A Comparative Survey of WLAN Location Fingerprinting Methods

Ville HONKAVIRTA, Tommi PERÄLÄ, Simo ALI-LÖYTTY and Robert PICHÉ

Tampere University of Technology, Finland

Email: simo.ali-loytty@tut.fi

Abstract—The term "location fingerprinting" covers a wide variety of methods for determining receiver position using databases of radio signal strength measurements from different sources. In this work we present a survey of location fingerprinting methods, including deterministic and probabilistic methods for static estimation, as well as filtering methods based on Bayesian filter and Kalman filter. We present a unified mathematical formulation of radio map database and location estimation, point out the equivalence of some methods from the literature, and present some new variants. A set of tests in an indoor positioning scenario using WLAN signal strengths is performed to determine the influence of different calibration and location method parameters. In the tests, the probabilistic method with the kernel function approximation of signal strength histograms was the best static positioning method. Moreover, all filters improved the results significantly over the static methods.

I. INTRODUCTION

Location-aware services have become popular with the development of modern communication technology. The increased variety of commercial applications has established the demand for indoor localization services. Weak signal reception and missing line-of-sight between the user and the satellites causes Global Positioning System (GPS) to perform poorly indoors and thus different indoor localization systems have been developed.

Systems such as the Active Badge, the Cricket, the Bat and the Ekahau positioning engine (EPE) rely highly on an infrastructure that is specially designed for indoor localization [1, 2]. However, these kind of purpose-built systems can be expensive and thus hard to implement on a world-wide scale. There are also other dedicated indoor positioning systems, see e.g. survey [3].

Localization systems can exploit different kinds of measurements; systems based on the angle of arrival (AOA), time of arrival (TOA) and the time difference of arrival (TDOA) have been proposed [4]. However, the reliability of these measurements suffers from the complex signal propagation environments [5].

The increased deployment and the popularity of wireless local area networks (WLAN) have opened a new opportunity for location-aware services. Although WLAN was not designed for localization, it can be used for location estimation by exploiting the received signal strength indicator (RSSI) value. RSSI also allows the utilization of the existing infrastructure, because no additional hardware is needed. Signal to noise ratio (SNR) is also available, but it is often omitted because RSSI has stronger correlation with the location than SNR [5].

Location fingerprinting differs from other localization principles. Instead of determining the distances between the user and the transmitting access points (AP) and triangulating the user's location, the location of the user is determined by comparing the obtained RSSI values to a radio map. The radio map is constructed in an offline phase and it contains the measured RSSI patterns at certain locations. This way the characteristics of the signal propagation in indoor environments are captured and the modeling of the complex signal propagation is avoided. However, the offline phase is quite laborious and the radio maps have to be stored in memory.

Many of the existing location fingerprinting methods lack a proper mathematical formulation and theoretical basis. The first purpose of this work is to present the mathematical formulation of the location fingerprinting methods covered in this paper. The second goal is to apply Bayesian and Kalman filters to location fingerprinting. The third objective in this work is to implement the different algorithms and to test them in varying circumstances.

This paper is organized as follows. In Section II, the mathematical formulation of the radio map is presented. In Section III, the location estimation phase is covered from the deterministic (III-A) and the probabilistic (III-B) point of view. In Section IV, the traditional location fingerprinting is extended to computating the location estimates in time series by applying different filters. Different state models are combined with the static location estimation algorithms. The test results are presented in Section V. Section VI summarizes the results and suggests guidelines for the future work.

II. RADIO MAP

The construction of the radio map begins by dividing the area of interest into cells with the help of a floor plan. RSSI values of the radio signals transmitted by APs are collected in calibration points inside the cells for a certain period of time and stored into the radio map. The ith element in the radio map has the form

$$\mathcal{M}_i = \left(B_i, \underbrace{\{\vec{a}_{ij} | j \in N_i\}, \theta_i}_{R_i \in R} \right), \qquad i = 1, \ldots, M, \quad (1)$$

where B_i is the ith cell, whose center p_i is the ith calibration point. Vector $\vec{a}_{ij}$ holds the RSSI values measured from the access point AP_j and its kth element is denoted by a_{ij}^k. The parameter θ_i contains any other information needed in

the location estimation phase. This can be for example the orientation $\theta_i = d_i \in \{\text{north, south, east, west}\}$ of the mobile unit (MU), such as in the RADAR system [5]. We denote the ith fingerprint by R_i and the set of all fingerprints by $R = \{R_1, \ldots, R_M\}$, so the ith element of the radio map is $\mathcal{M}_i = (B_i, R_i)$.

The radio map can be modified or preprocessed before applying it in the location estimation phase. The motivation can be the reduction of the memory requirements of the radio map or the reduction of the computational cost of location estimation. In addition, different location estimation methods use different characteristics of the fingerprint histogram, such as the mean and the variance.

III. LOCATION ESTIMATION

Given the radio map, the objective of the location estimation phase is to infer the state (location) of MU from the received *measurements* vector y which includes RSSI samples y_j from several APs. In some cases we collect several RSSI measurement from the jth AP into vector $\vec{y}_j$ before we compute the state estimate. The mean of these measurements is denoted by $\bar{y}_j$ which is the jth component of vector $\bar{y}$.

