Using Wi-Fi Signal Strength to Localize in Wireless Sensor Networks

Eddie C.L. Chan, George Baciu, S.C. Mak

The Hong Kong Polytechnic University – Hong Kong
{csclchan, csgeorge, csscmak}@comp.polyu.edu.hk

Abstract

Wireless sensor network (WSN) is widely used in many applications such as localization and real-time tracking system. Previous researches commonly suffer the line-of-sight (LOS) problem and dependence on contrast of the background light intensity. Location Fingerprinting (LF) method uses a training dataset of received signal strength (RSS) at different location to track the target. The drawbacks of LF method are needed to have extensive training dataset surveying and highly affected by the changing of internal building infrastructure. In this paper, a sensor-based LF method will be implemented to replace extensive site-surveying. Using a Kalman Filter tracks multiple points to characterize a trajectory. Our experimental result shows that the effectiveness of our method leads to have more accurate and effective tracking system.

1. Introduction

Common approaches to WSN localization such as GPS[1], acoustic [3],[4],[5] and light-based approaches are most effective in relatively open and flat outdoor environments but are much less effective non-line-of-sight (NLOS) environments such as hilly, mountainous or built-up areas. Moreover, most acoustic localization applications not only require the sound source to have a high intensity and to be continuously propagated, they also are limited to localizing only within the area covered by the sound. The drawbacks of light-based localization approaches include their dependence on contrast of the background light intensity [4], [6]. Sequential Monte Carlo (SMC) approaches [12], [13] need to have many sampling, weighting and filtering steps to have updated distribution of sensors. After building an overview of sensors’ distribution, the sensor estimates its location by the weighted average of all samples. It is not effective and has high computational cost in sensor networks.

The task of localization is not limited to WSNs but is also carried out on other types of networks, in particular, on Wi-Fi - IEEE 802.11b. These networks are increasingly ubiquitous in public places, for example, airports, malls, cafes, campuses, and even public squares. They are fuelling a wide range of location-aware computing applications. Currently, Wi-Fi-enabled devices can be located by applying one of two types of location-sensing techniques, propagation-based [2], [8], [9] and location fingerprinting (LF) [1], [8], [10]. Propagation-based techniques measure the received signal strength (RSS), angle of arrival (AOA), or time difference of arrival (TDOA) of received signals and apply mathematical models to determine the location of the device. The drawbacks of propagation-based method are needed to compute every condition that can cause wave signal to blend in order to achieve accurate localization. Location fingerprinting allows a person to locate himself by using a device to access a database containing the fingerprint (i.e., the RSSs and coordinates) of other devices within the Wi-Fi footprint and then calculate its own coordinates by comparing with the LF database. The drawbacks of LF method are needed to have extensive training dataset surveying and highly affected by the changing of internal building infrastructure.

In this paper, a localization approach that makes use of the increasingly ubiquitous Wi-Fi network is implemented with a WSN that estimates the location of sensors using a LF methodology. The approach uses LF-based techniques and sensors in two phases. The first phase detects IEEE 802.11b Wi-Fi signal strength. Then using a set of static location fingerprint sensors collects the location fingerprints to a training database. In the second phase, the location fingerprints are retrieved by the mobile Wi-Fi-enabled device and estimate the location by applying the k-nearest neighbor algorithm to LF training database. Finally, a Kalman Filter is used to track multiple points to characterize a trajectory. This WSN-based localization
approach offers a number of benefits. First, it obviates the need for extensive manual site-surveying. Second, it is potentially suitable for every environment, indoor or outdoor, notwithstanding topography, the presence of man-made structures, or environmental conditions. Finally, it is accurate and cost-effective.

The rest of this paper is organized as follows: Section 2 describes the design of our proposed WSN localization approach. Section 3 presents the localization methodology. Section 4 presents Trajectory accuracy improvement with a Kalman Filter. Section 5 discusses the performance evaluation of location estimation accuracy. Finally, conclusion and future work of the paper are presented in Section 6 respectively.

2. The Design of WSN Localization Approach

First of all, the experiment environment is described. The test bed was established in a laboratory on 7th floor of the PQ building, in Department of Computing, at The Hong Kong Polytechnic University. The layout of the laboratory is shown in Figure 1. The dimension of room is approximately 10m by 4m.

The radio frequency channels of IEEE 802.11b are in the 2.4GHz band. The number of non-overlapping channels for 802.11b is three. The received signal sensitivity also limits the range of the RSS to be between -90 dBm and -30 dBm. Nevertheless, the highest typical value of the RSS is approximately -40 dBm at one meter from any AP. Samples at 33 locations are collected which shown in Figure 1, each data set of samplings consists 20 times of sampling interval. In our case, four access points are distributed in the room. Figure 1 illustrates the position of the access points on the grid. They are placed in (2, 4), (1, 0), (8, 4) and (10, 1) respectively.

