

The Research of Wi-Fi Indoor Positioning Algorithm based on Position Fingerprint

Yujie Liu^{1,a}, Meng Wu^{1,b}

¹College of Telecommunications & Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China

^amygoodhelper@hotmail.com, ^bwum@njupt.edu.cn

Keywords: Position fingerprint, Indoor positioning, Kalman filter, Probability weighted, Naive Bayes algorithm.

Abstract. To improve the positioning accuracy and robustness of indoor positioning technology, a fingerprint positioning method is proposed based on the Kalman filter and the probability weighted Bayes algorithm. Firstly, the optimal estimation of signal strength is calculated using the prediction and correction model of Kalman filter, and then the position is estimated using the posterior probability weighted Naive Bayes algorithm. The experimental results show that the algorithm based on Kalman filter and the probability weighted Bayes gains pretty good positioning accuracy and robustness.

Introduction

With the rapid development of mobile data services and wireless local area networks (WLAN), the location-based services (LBS) based on Wi-Fi has become the research hotspot of indoor positioning. Wi-Fi indoor positioning technology based on position fingerprint with advantages of low costs, high positioning accuracy and pretty good robustness meets the requirements of most indoor positioning, so it is widely used in actual indoor positioning [1].

Position fingerprint technology is to calculate the position by using the mapping relationship between the received signal strength (RSS) and the physical location of access points (AP) and it contains two phases, offline training and online positioning [2]. In offline training phase, the scheme of position fingerprint acquisition with higher estimation accuracy and less computation is usually adopted to establish the database of position fingerprint [3]. In online positioning phase, the real-time RSS of mobile terminal should be collected and then be matched with the position fingerprint quickly for calculating the positioning trajectory. Currently, the position matching algorithm mainly includes Nearest Neighbor in Signal Space (NNSS) and Naive Bayes. NNSS is the most basic deterministic fingerprint matching algorithm, which is proposed in RADAR system developed by Microsoft [4]. However, due to the RSS's volatility caused by environmental interference, K-NNSS (K-Nearest Neighbor in Signal Space) and weighted K-NNSS are proposed soon after [5]. In 2002, Roos proposed a probability statistical model by improving the positioning method of fingerprint and adopted the Bayes algorithm for the first time in indoor positioning [6]. Subsequently, Horus positioning system based on probability statistical model designed by Youssef et al. Because of the thought of clustering, the accuracy of the system is improved significantly [7]. However, most fingerprint positioning methods based on the Bayes algorithm ignore the contributions of other reference points and the reliability of real-time measurement of RSS, there are obvious limitations for these positioning methods.

In this paper, we propose a positioning algorithm based on Kalman filter and the probability weighted Bayes algorithm. According to the measured RSS of current position and the estimated RSS of previous position, we determine the amount of correction and transfer process using Kalman filter algorithm, and then calculate the optimal estimated RSS of current position to improve reliability of measured RSS in online phase. Naive Bayes algorithm are used to calculate the posterior probability of the optimal estimated RSS sample vector. And we estimate the real position by posterior probability weighted approach.

Kalman Filter and the Probability Weighted Bayes Algorithm

Kalman Filter Algorithm.

The Kalman filters are based on linear dynamic systems discretized in the time domain and hidden Markov model. The basic dynamic system are modelled on a Markov chain built on linear operators perturbed by errors that may include Gaussian noise(normal distribution noise) [8].

Neglecting the system control function, stochastic linear discrete system equations are expressed as follows [9].

$$X_k = \Phi_{k,k-1}X_{k-1} + \Gamma_{k,k-1}W_{k-1} \quad (1)$$

$$Z_k = H_kX_k + V_k \quad (2)$$

Where X_k is the system's n dimensional state sequence, Z_k is the system's m dimensional observation sequence, $\Phi_{k,k-1}$ is the system's $n * n$ dimensional state transition matrix, $\Gamma_{k,k-1}$ is the system's $n * p$ dimensional noise input matrix, H_k is $m * n$ dimensional observation matrix, W_{k-1} is the system's p dimensional process noise sequence, $w_k \sim N(0, Q_k)$, V_k is the system's m dimensional observation noise sequence, $v_k \sim N(0, R_k)$.