A. Deterministic framework

In the deterministic framework, the state x is assumed to be a non-random vector [5]. The main objective is to compute the estimate $\hat{x}$ of the state at every time step. Usually the estimate is a convex combination of the calibration points p_i, that is

$$\hat{x} = \sum_{i=1}^{M} \frac{w_i}{\sum_{j=1}^{M} w_j} p_i, \qquad (2)$$

where all weights w_i are nonnegative. One possible weight w_i is the inverse of the norm of the RSSI innovation [6], that is

$$w_i = \frac{1}{||\bar{y} - \bar{a}_i||}, \qquad (3)$$

where $\bar{y}$ is the measurement vector, $\bar{a}_i$ is the vector of the means of the RSSI values of each AP at the ith calibration point, and the norm $||\cdot||$ is arbitrary. Examples of possible norms are given in Table I. The Euclidean norm (2-norm)

Table I
NORMS

Name	$||x||$ $(x \in \mathbb{R}^{n_x})$
p-norm [5]	$||x||_p = \left(\sum_{i=1}^{n_x} |x_i|^p\right)^{\frac{1}{p}}$
modified p-norm [7]	$||x||_{mp} = \left(\sum_{i=1}^{n_x} \frac{1}{w_i} |x_i|^p\right)^{\frac{1}{p}}$
infinity-norm	$||x||_\infty = \max_i (|x_i|)$
Mahalanobis-norm [8]	$||x||_M = \sqrt{x^T \Sigma^{-1} x}$

is widely used, but the Manhattan norm (1-norm) is also common [5, 6, 7]. In this paper, the Mahalanobis norm is used by applying the sample means and the sample variances of the fingerprints. Because samples from different APs are assumed to be mutually independent, the covariance matrix

$\Sigma_i \in \mathbb{R}^{n_y \times n_y}$ used in the computation of the norm $||x||_M$ is the diagonal matrix

$$\Sigma_i = \begin{bmatrix} \hat{\sigma}_{i1}^2 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \hat{\sigma}_{in_y}^2 \end{bmatrix}.$$

The estimator (2), which keeps K biggest weights and sets the others to zero is called the *Weighted K-nearest neighbor* method (WKNN) [6]. WKNN with equal weights is called the *K-nearest neighbor* method (KNN) [5]. The simplest method, where $K = 1$, is called the *Nearest neighbor* method (NN) [9].

In general the KNN and the WKNN can perform better than the NN method, particularly with parameter values $K = 3$ and $K = 4$ [6]. However, if the density of the radio map is high, NN method can perform as well as the more complicated methods.

B. Probabilistic framework

In the probabilistic (or statistical) framework the state $\mathbf{x}$ is assumed to be a random vector [10]. The idea in the probabilistic framework is to compute the conditional pdf $p_{\mathbf{x}|\mathbf{y}}(x|y) \triangleq p(x|y)$ (*posterior*) of the state $\mathbf{x}$ given measurements $\mathbf{y} = y$. The posterior contains all the necessary information to compute an arbitrary estimate of the state and an estimate of the error. Using the Bayes' rule we get

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)}, \qquad (4)$$

where $p(y|x)$ is the *likelihood*, $p(x)$ is the *prior* and $p(y)$ is a normalizing constant. In this section, we use a uniform prior which is a conventional choice for the prior [10], that is

$$p(x) = \frac{\sum_{i=1}^{M} \chi_{B_i}(x)}{\sum_{j=1}^{M} |B_j|}, \qquad (5)$$

where $|B_i|$ is the volume of B_i and

$$\chi_{B_i}(x) = \begin{cases} 1, & x \in B_i \\ 0, & x \notin B_i \end{cases}.$$

In Section IV, the posterior distribution is computed in time series, and the prior is computed using the state model and the posterior from the previous time step. In this section the emphasis is on the computation of the likelihood. We assume that the measurements collected at the calibration point represent the distribution of the RSSI in the whole cell. That is, the likelihood is constant inside each cell B_i, so

$$p(y|x) = \sum_{i=1}^{M} p(y|i)\chi_{B_i}(x), \qquad (6)$$

where $p(y|i) = p_{\mathbf{v}_i}(y - \bar{a}_i)$ and $\mathbf{v}_i = \mathbf{y} - \bar{a}_i$. We assume that the components of the random vector $\mathbf{v}_i$ are independent. Thus,

$$p(y|i) = \prod_{j=1}^{n_y} p_{\mathbf{v}_{ij}}(y_j - \bar{a}_{ij}),$$

where $y \in \mathbb{R}^{n_y}$. The pdf $p_{\mathbf{v}_{ij}}$ of the measurement error of the RSSI signal from the jth station in the ith cell is proportional to the current centralized histogram in the radio map if the calibration time is long enough. Hence, the normalized centralized histogram $H_{\mathbf{v}_{ij}}(x)$ is the natural approximation of $p_{\mathbf{v}_{ij}}$. However, usually the calibration time is not long enough, and it is rational to approximate the incomplete calibration data somehow. Often the calibration phase is incomplete in the sense that bins with zero probability occur in some parts of the histograms. That can be prevented by applying a uniform prior which assigns a small fraction of the total probability mass, e.g. the inverse of the sample size, to all bins [10].