3. Localization Methodology

In this section, the algorithms are used in the proposed approach: the signal propagation theorem, K-Nearest Neighbors Fingerprinting Estimation, and Probabilistic Estimation.

3.1 Signal Propagation Theorem

A controlled environment offers no background signal interference. Further, instead of using sound and visible light as medium to track objects, Wi-Fi RSS can be retrieved from every grid point. The radio frequency signal obeys propagation-based theorem [11].

\[ r_j(d_{j,k}) = r_k(d_0) - 10\alpha \log_{10}(d_{j,k}) - \text{wallLoss} \]

where \( r_k \) is the initial RSS at the reference distance of \( d_0 \) is 1 meter, the variable \( \alpha \) denotes the path loss exponent. Under other circumstances, the indoor path loss exponent \( \alpha \) can be between 1 and 6. wallLoss is the sum of the losses introduced by each wall on the line segment drawn at Euclidean distance \( d_{j,k} \).

Figure 1. Test bed environment with 33 grid points

3.2 K-Nearest Neighbors Fingerprinting Estimation

To estimate the positions of sensors, K-Nearest Neighbor (K-NN) algorithm is applied to two sets of data. The first set of data is the samples of RSS from N APs in the area; sampling vector \( r \) is called throughout the paper. This vector is denoted as \( R = [r_1, r_2, r_3...r_N] \), each element in vector is the independent RSS (in dBm) collected from APs in the location. The second set of data is for the location fingerprinting, it contains all of the average RSS from N APs at a particular location. The data forms the LF database. Location fingerprint vector is called throughout the paper. This vector is denoted as \( F = [f_1, f_2, f_3...f_N] \) at the position \( D = [d_1, d_2, d_3...d_N] \).

There are several types of LF algorithm. The fundamental method of positioning algorithm is the K-Nearest Neighbor (K-NN) algorithm. The location \( d \) can be estimated by clustering the Euclidean distance \( |r - f_i| \) between sampling fingerprint vector \( r \) and location fingerprint vector \( f_i \) with position \( d_i \) as follows:

\[ d = \frac{\sum_{i=1}^{N} \frac{d_i}{|r - f_i|}}{\sum_{i=1}^{N} \frac{1}{|r - f_i|}} \]
3.3 Probabilistic Estimation

In this section, probabilistic estimation is described to calculate the probability that picking up the correct fingerprint. This is done using the Euclidean distance. For example, if our sampling vector $r$ and location fingerprint vector $f$ are the same, it means that the correct location can be known exactly. However, the sampling vector may not be the same as the location fingerprint vector. The Euclidean error distance between the sampling and correct location fingerprint is $|r - f_{\text{correct}}|$. There is also a possibility to get the incorrect neighbor location fingerprints from database. The Euclidean error distance between the sampling and incorrect neighbor location fingerprint is $|r - f_{\text{neighbor}}|$. The Euclidean distance error of correct location fingerprint must be smaller or equal to incorrect neighbor location fingerprint, such that (3) always holds.

$$ |r - f_{\text{correct}}| \leq |r - f_{\text{neighbor}}| $$

By evaluating (3),

$$ (r - f_{\text{correct}})^2 \leq (r - f_{\text{neighbor}})^2 $$

$$ (f_{\text{correct}} - f_{\text{neighbor}})(f_{\text{correct}} + f_{\text{neighbor}} - 2r) \leq 0 $$

Recent research [5],[6],[7] support Wi-Fi RSS obeys Gaussian distribution when the sample size of RSS is large. Assume Wi-Fi RSS obeys Gaussian distribution, the mean and variance of Wi-Fi RSS distribution are $\mu_s$ and $\sigma_s$. The squared error also obeys Gaussian distribution. By applying the properties of the sum of multiple independent Gaussian random variables [11], the mean and variance of squared error as follow:

$$ \mu = (f_{\text{correct}} - f_{\text{neighbor}})(f_{\text{correct}} + f_{\text{neighbor}} - 2r) $$

$$ \sigma^2 = 4(f_{\text{correct}} - f_{\text{neighbor}})^2\sigma_s^2 $$

The Euclidean distance is calculated with K-nearest neighbors. Using (4), the new mean and variance are:

$$ \mu_k = \frac{1}{k} \sum_{i=1}^{k} (f_{\text{correct}} - f_i)(f_{\text{correct}} + f_i - 2r) $$

$$ \sigma_k^2 = \frac{4}{k} \sum_{i=1}^{k} (f_{\text{correct}} - f_i)^2\sigma_s^2 $$

Thus, by using (5), the probability of getting the correct location estimation is calculated with the RSS coverage in the building as follows:

$$ \Pr[X] = \Pr[d_i | F] = \frac{1}{\sqrt{2\pi\sigma}} \exp \left[ -\frac{x - \mu_k}{2\sigma_k^2} \right] $$

where $X$ is the event of returning the correct location.