Assuming that we have acquired the k times RSS measurement value sequence Z_1, \dots, Z_{k-1}, Z_k of one certain AP and found the optimal linear estimation \hat{X}_{k-1} of X_{k-1} . That is, \hat{X}_{k-1} is the linear function of Z_1, \dots, Z_{k-1} sequence. According to equations (1) and (2), it is evident that we can get the following equations.

$$\hat{X}_{k,k-1} = \Phi_{k,k-1}\hat{X}_{k-1} \quad (3)$$

$$\hat{Z}_{k,k-1} = H_k\hat{X}_{k,k-1} \quad (4)$$

So, the current estimated RSS value \hat{X}_k for this certain AP is expressed as follows.

$$\hat{X}_k = \hat{X}_{k,k-1} + K_k[Z_k - \hat{Z}_{k,k-1}] \quad (5)$$

$$\tilde{X}_k = X_k - \hat{X}_k \quad (6)$$

$$P_k = E[\tilde{X}_k\tilde{X}_k^T] \quad (7)$$

Where \tilde{X}_k is the estimation error of X_k after acquiring the current RSS measurement value Z_k , K_k is the filter gain matrix, P_k is the error variance matrix of Kalman filter. The formulas of gain matrix and variance matrix are given as follows.

$$K_k = P_{k,k-1}H_k^T[H_kP_{k,k-1}H_k^T + R_k]^{-1} \quad (8)$$

$$P_{k,k-1} = \Phi_{k,k-1}P_{k-1}\Phi_{k,k-1}^T + \Gamma_{k,k-1}Q_{k-1}\Gamma_{k,k-1}^T \quad (9)$$

The basic equations of Kalman filter are the recursive form of time domain and the calculation is a process of continuous “prediction and correction”. Also there is no need to store large amounts of data when positioning. Once the latest measurements are observed, the estimated value of state prediction can be corrected by using the prediction error. That is, Kalman filter is propitious to real-time online processing for the reason that it can calculate the current optimal estimated RSS values in time. At the same time, the filter gain matrix K_k is independent of the observed values, which can be calculated in advance in offline phrase. This will reduce the computational quantity of the online positioning phrase automatically.

Probability Weighted Bayes Algorithm.

Traditional Bayes algorithm regards the reference point with maximum posteriori probability $P(l|r)$ as the positioning location. For ignoring the contribution of other reference points to the location, there are obvious limitations for traditional Bayes algorithm. Based on the traditional Bayes positioning algorithm, this paper proposes an improved probability weighted Bayes algorithm. We define the fingerprint database position sample of offline phrase as $L = \{l_1, l_2, \dots, l_n\}$, where n is the amount of the selected reference points, (x_i, y_i) indicates the coordinate position

of reference point l_i . And we define the optimal RSS estimation sample vector which is filtered by Kalman filter in online phrase as $\mathbf{r} = \{r_1, r_2, \dots, r_m\}$. According to maximum likelihood probability, the reference point with maximum posteriori probability $P(l_i|\mathbf{r})$ can be regarded as the user's position.

$$(\hat{x}, \hat{y}) = \max_{l_i} P(l_i|\mathbf{r}) \quad i = 1, 2, \dots, n \quad (10)$$

According to Bias's theorem, the posterior probability can be calculated by a priori probability. And the equation is given, as follows.

$$P(l_i|\mathbf{r}) = \frac{P(\mathbf{r}|l_i) P(l_i)}{P(\mathbf{r})} = \frac{P(\mathbf{r}|l_i) P(l_i)}{\sum_{i=1}^n P(\mathbf{r}|l_i) P(l_i)} \quad (11)$$

Where $P(l_i)$ is the prior probability of user's reference point l_i . Let's assume that user appears in each reference point with same probability. That is, we can take $P(l_i)$ as uniform distribution, and each AP's RSS is independent of the others. The calculation of the posterior probability can be realized in two ways, histogram and kernel function. Because some RSS likelihood probability often equals to zero in histogram, the accuracy of the posterior probability of the reference point and the positioning error may be seriously affected by histogram method.

Since the signal strength distribution of reference point can be fitted as Gauss distribution, we adopt Gaussian kernel as kernel function and take sum of Gaussian kernel function as likelihood function.

$$P(r_k|l_i) = \frac{1}{s} \sum_{j=1}^s K(r_k^j, r_k^{l_i}) \quad (12)$$

$$K(r_k^j, r_k^{l_i}) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(r_k^j - r_k^{l_i})^2}{2\sigma^2}} \quad (13)$$

Where $K(r_k^j, r_k^{l_i})$ is Gaussian kernel function, parameter σ which determines the envelope shape of likelihood function describes the window width of the kernel function [10]. r_k^j represents the j -th collected RSS data in online phrase for k -th AP. $r_k^{l_i}$ represents the offline reference RSS data of k -th AP for reference position l_i .