There are several approaches for computing the likelihood $p(y|i)$. Examples of these methods are given in Table II, where $K(\cdot)$ denotes the kernel function, a_{ij}^k is the tth element of the vector a_{ij}, $|a_{ij}|$ denotes the number of elements in vector a_{ij}, and $h > 0$ is a smoothing parameter which determines the width of the kernel [11].

<div align="center">

Table II
LIKELIHOOD

</div>

Name of the method	$p(y|i) = \prod_{j=1}^{n_y} p_{\mathbf{v}_{ij}}(y_j - \bar{a}_{ij})$
Histogram [10]	$p_{\mathbf{v}_{ij}}(x) = H_{\mathbf{v}_{ij}}(x)$
Kernel [10]	$p_{\mathbf{v}_{ij}}(x) = \dfrac{\sum_{k=1}^{|a_{ij}|} K\left(\frac{x+\bar{a}_{ij}-a_{ij}^k}{h}\right)}{|a_{ij}|h}$
Gaussian [12]	$p_{\mathbf{v}_{ij}}(x) = \dfrac{1}{\sqrt{2\pi\hat{\sigma}_{ij}^2}} \exp\left(-\frac{x^2}{\hat{\sigma}_{ij}^2}\right)$
Log-normal [13]	$p_{\mathbf{v}_{ij}}(x) = \dfrac{1}{x\sqrt{2\pi\hat{\sigma}_{ij}^2}} \exp\left(-\frac{(\log(x)-\mu)^2}{\hat{\sigma}_{ij}^2}\right)$
Inverse function	$p_{\mathbf{v}_{ij}}(x) = \begin{cases} \frac{2-|x|}{2\log(t)+3}, & |x| \leq 1 \\ \frac{1}{(2\log(t)+3)|x|}, & 1 < |x| \leq t \\ 0, & t < |x| \end{cases}$
Exponential	$p_{\mathbf{v}_{ij}}(x) = \frac{1}{2}e^{-|x|}$
Histogram comparison	$p(y|i) = \prod_j d_{\mathrm{pdf}}\left(H_{y_{ij}}, H_{a_{ij}}\right)$

Different probability density distance measures $d_{\mathrm{pdf}}(\cdot, \cdot)$ for the histogram comparison method are given in Table III [14].

<div align="center">

Table III
PROBABILITY DENSITY DISTANCES

</div>

Name	$d_{\mathrm{pdf}}(f, g)$
Infinity norm	$\sup |f(x) - g(x)|$
Lissack-Fu (LF)	$\sum |f(x) - g(x)|^p$
Bhattacharayya	$\sum \sqrt{f(x)g(x)}$
Kullback-Leibler (K-L)	$\sum f(x)\frac{f(x)}{g(x)}$
Simandl	$1 - \sum \min(f(x), g(x))$

Another way to handle the problem of incomplete data is non-parametric kernel approximation of the histograms. The RSSI fingerprint histograms are discontinuous and sensitive to disturbances due to varying MU's orientation and obstacles that affect the radio signal propagation. Thus it is reasonable to smoothen the discrete histogram to a continuous function and fill the gaps.

In the parametric approach, the RSSI histogram is approximated as a known function. Fingerprints include a lot of data, and thus it would be practical to use some parametric methods to approximate the histograms. Unfortunately, the distribution of the RSSI varies as a function of location and time because of the complex radio signal propagation environment. RSSI histograms can be asymmetric and multimodal. Thus, the modeling of the measurement error distribution with a known function at every location becomes challenging. In this paper the left-skeweness of some of the histograms is exploited by approximating the histograms with a log-normal distribution [13], when the absolute values of the signal strengths are used. The left-skewness occurs due to the observation that the variations of the weaker RSSI values are larger than the stronger RSSI values. However, assuming that can be problematic in the outer limits of the AP's range. The MU has a lower limit to the received RSSI and below that limit the RSSI values are too weak for MU to register. This limitation of MU restricts the tail of RSSI histograms from the left. Thus RSSI histograms can be left skewed, symmetric or right-skewed, depending on the distance and obstacles between the MU and APs.

The Gaussian approximation of the histograms has been widely discussed in the literature [6, 12, 13]. The Gaussian approximation can provide a good fit with some of the histograms, but in the case of skewed histograms the fit can be bad. Moreover, some of the histograms are quite concentrated around their mean, and the Gaussian approximation tends to spread the distribution too much. In those cases, the exponential function provides better approximation and improves the resolution of different fingerprints close to each other.

Substituting equations (5) and (6) into equation (4) we get the posterior

$$p(x|y) = \sum_{i=1}^{M} \beta_i \frac{\chi_{B_i}(x)}{|B_i|}, \qquad (7)$$

where

$$\beta_i = \frac{p(y|i)|B_i|}{\sum_{j=1}^{M} p(y|j)|B_j|}.$$

From this posterior we can compute an estimate of the state. One possible estimate is the *maximum-a-posterior* (MAP) estimate. In this case, it is the same as the *maximum-likelihood* (ML) estimate, because the prior is uniform. The posterior is piecewise constant, and thus the MAP estimator is ambiguous. If the posterior has the maximum value only in one cell, say B_i, it is reasonable to use the center of the cell as MAP estimate, that is

$$\hat{x}_{\mathrm{MAP}} = p_i.$$

Another commonly used estimate is the mean of the posterior,

that is

$$\hat{x}_{\text{MEAN}} = \int x p(x|y) \mathrm{d}x = \sum_{i=1}^{M} \beta_i p_i. \tag{8}$$

We see that there are many relations between the deterministic and the probabilistic methods and with particular selections of parameters both methods give identical estimates. For example, the probabilistic MAP estimate has the same form as the deterministic Nearest Neighbor estimate. Secondly, if the volumes of the cells are equal and $w_i = p(y|i) \ \forall i$, then the deterministic estimate (2) and the posterior mean (8) are identical. In addition, if the number of values of i, where the likelihood $p(y|i)$ is nonzero is K, then the posterior mean (8) has the same form as the deterministic Weighted K-nearest neighbor estimate.

