A Survey of Parametric Fingerprint-Positioning Methods

ISSN 2075-1087, Gyroscopy and Navigation, 2016, Vol. 7, No. 2, pp. 107–127. © Pleiades Publishing, Ltd., 2016.
Published in Giroskopiya i Navigatsiya, 2016, No. 1, pp. 3–35.

A Survey of Parametric Fingerprint-Positioning Methods[1]

Ph. Müller[a], M. Raitoharju[a], S. Ali-Löytty[b], L. Wirola[a], and R. Piché[a]

[a] *Department of Automation Science and Engineering (ASE), Tampere University of Technology, Tampere, Finland*
[b] *Department of Mathematics, Tampere University of Technology, Tampere, Finland*
e-mail: philipp.muller@tut.fi, matti.raitoharju@tut.fi, robert.piche@tut.fi, simo.ali-loytty@tut.fi

Received August 31, 2015

Abstract—The term fingerprint-based (FP) positioning includes a wide variety of methods for determining a receiver's position using a database of radio signal strength measurements that were collected earlier at known locations. Nonparametric methods such as the weighted k-nearest neighbor (WKNN) method are infeasible for large-scale mobile device services because of the large data storage and transmission requirements. In this work we present an overview of parametric FP methods that use model-based representations of the survey data. We look at three different groups of parametric methods: methods that use coverage areas, methods that use path loss models, and methods that use Gaussian mixtures. Within each group we study different approaches and discuss their pros and cons. Furthermore, we test the positioning performance of several of the analyzed approaches in different scenarios using real-world WLAN indoor data and compare the results to those of the WKNN method.

DOI: 10.1134/S2075108716020061

I. INTRODUCTION

Over the last decades positioning techniques have received extensive attention, because they have become the backbones of an increasing number of position-aware applications in commercial, public service, and military networks [1, 2]. Those applications include vehicle navigation, intelligent transport systems, inventory tracking, vehicle tracking, fleet management, resource management, environment monitoring, emergency services (E911 in North America, E112 in Europe), medical services (e.g. patient and equipment surveillance in hospitals), and rescue operations (e.g. locating fire fighters in burning buildings). For smart phone users applications include location identification, local search, suggesting local points of interest, geo-tagging of photos and videos, location sensitive billing, and targeted advertising [2–8]. Because many of these applications have to run on small mobile devices, the positioning algorithms have strict limits on allowed energy, memory, bandwidth, and computational resources.

In outdoor environments positioning techniques mainly rely on Global Navigation Satellite System (GNSS) signals. Because nowadays almost all new smart phones are equipped with a GNSS receiver precise positioning of the mobile user equipment (UE) outdoors is continuously possible. However, GNSS receivers use significant amounts of energy. Furthermore, indoors and also under forest canopies and in certain urban settings, such as urban canyons, poor signal penetration by GNSS generally results in unavailable or unreliable location information. Therefore, positioning in those environments must rely on other measurements, e.g. from an inertial measurement unit (IMU) or from radio signals such as cellular networks, Bluetooth, wireless local area networks (WLAN), or ultra-wideband (UWB).

Depending on the used algorithm, positioning using cellular telephone measurements can be simple, economic and can be implemented without upgrading UEs or network equipment [3, 9, 10]. Although the cellular network was not originally designed with positioning in mind,[2] it provides in most environments sufficient accuracy (around 100 m indoors and in urban areas possible, around 200 m in suburban areas) for applications such as local search or weather forecast [12]. Since WLANs are ubiquitous in urban areas and the coverage areas of WLAN access points (APs) are much smaller than cellular network cells, WLAN-based positioning techniques have come to be preferred over cellular as alternatives to GNSS [13] when higher accuracy positioning is desired. Like cellular networks, WLANs were not originally designed for positioning. Also WLAN-based positioning uses the already existing infrastructure, and WLAN APs and receivers in UEs are widely available. In the remainder we focus on WLANs although many of the concepts apply also to cellular networks.

[1] The article is published in the original.

[2] An exception is the Long Term Evolution (LTE) standard, which specifies in release 9 the positioning reference signal (PRS), which makes a positioning accuracy on the order of 10 m possible [11].

Most WLAN-based positioning algorithms exploit the correlation between the received signal strength (RSS) and the UE's location (see e.g. [14, p. 47]). Because modeling signal propagation, especially in indoor environments, is rather complex, nonparametric location fingerprinting methods are widely applied for positioning [15]. Those methods estimate the user's position by comparing the list of current AP received signal strength indicator (RSSI) or RSS measurements to a database (called a radio map) of information (called fingerprints) on APs and their corresponding signal strength values for known positions.