The reference point's signal strength between APs is independent from each other, so the single-dimensional likelihood function model can be extended into multi-dimensional model. Joint probability distribution estimation can be converted into marginal probability distribution estimation.

$$P(\mathbf{r}|l_i) = P(r_1, r_2, \dots, r_m|l_i)P(l_i) = P(r_1|l_i)P(r_2|l_i) \dots P(r_m|l_i)P(l_i) \quad (14)$$

As mentioned, traditional Bayes algorithm only take the reference point with maximum posteriori probability into consideration. Since ignoring the contribution of other reference points to the location, the positioning error is multiplied by the separation distance of reference point. With the increasing of the reference point's spacing, the positioning error can be rapidly deteriorated. In this paper, we calculate the position coordinates based on weight sum of posterior probability (top K) using the analogy of weighted K-nearest neighbor method. With the weight sum of posterior probability, our method has more higher accuracy than weighted K-nearest neighbor method measuring the Euclidean distance based on signal strength. And then the more accurate positioning coordinates $l(\hat{x}, \hat{y})$ is estimated by calculating the posterior probability of top K reference nodes and normalizing probability, where $l(\hat{x}, \hat{y})$ is shown in equation (15).

$$l(\hat{x}, \hat{y}) = \frac{\sum_{i=1}^K [P(\mathbf{r}|l_i) \cdot l_i]}{\sum_{i=1}^K P(\mathbf{r}|l_i)} \quad (15)$$

Experiment and Result

Indoor Experiment Process.

The laboratory with intensive public access points is selected as our experimental environment,

which is regarded as the rectangular area of 42m*18m. In offline phrase, to test the sensitivity to the reference point interval of the improved weighted Bayes algorithm, we select different reference point intervals (2m and 4m). In online phrase, we use HTC ONE XC (Android 4.0, Linux Kernel 3.0) cellphone and set the RSS measurement period $T = 500ms$. And the Naive Bayes positioning algorithm (NB) and the Kalman filter and the probability weighted Bayes algorithm (KWB) were adopted respectively.

Experiment Result and Analysis.

Position Error Mean (MPE) is used to measure the accuracy of the measurement. Equation is shown as follows.

$$MPE = \frac{1}{N} \sum_{i=1}^N \sqrt{(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2} \quad (16)$$

Where N is positioning times, (\hat{x}_i, \hat{y}_i) and (x_i, y_i) respectively represent the i -th estimate position coordinates and real position coordinates.

Error Standard Deviation Position (PESD) is used to measure the robustness of the measurement. Equation is shown as follows.

$$PESD = \sqrt{\frac{1}{N} \sum_{i=1}^N (PE_i - MPE)^2} \quad (17)$$

Where $PE_i = \sqrt{(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2}$ is the i -th Position Error (PE).

Table 1 shows the MPE and PESD at 2m and 4m intervals for NB algorithm and KWB algorithm. It can be directly seen that our proposed algorithm is better than Naive Bayes algorithm on MPE and PESD, which shows that our algorithm has better positioning accuracy and robustness.

Table 1 MPE and PESD of positioning algorithm

positioning algorithm	MPE/m(2m/4m)	PESD/m(2m/4m)
KWB	2.14/3.74	1.46/2.63
NB	2.92/5.02	1.79/3.56

Figure 1 shows the error accumulation probability of NB algorithm and KWB algorithm. Obviously, KWB is better than NB on the error accumulation probability whether interval is 2m or 4m. For using the posterior probability weighted Bayes algorithm, KWB algorithm is particularly effective at 4m interval and it is proved that KWB is less sensitive to the interval changing of the reference point.