IV. FILTERING APPROACH

In this section, the accuracy of localization is enhanced by exploiting also the previous measurements in addition to the current ones.

A. Bayesian filter

The idea in the Bayesian filtering is to compute the posterior $p(x_k|y_{1:k})$, where the subscript k represents time t_k and

$$y_{1:k} = \{y_1, \ldots, y_k\}$$

denotes the set of all measurements. Assuming independences and using Bayes' rule we get [15, 16]

$$p(x_k|y_{1:k}) = \frac{p(y_k|x_k)p(x_k|y_{1:k-1})}{p(y_k|y_{1:k-1})}, \tag{9}$$

where $p(y_k|x_k)$ is the measurement likelihood (6), and the prior is

$$p(x_k|y_{1:k-1}) = \int p(x_k|x_{k-1})p(x_{k-1}|y_{1:k-1})\mathrm{d}x_{k-1},$$

where $p(x_k|x_{k-1})$ is the transition density. Furthermore, the density function of the initial state is $p(x_0) \triangleq p(x_0|y_{1:0})$.

We use the same likelihood (6) as in the stationary case. We approximate all the priors and the posteriors as piecewise constant densities. Let the previous posterior be denoted as

$$p(x_{k-1}|y_{1:k-1}) = \sum_{j=1}^{M} \frac{\beta_{k-1}^{j+}}{|B_j|} \chi_{B_j}(x_{k-1})$$

and the current prior be denoted as

$$p(x_k|y_{1:k-1}) = \sum_{i=1}^{M} \frac{\beta_k^{i-}}{|B_i|} \chi_{B_i}(x_k),$$

where

$$\beta_k^{i-} = \sum_{j=1}^{M} \int_{B_i} \int_{B_j} \frac{p(x_k|x_{k-1})}{|B_j|} \mathrm{d}x_{k-1}\mathrm{d}x_k \beta_{k-1}^{j+}$$
$$\triangleq \sum_{j=1}^{M} \pi_{i,j} \beta_{k-1}^{j+} \triangleq \pi_i^T \beta_{k-1}^+.$$

Interestingly, in our case it is enough to model transition probabilities between the cells,

$$\pi_{i,j} \triangleq P(x_k \in B_i | x_{k-1} \in B_j).$$

This can be made using the floorplan of the building. We call the resulting state model the graph state model since it is based on the graph adjacency matrix

$$[\mathrm{G}]_{ij} = \begin{cases} 0, & \text{no connection between } \mathrm{B}_i \text{ and } \mathrm{B}_j \\ 1, & \text{connection between } \mathrm{B}_i \text{ and } \mathrm{B}_j. \end{cases}$$

The state transition matrix is obtained from the matrix T as

$$[\mathrm{T}]_{ij} = \frac{[\mathrm{G}]_{ij}}{\sum_i [\mathrm{G}]_{ij}}.$$

Thus, the prior is obtained from the previous posterior as

$$\beta_k^- = \mathrm{T}\beta_{k-1}^+.$$

B. Kalman filter

In our Kalman filter approach the Position Kalman Filter (PKF) [17] [18, II.D] is used to smoothen the estimates of the static estimators by using the static solutions as measurements for the conventional KF.

In this paper two state models were used, namely the stationary state model and the constant velocity model [19]. The stationary state model can be formulated as

$$\mathbf{x}_k = \mathbf{x}_{k-1} + \mathbf{w}_{k-1}, \tag{10}$$

where the state $\mathbf{x}$ includes only the position coordinates. This model can be used when there are no velocity measurements and the MU is mainly located indoors, and thus the velocity often stays small. The constant velocity model can be formulated as

$$\mathbf{x}_k = \mathrm{F}_{k-1}\mathbf{x}_{k-1} + \mathbf{w}_{k-1},$$

where the state $\mathbf{x}$ contains both the position and the velocity coordinates and

$$\mathrm{F}_{k-1} = \begin{bmatrix} \mathrm{I}_{2\times2} & \Delta t \mathrm{I}_{2\times2} \\ 0_{2\times2} & \mathrm{I}_{2\times2} \end{bmatrix},$$

where $\Delta t = t_k - t_{k-1}$ and I is identity matrix. In the constant velocity model of this paper, the state $\mathbf{x}$ also contains the 2-dimensional velocity in addition to the 2-dimensional position. The state noise is $\mathbf{w}_k \sim \mathrm{N}(0, \mathrm{Q}_k)$, where the state noise covariance matrix is

$$\mathrm{Q}_k = \begin{bmatrix} \frac{1}{3}\Delta t^3 \mathrm{I} & \frac{1}{2}\Delta t^2 \mathrm{I} \\ \frac{1}{2}\Delta t^2 \mathrm{I} & \Delta t \mathrm{I} \end{bmatrix} \sigma_C^2, \tag{11}$$

where σ_C^2 describes the error in the axis directions.