Parametric methods include various approaches. In contrast to the nonparametric methods, they only store some parameters (e.g. the parameters of a signal propagation model) in the radio map that summarize the fingerprints (also known as allocation reports, reception reports, or observations) in a certain way, reducing the radio map's size significantly. The user's position can be estimated using parametric methods, for example, by computing distance estimates between the user and the APs using the received RSS/RSSI measurements in a signal propagation model. Within this paper we present an overview of these parametric methods, and compare them with each other and a widely-used nonparametric method.

The main contribution of our paper is to give an overview about recent developments in the field of parametric fingerprint-positioning methods, and analyses their strengths and weaknesses. For nonparametric fingerprinting methods overviews can be found, for example, in [15, 16].

The results of the field tests have been already published in [17]. However, in this paper we provide a more detailed analysis of the results and the parametric methods that were compared. Furthermore, we analyze additional parametric methods, which have not been considered in [17], and compare them with those considered in [17].

The outline of this paper is as follow. We discuss similarities and differences of nonparametric and parametric fingerprinting methods as well as methods used for fingerprint collection and related issues in Section II. The parametric methods presented use Bayes' rule and Bayesian filtering, thus we briefly summarize the idea of positioning using Bayes' rule and Bayesian filtering in Section III. In Section IV we look at parametric FP approaches that use rather simple models for describing the area that an AP covers, i.e. in which it can be heard. Those methods reduce the radio map's size tremendously while providing sufficient accuracy for many applications in the positioning phase. Section V is dedicated to signal propagation path loss models, which are calibrated from FP data. In Section VI we present an overview of approaches that rely (partly) on parametric fingerprinting techniques, and use mixtures of distributions for modeling position estimates etc. These techniques are useful for nonlinear and/or non-Gaussian systems for which traditional approaches such as Kalman filter (KF) and extended KF (EKF) perform poorly. The performance of a selection of different parametric FP techniques, with and without filtering, is compared in Section VII using benchmark tests using real-world WLAN measurements in indoor environments. Section VIII summarizes and concludes.

Notation: Scalar variables are italic, **x** denotes column vectors, and **H** denotes a matrix.

LIST OF USED ABBREVIATIONS

AP—access point
AP-ID—access point identifier
AS—adaptive splitting
CA—coverage area
CN—communication node
EGMF—efficient Gaussian mixture filter
EKF—extended Kalman filter
EM—expectation maximisation
FP—fingerprint
GGM—generalised Gaussian mixture
GGMF—generalised Gaussian mixture filter
GM—Gaussian mixture
GMA—Gaussian mixture allowing negative weight
GMB-REM—Gaussian mixture Bayes' with regularised expectation maximisation
GMEM—signal strength estimation model from [44]
GMF—Gaussian mixture filter
GNSS—global navigation satellite system
ID—identifier
IMU—inertial measurement unit
IRLS—iterative reweighted least squares
KF—Kalman filter
KNN—k-nearest neighbour
LS—least squares
MAC—media access point
ML—maximum likelihood
MMSE—minimum mean square error
NN—nearest neighbour
non-LOS—non-line of sight
PL—path loss
RSS—received signal strength
RSSI—received signal strength indicator
SPGMF—sigma point Gaussian mixture filter
SSM—state space model
TP—test point
UE—user equipment
UKF—unscented Kalman filter
UWB—ultra-wideband
WKNN—weighted k-nearest neighbour
WLAN—wireless local area network

II. WHAT ARE PARAMETRIC AND NONPARAMETRIC APPROACHES?

In this section we define what we mean by nonparametric and parametric fingerprinting positioning methods, what they have in common and how they differ. Furthermore, we will discuss issues and possible solutions related to fingerprinting.

The aim of both parametric and nonparametric positioning methods is to determine the $n_{\mathbf{x}}$-dimensional state $\mathbf{x}_k \in \mathbb{R}^{n_{\mathbf{x}}}$ given the $n_{\mathbf{y}}$-dimensional measurements in vector $\mathbf{y}_k \in \mathbb{R}^{n_{\mathbf{y}}}$. The fingerprinting positioning methods have an offline and an online phase. In the offline phase the state $\mathbf{x}_k$ denotes a vector of parameters, while in the online phase the state $\mathbf{x}_k$ denotes the UE position at time t_k and possibly additional information such as the UE velocity.