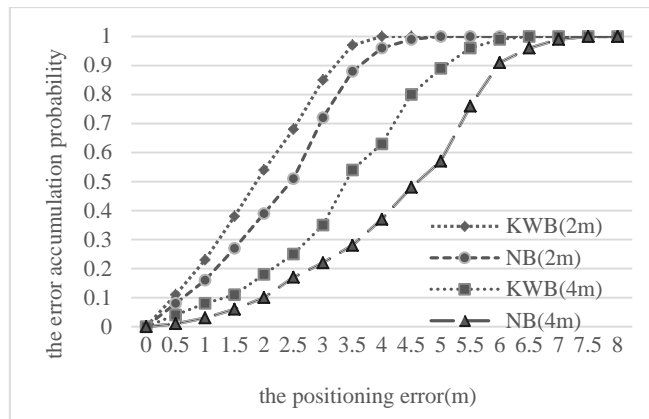


Fig. 1 The error cumulative probability comparison chart

Figure 2 shows the estimated trajectory map of NB algorithm and KWB algorithm (where interval is 2m). Compared with the NB algorithm, KWB algorithm is more consistent with the real motion trajectory.

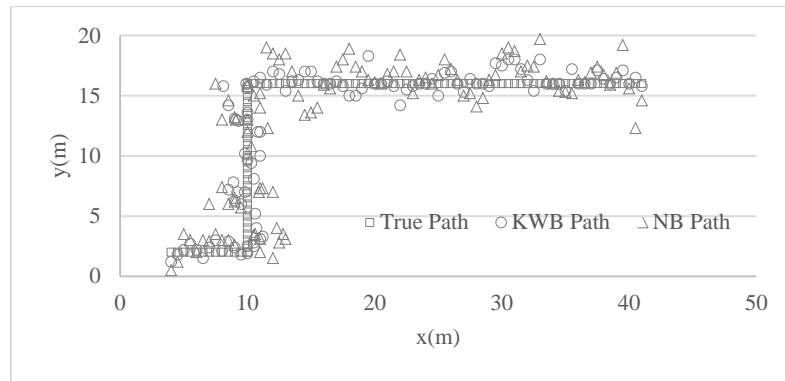


Fig. 2 The trajectory comparison chart

Conclusions

In this paper, we proposed a position fingerprint positioning method based on the Kalman filter and the probability weighted Bayes algorithm. The optimal estimation value of RSS is calculated using Kalman filter algorithm by establishing the relationship among the gain matrix K_k , the error variance matrix $P_{k,k-1}$, the system process noise matrix Q_k and the observation noise matrix R_k . We improve the traditional Naive Bayes algorithm and the positioning location is estimated using the posteriori probability weighted method by analogy with weighted K nearest neighbor method. Compared with the traditional Bayes positioning algorithm, our algorithm has higher precision, stronger applicability and better robustness.

References

- [1] Wirola L, Laine T A, Syrjarinne J. Mass-Market Requirements for Indoor Positioning and Indoor Navigation[C]// 2010 International Conference on Indoor Positioning and Indoor Navigation. Zurich, 2010: 1-7.
- [2] Tagashira S, Kanekiyo Y, Arakawa Y, et al. Collaborative Filtering for Position Estimation Error Correction in WLAN Positioning Systems[J]. IEICE Transactions on Communications. 2011, 94-B(3): 649-657.
- [3] Hossain S, Ariffin S H S, Faisal N, et al. Accuracy Enhancement of Fingerprint Indoor Positioning System[C]// Intelligent Systems, Modelling and Simulation (ISMS), 2012 Third International Conference on. IEEE, 2012: 600-605.
- [4] Bahl P, Padmanabhan V N. RADAR: An in-building RF-based user location and tracking system[C]// INFOCOM 2000, Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE, 2000, 2: 775-784.
- [5] Brunato M, Battiti R. Statistical Learning Theory for Location Fingerprinting in Wireless Lans[J]. Computer Networks, 2005(47) : 825-845.
- [6] Roos T, Myllymäki P, Tirri H, et al. A probabilistic approach to WLAN user location estimation[J]. International Journal of Wireless Information Networks, 2002, 9(3): 155-164.
- [7] Youssef M. Horus: A WLAN-Based Indoor Location Determination System[D]. Ph.D. dissertation, University of Maryland, 2004.
- [8] Wikipedia. Kalman filter [EB/OJ]. [2015-10-12]. https://en.wikipedia.org/wiki/Kalman_filter.
- [9] Fu Mengyin, Deng Zhihong, Yan Liping, et al. Kalman filtering theory and its application in navigation system[M]. 2nd edition. Beijing: Science Press, 2010: 17-21.
- [10] Zhao Fang, Luo Haiyong, Linqun, et al. Node localization algorithm based on kernel function and Markov chains[J]. Journal on Communications, 2010, 31(11): 195-204.