V. IMPLEMENTATIONS AND RESULTS

The algorithms and methods discussed in Sections III and IV were implemented and tested in the wireless local area network (WLAN). In Sections V-A and V-B, the collection of the radio maps and the test data is explained. In Section V-C, the test results of the static location estimation algorithms are presented. Section V-C also covers the impact of several environment parameters on the performance of the algorithms. In Section V-D, the filters are applied to the test data. The data was collected at the calibration phase and at the location estimation phase with a MacBook laptop containing an AirPort network card. Wireshark software was used to collect the radio map and the test data.

A. Radio maps

The radio maps were collected in an office area at the Tampere University of Technology. The division of the cells was done by exploiting the floor plan. Four radio maps were constructed to test the effect of the density of the radio map, orientation of the MU and the duration of the calibration time at the each CP.

The radio map 30_1 was collected at 40 CPs, with fixed orientation and 30 s measurement period. The radio map 60_{r1} was collected at the same CPs, but with 60 s measurement period and by rotating at the CPs. The rotation was done to minimize the effect of the orientation of the user and the calibration time was lengthened to increase the reliability of the fingerprint. The radio map $sparse$ was constructed by removing every second entry from the radio map 60_{r1}. The radio map 60_{r2} was also done with the 60 s measurement period and by rotating the MU, but the density of the radio map was increased to 77 CPs. The fingerprints were measured from 20 APs from which 11 were located on the same floor as the CPs, 5 on the floor below and 4 on the floor above. The radio maps are summarized in Table IV.

Table IV
RADIO MAPS

Radio map	Duration(s)	Rotation	CPs
$sparse$	60	√	20
30_1	30		40
60_{r1}	60	√	40
60_{r2}	60	√	77

B. Test data

Tracking the MU was tested by walking at nearly constant speed in the neighborhood of the cells. One of the four tracks collected is shown in Figure 1.

The raw data was preprocessed later, and the weakest signals of the test data were ignored to increase the reliability of the measurements.

The same set of 20 APs was used as in the calibration phase. The number of APs heard varied from 2 to 15. The most common number of APs heard was 9. For further details see [20].

Figure 1.   One of the four test tracks.

C. Static location estimation algorithms

The different methods described in Section III were tested. The radio map 60_{r2} was used in most of the tests. If the solver was not able to produce an estimate at some time step, the estimate from the previous time step was used. The initial state was set to be the location of the AP transmitting the strongest sample among all samples in the first 1 s time interval.

Each of the location estimation algorithms has several parameters which were varied to find the best performance. The $\hat{x}_{\mathrm{MEAN}}$ was used as an estimator in the tests, because it provided only slightly different results than $\hat{x}_{\mathrm{MAP}}$ throughout the tests. The mean of the norms of the 2-dimensional error vectors, denoted as ME, was used as the primary performance measure.

1) K-nearest neighbor: The K-nearest neighbor method and its modifications were tested. Different numbers of CPs were considered in the computation of the estimate by varying the parameter K. Different measures were used to compute the distances in the signal space.

The effect of varying the parameter K in KNN method is studied first. It is examined whether consideration of multiple "nearest" neighbors has any benefit to the performance. The MU was moved through the CPs and considering multiple CPs did not provide better performance.

The impact of the distance measure used is considered next. The parameter p was varied from 1 to 5 in the p-norm distance measure in the NN method; ∞-norm and Mahalanobis norm were also tested. The 1-norm resulted in the smallest ME.

The sample size of the fingerprint can be interpreted as a measure of the reliability of that fingerprint [7]. Direct

weighting with the inverse of the sample size improved the results slightly; ME decreased from 5.6 m to 5.4 m. The test was done with the NN method using the 1-norm as a distance measure.

2) Histogram method: The binwidth of the histograms was varied and the binwidth $b = 1$ gave the smallest ME. The use of the uniform distribution also improved the results and prevented the cells from having zero likelihood.

3) Histogram comparison method: The Simandl and the Bhattacharayya distances provided the smallest ME (Table V).

Table V

HISTOGRAM COMPARISON METHOD

Distance	ME (m)
Infinity norm	8.5
Lissack-Fu	9.7
Bhattacharayya	7.0
Kullback-Leibler	11.1
Simandl	7.0

4) Kernel method: The kernel method was tested with several kernel functions and kernel widths. The kernel functions were applied to the normalized fingerprint histograms with the histogram binwidth $b = 1$.