The measurements in the offline are so-called fingerprints (FPs) collected at known locations. Therefore, they are sometimes also called location FPs. For WLAN-based indoor localization, FPs are generally collected in grid points with one grid point per square meter [18]. The FP data for one of these points consists of identifiers (IDs) of APs from which signals are received by the UE in that specific grid point together with the corresponding signal strength values. From the collected FPs a radio map is generated. The radio map generation of various parametric and nonparametric methods is explained in the remainder of this paper.

In the online phase, measurements in form of a FP are collected in the UE's unknown location. The FP's data depends on the UE location. This dependency between UE position $\mathbf{x}_k$ and measurements $\mathbf{y}_k$ is used for estimating the UE position, and the estimation process uses the radio map entries. In the remainder of this paper several methods for determining position estimates are explained and analyzed.

A. Parametric vs. Nonparametric Fingerprinting

Nonparametric fingerprinting methods use radio maps in which FPs $\mathbf{y}_k$ are stored. For a WLAN the radio map contains, in general, location coordinates, IDs of APs from which signals were received in this location and corresponding signal strength values. In the positioning phase the UE's measurement (AP-IDs and corresponding RSS values) are compared with the radio map entries to infer a position estimate. The simplest approach is the nearest neighbor (NN) method. It returns the location of the FP from the radio map whose measurement is most similar to the UE's measurement as position estimate. This FP is called the nearest neighbor and is found by optimizing a given cost function [19].

A more advanced, widely applied version of the NN is the weighted k-nearest neighbor (WKNN) method. Here, the position estimate is the weighted mean of k locations whose FPs are most similar to the UE's measurement [15]. According to Liu et al. [7] the WKNN combines medium complexity and medium cost with a good robustness and accuracy. Therefore, we use it in Section VII for comparison with parametric localization methods. For overviews on nonparametric location fingerprinting methods we refer the reader to [15, 20] and references therein.

Parametric fingerprinting methods use radio maps in which parameters $\mathbf{x}_k$ are stored that summarize the FP data $\mathbf{y}_k$. Instead of having one FP per grid point, the radio map contains a set of parameters for each AP observed in the FP data. For positioning various methods, dependent on the parameters used in the radio map, can be used. In the following sections we will discuss those methods further.

Nonparametric fingerprinting methods have the advantage that modeling the signal propagation is not needed. These algorithms have been shown to be reasonably precise and reliable in indoor environments (see e.g. [15, 21]), where non-line of sight (non-LOS) situations are very common. Many parametric fingerprinting methods require modeling the signal propagation, which causes problems in challenging indoor environments. In the latter sections of this paper we will consider this topic in more detail.

A major difference between parametric and nonparametric methods is that the radio map's size depends for nonparametric methods on the number of FPs (i.e. grid points), while for parametric methods it depends on the number of APs and number of parameters stored per AP. This means, if more FP data is collected to improve the radio map quality the map's size increases for nonparametric methods while for parametric methods it stays the same, provided that no additional APs are detected.

To get a better understanding on what kind of sizes we are talking it should be noted that for each grid point around 100 samples have to be collected [18] to obtain a reliably FP for the radio map. Since nonparametric FP positioning works directly with the FP data, the size of this data can be a critical issue when FP-based positioning is offered as a large-scale service for mobile devices. Wirola et al. [12] point out that a radio map of just 0.1% of the earth's surface (approximately 130000 km^2 or the size of Greece or Louisiana, USA), with an average density of one FP per m^2, and at least one 6-byte AP Media Access Point (MAC) address per point, needs at least 780 GB. For indoor environments often signals from more than 5 APs can be received. In addition, signal strength values are stored besides the APs' MAC addresses.

One approach for shrinking the radio map for nonparametric method is the usage of data-compression techniques [22, 23]. A more fundamental way to address the issue is to use parametric (model-based) FP methods. Since an AP's signal is generally receivable in many grid points the radio map for parametric methods is in general significantly smaller [17].

Before moving on to the discussion of various parametric methods let us consider how FPs are collected and some of the issues plus solutions related to it.

B. Fingerprinting and Related Issues

For location fingerprinting the fingerprints are collected in an offline phase by site survey, war-driving or crowd-sourcing. In addition to the UE's current position, each fingerprint contains radio characteristic records. In WLANs those records, in general, contain at least AP-IDs and corresponding RSS or RSSI values. Site survey means that the FPs are collected at various locations for generating a fingerprint database. Most often the location has to be entered by hand. This is a significant difference to war-driving, where the location generally is a GNSS-based estimate. As the name suggests war-driving is the act of collecting measurements in a moving vehicle. A disadvantage of both site survey and war-driving is that the data collecting is tedious and expensive. Thus, crowd-sourcing is preferred. In this form of data collection several persons (for example the users of the FP-based localization method) are collecting the FPs. In practice a combination of those methods can be used.

