InfoGain algorithm is based on information gain in information theory to examine an AP’s discriminability according to its average RSS (Deng, 2012). Information gain is the standard to value an AP’s discriminability.

Suppose a localization system based on cells, n represents the number of the cells and m represents the number of APs that can be detected in this localization environment. An AP is considered as a feature and then every cell will have m features. In fact, some AP can’t be detected at some position, thus the absent RSS will be represented as the minimum RSS that can be detected in this environment (Wang and Zhang, 2009). The way to calculate information gain is shown as follows:

\[ H(L) = - \sum_{i=1}^{m} P(L_i) \log P(L_i) \]  
\[ H(L|AP) = - \sum_{i} \sum_{v} P(L_i, AP, v) \log P(L_i | AP, v) \]

Here, \( L \) represents the point’s coordinate in two dimensions, \( v \) represents all the RSS values which are sent from AP, \( c \) represents the number of cells. \( P(L_i) \) represents the prior probability for cell L. If cells are distributed uniformly, then it can be regarded as mean distribution (Wygłinski *et al.*, 2004). For every AP, use equation (1) to calculate its information gain and select the first k AP with the highest information gain.

**PROPOSED ALGORITHMS**

**Summary of existing algorithms:** MaxMean algorithm is the most widely used AP selection algorithm because it's easy to understand. MaxMean algorithm works at the offline phase, certain number of APs will be selected at every sample point to be the result of AP selection. But this algorithm doesn’t cooperate well with localization algorithms when there are too many APs. Such as when using the KNN localization algorithm, the selected AP group may not be the same at both reference and sample point, this will decrease the accuracy for localization.

InfoGain algorithm is another classical AP selection algorithm which is proposed after MaxMean algorithm. InfoGain algorithm also works in the offline phase, it selects a certain number of APs which have the highest information gain to be the input of localization algorithm. This algorithm cooperates well with the localization algorithms, but it also has disadvantages such as not taking the correlation between APs into consideration.

**JointChoice algorithm:** JointChoice algorithm is an improved algorithm based on InfoGain algorithm. In this algorithm, another variable called differentiation is combined with information gain to represent the discriminability of every AP. The correlation between two different APs can be considered using this way. The new variable can be calculated in the following equations:

\[ \text{Dis}(AP) = \text{InfoGain}(AP) + r \ast \text{Diff}(AP) \]  
\[ \text{Diff}(AP) = \sum_{i} \frac{|RSS_i - RSS_{\text{avg}}|}{n} \]

Here, \( n \) represents the number of APs which are in the nearby area of AP, \( (x_i, y_i) \) represents the two dimensional coordinate for AP, \( r \) represents the weight.
between information gain and differentiation. Then, q APs with the highest discriminability will be selected to be the input of localization algorithm.

**CorMin algorithm:** CorMin algorithm is another AP selection algorithm based on InfoGain algorithm. The correlation between different APs is fully considered in this algorithm to get the best combination of APs and delete the useless APs.

There are two steps in the CorMin algorithm:

- Use Eq. 1 to calculate every AP's information gain and select the first p APs with the highest information gain
- Use Eq. 6 to calculate the correlation between different APs. When the correlation of two APs is high, delete the AP with lower information gain. Then select the first q APs with the highest information gain in this new range to be the input of localization:

\[
\rho = \frac{\text{Cor}(X|AP_j, X|AP_i)}{\sigma X|AP_j| \cdot \sigma X|AP_i|}
\]  

**EXPERIMENTS RESULTS**

**Experiment environment:** These experiments run in a 25×28 m laboratory room with two separate rooms in the left-above corner and 3×28 m passage, which is shown in Fig. 1. The number of APs that can be detected in this environment is 192 and the number of sample points is 28.

**Test results:** This section reports on the test results. The experiments focus on the contrast among InfoGain, JointChoice and CorMin algorithms. Figure 2 shows the relationship between average localization error and the value of q. From the figure, the average localization error first decreases then increases with the value of q increases. That is to say, it’s not good to select the group of APs as big as possible. More APs mean more information in ideal environment, however in fact, some APs are easily influenced by the environment. Thus, choosing too much this kind of APs will decrease the accuracy of localization. So, deleting some bad APs may not only increase the localization accuracy but also decrease the amount of calculation. Figure 3 shows that when q = 10, the influence on the average localization error when p equals to different values using CorMin algorithm. Deleting the APs with high correlation can increase the localization accuracy. But there shouldn't be a too large difference between p and q. When p/q=1.5, the minimum average localization error can be achieved. In the real circumstance, to decrease the amount of calculation and ensure the accuracy, p/q=1.2 can be used. When p<16, the average localization error decreases as p increases which implies that deleting APs with high correlation can reject similar information and introduce new useful information. But when p>16, the average localization error decreases as p increases which implies that the information introduced by new APs is less useful than the information provided by the deleted APs.
APs. Under ideal circumstances, selecting more APs represents more useful information, however in the real circumstances, some APs’ stabilities are too bad to decrease the accuracy of localization. In CorMin algorithm, when q stays unchanged, the average localization error first decreases then increases as q increases, that is to say, when p/q equals to certain value, the average localization error reaches the minimum. And this certain value may depend on the number of APs that can be detected in the environment. If the number of APs that can be detected is small, then considering the correlation between different APs is useless. But when the number is big, the procedure is really useful to increase accuracy and decrease the amount of calculation. The JointChoice algorithm is another way to deleting the correlation information. It's easier to achieve than CorMin algorithm. CorMin algorithm needs to include lots of calculation, so when there's a limitation in energy and storage, JointChoice algorithm can be used rather than CorMin algorithm to increase accuracy and decrease the amount of calculation.

REFERENCES


CONCLUSION

From the results, it's easy to find the important meaning of considering the correlation between different