Different kernel functions were tested with the kernel width $h = 2$ which provided the smallest ME (Figure 2). The exponential kernel function is claimed to provide good results [21], and in this work, it also gave better performance compared to the other kernel functions.
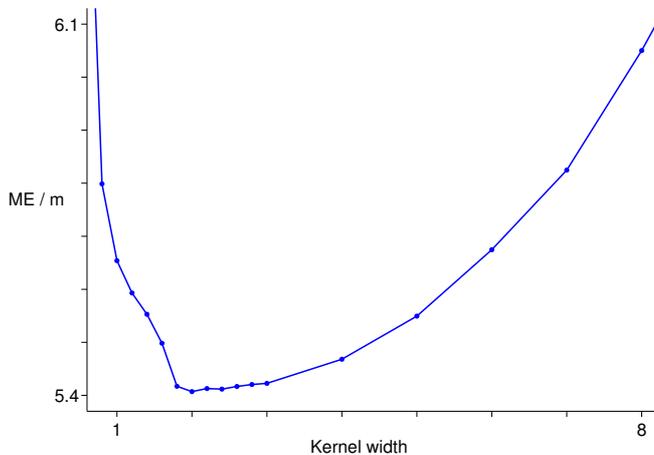
Figure 2.   Effect of kernel width with exponential kernel function

5) Parametric approximation of measurement noise: The measurement likelihood was computed using different parametric approximations of the normalized histograms. In the tests, the exponential approximation produced the smallest ME (Table VI).

6) Radio map density: The density of the radio map was varied by changing the number of CPs considered during

Table VI

PARAMETRIC APPROXIMATION OF HISTOGRAMS

Function	ME (m)
Log-normal	8.7
Gaussian	6.0
Exponential	5.5
Inverse	6.4

the location estimation phase; radio maps $sparse$, 60_{r1} and 60_{r2} were used. As seen in Figure 3, the denser the radio map the better the results with every tested method. However, the improvement between the radio maps $sparse$ and 60_{r1} is larger than the improvement between the radio maps 60_{r1} to 60_{r2}. Thus the improvement in the results does not grow linearly as a function of the radio map density. Histogram and histogram comparison methods benefit the most from the refinement of the radio map.
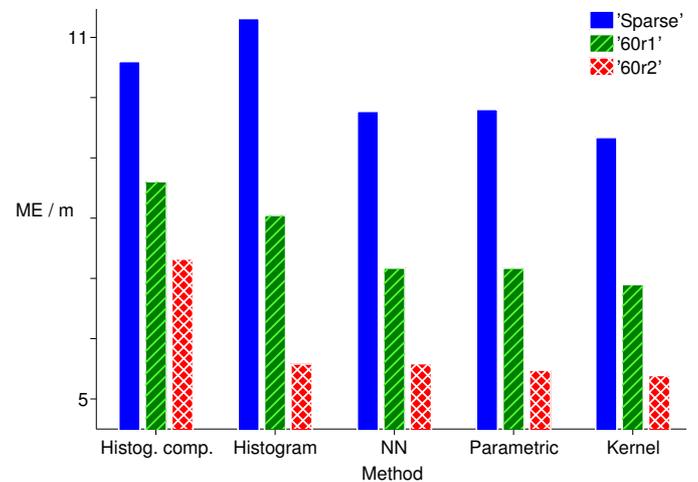
Figure 3.   Effect of radio map density

7) Single orientation vs. varying orientation: It was tested whether the orientation of the user during the calibration phase has any effect on the location estimation phase. This was done by using radio maps 30_1 and 60_{r1} with calibration time truncated to 30 s. The purpose of the rotation was to equalize the impact of the MU's orientation to measure more reliable fingerprints.

Figure 4 illustrates how all tested methods benefit significantly from the rotation of the MU during the calibration phase. Especially the histogram comparison and the histogram method gain a lot from the varying orientation. The kernel and the parametric methods resulted in smaller ME compared to the NN method which has the smallest ME compared to other methods when the MU is not rotating.

8) Calibration time: The effect of varying the calibration time at each CP was examined by parsing different calibration times from the radio map 60_{r2}. Figure 5 shows ME of the different methods as a function of the calibration time. The histogram method needs a lot of data to produce reliable like-
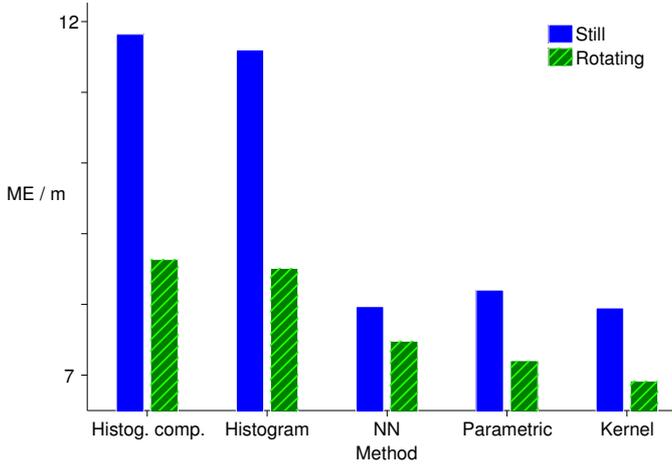
Figure 4. Effect of orientation

Figure 6. Effect of number of APs in the test data

lihoods and the prolonging of the calibration time up to about 30 s clearly improves its estimates. However, the histogram method reaches its minimum ME with the maximum 60 s calibration time. The histogram comparison method reaches its best performance already in about 10 s.
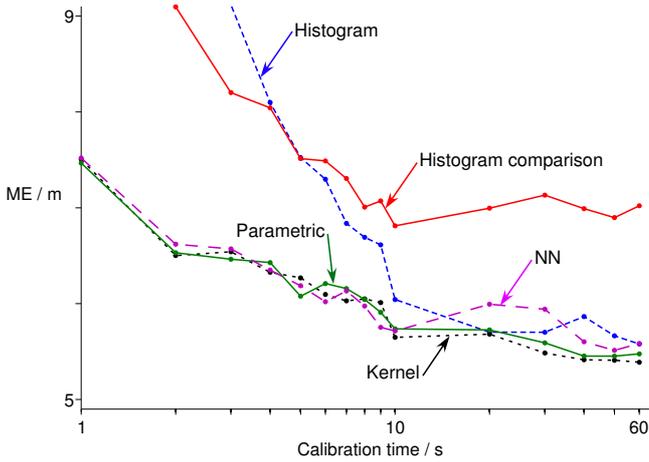