4. Trajectory Estimation with a Kalman Filter

To improve trajectory accuracy, the Kalman Filter [14] is used to estimate the state $x \in \mathbb{R}^2$ of a discrete-time tracking process that is governed by the linear stochastic difference equation. Our time update equations are:

$$ x_k = x_{k-1} $$

$$ P_k = P_{k-1} + Q $$

And our measurement update equations are:

$$ K_k = P_k^{-1}(P_k^{-1} + R)^{-1} $$

$$ \hat{x}_k = x_k + K_k(z_k - \hat{x}_k) $$

$$ P_k = (1 - K_k)P_k $$

where $Q$ is the process noise covariance and $R$ is measurement noise covariance. $P_k$ is the priori estimate error covariance. $\hat{x}_k$ is the posteriori estimate error covariance. $x_k$ is the priori state estimate at step $k$. $\hat{x}_k$ is the posteriori state estimate at step $k$. $K_k$ is the Kalman blending factor.

5. Performance Evaluation

In this section, experiment results are described for the probability of getting correct location fingerprint according to RSS features and the Kalman Filter influence the localization precision.

5.1 Results for the probability of getting the correct location fingerprint

Figure 2 shows the influence on probability of varying the standard deviation $\sigma$. Assuming that four nearest access points are used for calculation, when the standard deviation $\sigma$ increases, the probability of getting correct location fingerprint decreases. Therefore, for the high accuracy of localization, the suggested standard deviation $\sigma$ should be below than 4dBm (In other reported measurement, the $\sigma$ assumed as 2.13dBm [1]), and it needs over 3 neighbors for estimation. However, the standard deviation depends on the real environment, for example, because the human body can absorb the signal strength, so if the environment is in busy traffic state, the standard deviation will be large.
Figure 3 shows the influence on probability of varying path loss exponent $\alpha$. The path loss exponent $\alpha$ can be varied between 1 and 6. When the path loss exponent $\alpha$ increases, the accuracy also increases. Path loss exponent $\alpha$ represents the attenuation rate of the RSS. If the path loss exponent is large, it means that the RSS has a great change by a small shift in a distance. Therefore, high path loss exponent is easy to recognize the location fingerprint and to increase the probability of getting the correct location fingerprint.

Figure 4 shows the influence of varying the number of neighbors on the accuracy probability. In our case, there are 8 neighbors from the 8 directions. Equations (4), (5) and (6) are applied to estimate the probability of getting the correct location fingerprint. If more neighbors are used, then the accuracy increases. This means that the more comparisons between neighbors and mobile sensors, the greater the probability of getting correct location fingerprint. Moreover, the probability not only depends on the number of neighbors, the number of access points is also a factor to affect probability. For example in our case, the probability in comparison with 8 neighbors increases from 76% to 95% approximately, when it increases the number of access points from 1 to 4. As a conclusion, more access points the probability is higher.

5.2 Results for adding the location tracking filter

The K-Nearest Neighbor (K-NN) algorithm is used to estimate the location of the mobile sensor. After the location is calculated, a Kalman filter is added to enhance the accuracy. The experiments in this section observe the trajectory accuracy when a user with mobile sensor walks around the room. Assume that the Gaussian RSS noise measurement is 4.

Figure 5 shows the actual and estimated (both with and without filter) paths of the user. Figure 6 shows the reduction of error when using a Kalman Filter. The use of (7) and (8) slightly improves the trajectory accuracy, reducing the error from 1.16m to 1.10m.

Figure 7 shows error on the X and Y coordinates. The error on the X-axis is greater than on the Y-axis (0.83m and 0.58m respectively). In the testing environment, the width of the room (X-axis) is longer than the length of the room (Y-axis). Therefore, the
distance between mobile sensors and the access points in X-axis is longer than in Y-axis. Because of a longer distance, the probability of blocking signal by furniture is higher.

Figure 6. Reduction of error when using a Kalman Filter

Figure 7. Error distance on the X and Y coordinates

6. Conclusion and Future Work

In this paper, a WSN localization system is established with using a Kalman filter for trajectory tracking. The contribution of the paper is to replace the extensive manual site-surveying and it is workable for every Wi-Fi environment. Localization application based on existing Wi-Fi indoor infrastructure becomes very challenging because of the extensive movement of obstacles and the interference between different frequency channels. Future work will be building a 3D pervasive tracking and a dynamic spatio-temporal filtering technique.

7. References