Figure 5. Effect of calibration time

The parametric approximation, NN and the kernel methods perform well even with short calibration time. The largest drop in the ME occurs between 1 s and 2 s. When more data is collected, the kernel method performs best. The parametric approximation, NN and the kernel methods perform well, because they fill the incomplete fingerprint data.

9) Number of access points in the test data: In this work a set of 20 APs was used. In this section the effect of varying the number of APs used in the test data is examined. In the test bed, none of the ranges of the 20 APs covers the whole test area. Thus to test the effect of number of available APs on the performance of the algorithms, the measurement from the strongest APs were used in the test data.

Figure 6 illustrates the ME as a function of the number of

available APs. All tested methods benefit from increasing the number of available APs. However, the ME is not a linear function of the number of APs: increasing the number of available APs beyond five does not improve the performance dramatically. The order of the different methods according to ME stays almost the same with any number of APs and the kernel method has the smallest ME throughout the test.

10) Summary of static location estimation algorithms: In this section, the algorithms are compared by using the best parameter values found in this work. The numerical results are shown in Table VII. The probabilistic histogram, kernel and parametric approximation methods performed slightly better than the deterministic NN method. The histogram comparison method resulted in the largest ME whereas the kernel method resulted in the smallest ME.

Table VII
STATIC ESTIMATORS, RADIO MAP 60_{r2}

Method	ME (m)	Median (m)	RMSE (m)	Max (m)	95th (m)
Histogram	5.5	4.3	8.4	104.5	14.2
Kernel	5.4	4.1	8.9	98.6	12.3
Parametric	5.5	4.3	8.7	106.3	13.1
Histogram comp	7.0	5.0	11.0	106.0	19.4
NN	5.6	4.4	8.8	119.6	13.7

D. Filtering algorithms

The filtering approach was tested by applying the Bayesian filter with the graph state model and PKF with the stationary and the constant velocity state models to the static location estimation algorithms, as discussed in Section IV.

The estimate of the static location estimation algorithm was given to PKF as a measurement with the covariance matrix $R_k = 4 \cdot I$. The time step in the tests was constant ($\Delta t = 1$ s) and the state noise covariance for the stationary state model was $V(\mathbf{w}_k) = 8.3 \cdot I$ (Eq. (10)). The state noise covariance for the constant velocity model is given in Eq. (11) where $\sigma_C^2 = 2$ was used as a parameter. The filtering approach smoothens
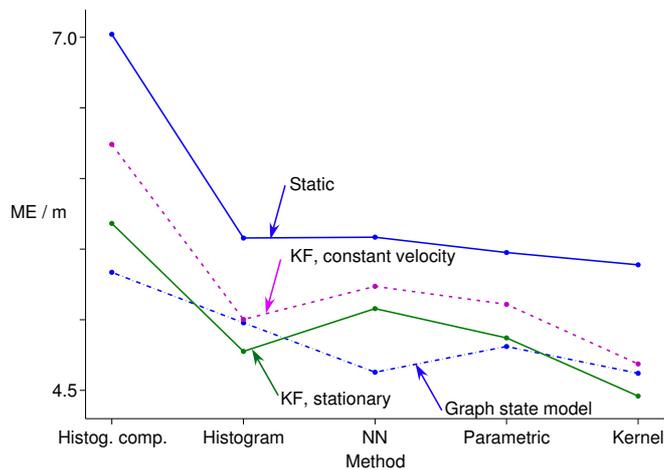
Figure 7. Filtering approach

the trajectory of the estimates, and thus reduces ME which follows especially from the reduction of the maximum error. Figure 7 illustrates the performance of the different filters compared to the static estimation algorithms. The smallest ME was achieved with the PKF that used the stationary state model.

Table VIII summarizes the performance of the filters when they were applied to the static kernel method.

Table VIII
FILTERS APPLIED TO KERNEL METHOD, RADIO MAP 60_{r2}

Method	ME (m)	Median (m)	RMSE (m)	Max (m)	95th (m)
Bayesian filter, graph	4.6	3.9	5.7	19.9	11.0
PKF, constant velocity	4.7	4.0	5.8	22.4	11.1
PKF, stationary	4.5	3.8	5.5	20.2	10.8

VI. CONCLUSIONS

In this paper, different location fingerprinting methods were considered by introducing the mathematical basis of the methods and testing them with WLAN RSSI measurement data. The mathematical formulation was carried out from different points of view, and the parameters of the methods were varied in the tests in order to obtain the best performance. The environment variables, such as the number of access points (APs) and the radio map density, were also varied and the methods were compared also in these varying circumstances.

The main goal in the mathematical formulation was to model the location as a continuous random variable which led to the division of the area of interest into rectangular cells instead of considering only the individual calibration points. However, deterministic methods, such as the K-nearest neighbor (KNN) method and the Weighted K-nearest neighbor (WKNN) method, were formulated with the discrete location variable, as they are presented in the literature.

The Bayesian filtering approach with the graph state model performed clearly better than the static algorithms. PKF was tested with two state models, namely, the stationary and constant velocity state models. Both of these state models resulted in better performance than the static estimators, but the stationary state model outperformed the constant velocity model. PKF with the stationary state model performed better than the Bayesian filter with the graph state model when the histogram and the kernel method were used as the static estimation algorithms.

The radio map needs good planning, but an adequate grid density is hard to determine because it depends on the floor plan and the locations and number of available APs. In the tests, calibration times longer than 10 s did not improve the performance dramatically, but the benefit of rotation during the calibration was significant. In addition, the increase in the number of available APs did not diminish ME linearly. These results can be considered as guidelines in the design process.

In this work the calibration points were chosen to be in the centers of the cells; more tests can be made to collect the data from all over the cell to obtain perhaps more reliable signal characterization inside the cells. Moreover, it would be interesting to examine how well the algorithms work when the measurements used in the location estimation phase are collected with a different receiver than what was used for the radio map collection. In that case one should pay attention to the receivers' different RSSI units.

REFERENCES

[1] J. Hightower and G. Borriello, "Location systems for ubiquitous computing," *IEEE Computer*, vol. 1, no. 34, pp. 57–66, August 2001.

[2] Ekahau. [Online]. Available: http://www.ekahau.com/

[3] M. Wallbaum, "Indoor geolocation using wireless local area networks," Ph.D. dissertation, RWTH Aachen University, February 2006.

[4] M. Vossiek, L. Wiebking, P. Gulden, J. Wieghart, C. Hoffman, and P. Heide, *Wireless local positioning*. IEEE Microwave magazine, 2003, pp. 77–86.

[5] P. Bahl and V. N. Padmanabhan, "Radar: An in-building RF-based user location and tracking system," in *INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies*, vol. 2, no. 10, March 2000, pp. 775–784.

[6] B. Li, J. Salter, A. G. Dempster, and C. Rizos, "Indoor positioning techniques based on wireless LAN," School of Surveying and Spatial Information Systems, UNSW, Sydney, Australia, Tech. Rep., 2006.

[7] P. Prasithsangaree, P. Krishnamurthy, and P. Chrysanthis, "On indoor position location with wireless LANs," Telecommunications Program, University of Pittsburgh PA 15260, Tech. Rep., 2002.

[8] W. M. Yeung, J. Zhou, and J. K. Ng, *Emerging Directions in Embedded and Ubiquitous Computing*. Springer, 2007, ch. Enhanced Fingerprint-Based Location Estimation System in Wireless LAN Environment, pp. 273–284.

[9] S. Saha, K. Chauhuri, D. Sanghi, and P. Bhagwat, "Location determination of a mobile device using IEEE 802.11b access point signals," Department of Computer Science and Engineering, Tech. Rep., 2003.

[10] T. Roos, P. Myllymäki, H. Tirri, P. Misikangas, and J. Sievänen, "A probabilistic approach to WLAN user location estimation," *International Journal of Wireless Information Networks*, vol. 9, no. 3, pp. 155–163, July 2002.

[11] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. John Wiley, Sons Inc., 2001.

[12] A. Haeberlen, E. Flannery, A. M. Ladd, A. Rudys, D. S. Wallach, and L. E. Kavraki, "Practical robust localization over large-scale 802.11 wireless networks," MobiCom'04, Tech. Rep., 2004.

[13] K. Kaemarungsi, "Distribution of WLAN received signal strength indication for indoor location determination," National Electronics and Computer Technology Center, Thailand, Tech. Rep., 2006.

[14] N. Sirola, "Mathematical methods for personal positioning and navigation," Ph.D. dissertation, Tampere University of Technology, August 2007. [Online]. Available: http://webhotel.tut.fi/library/tutdiss/pdf/sirola.pdf

[15] A. Doucet, N. de Freitas, and N. Gordon, Eds., *Sequential Monte Carlo Methods in Practice*, ser. Statistics for Engineering and Information Science. Springer, 2001.

[16] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter, Particle Filters for Tracking Applications*. Boston, London: Artech House, 2004.

[17] S. Ali-Löytty, N. Sirola, and R. Piché, "Consistency of three Kalman filter extensions in hybrid navigation," in *Proceedings of The European Navigation Conference GNSS 2005*, Munich, Germany, July 2005.

[18] J. Kwon, B. Dundar, and P. Varaiya, "Hybrid algorithm for indoor positioning using wireless LAN," *IEEE 60th Vehicular Technology Conference, 2004. VTC2004-Fall.*, vol. 7, pp. 4625–4629, September 2004.

[19] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. Academic Press, 1970, vol. 64.

[20] V. Honkavirta, "Location fingerprinting methods in wireless local area networks," M.Sc. thesis, Tampere University of Technology, 2008. [Online]. Available: http://math.tut.fi/posgroup/

[21] A. Kushki, K. Plataniotis, and A. Venetsanopoulos, "Radio map fusion for indoor positioning in wireless local area networks," *The 7th International Conference on Information Fusion*, pp. 1311–1318, 2005.