In outdoor environments FP data can be collected via crowd-sourcing or war-driving, making radio map updating less laborious. However, due to the absence of GNSS signals this technique cannot be used indoors and more complicated alternatives have to be employed. Note that it is, nevertheless, possible to use FPs without GNSS-based position (so called unlocated FPs) but then the position accuracy decreases [24].

An important issue in fingerprint positioning is maintaining the radio map. Since network topography (APs can be added, removed, moved or modified) and radio environment change constantly, constant updates of the FP radio map are required to prevent performance deterioration [15, 25]. Already small environment changes can have tremendous influence on the measured RSS values. If the FPs were, e.g., collected while the building was empty, and positioning is done in the same building while crowded with people, the RSS values in a certain position will differ significantly due to body shadowing [18]. The influence of the user's body on RSS has been analyzed, e.g., by Kaemarungsi and Krishnamurthy [26].

Device heterogeneity is another problem that has to be considered. It describes the fact that received signal strength values measured by different devices at the same location and time can vary significantly [13, 21, 27–29]. The heterogeneity includes, for example, the fact that many devices show only unitless RSSI values rather than RSS values, which are always in dBm. Different chipset providers use different RSSI scales with different limits and granularity, which hinders comparison of RSS measurements from different devices [27, 29–31]. To circumvent the problem of RSS heterogeneity some authors use either rankings of RSS values in a FP [21, 32] or RSS ratios [31] or RSS differences [33] instead of the measured RSS values. In the following we will distinguish between RSS and RSSI, depending on which values were used by the cited authors or are needed for the specific task.

To resolve the RSSI scaling problem, calibration methods within the devices could be used; see [28] for a brief overview. Earlier proposed techniques use manual calibration, meaning that users collect measurements at known locations in order to calibrate their device. These methods generally use least-square fitted linear calibration functions [27, 30]. Park et al. [13] argue that calibrating the signal strength levels is insufficient and that differences in signal strength dispersion between devices is also important, and they describe a kernel estimation method for device calibration. Approaches for automatic calibration, which enable simultaneous calibration and positioning, and eliminate the requirement of taking measurements at known reference locations, are proposed, for example, in [28] and [29]. In [29] Koski et al. suggest to construct a RSSI histogram from values collected over a long period. By assuming that distributions of RSSI values for all devices are the same up to translation and scale, these device-specific parameters can be fitted from the histogram. Laoudias et al. [28] make a more thorough study of the histogram approach. They show that assuming that the true RSS value distribution is the same for all users, the parameters for a linear mapping of RSS between devices can be estimated from the histogram. In contrast to Park et al. [13], they find that the device heterogeneity effect is sufficiently reduced by a linear mapping, and that even a simple origin-shift suffices. They describe a method to continually update the calibration on-line while positioning, and find that positioning with this method is as accurate as with a manual calibration except at the beginning of the learning phase.

III. POSITIONING AND BAYESIAN FILTERING

All of the parametric methods presented in this paper use Bayes' rule, for determining parameters and static positioning. For filtered positioning problems Bayesian filtering is used. Thus, we now summarize the Bayesian approaches for both static and filtered positioning problems.

For the static case the state of the system $\mathbf{x}_k$ is estimated as follows. The measurements in vector $\mathbf{y}_k$ are used to determine the posterior probability density function (pdf) of the state by applying Bayes' rule

$$p(\mathbf{x}_k|\mathbf{y}_k) \propto p(\mathbf{y}_k|\mathbf{x}_k)p(\mathbf{x}_k), \quad (1)$$

where $p(\mathbf{x}_k)$ can be an uninformative prior pdf, i.e. constant.

The filtered positioning problem is formulated as a Bayesian filtering problem for a discrete-time state space model (SSM). Within this section we consider

the general discrete-time SSM, namely the nonlinear non-Gaussian additive noise system

$$\mathbf{x}_k = \mathbf{f}_{k-1}(\mathbf{x}_{k-1}) + \mathbf{w}_{k-1}, \quad (2a)$$

$$\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k) + \mathbf{v}_k, \quad (2b)$$

where the errors $\mathbf{w}_{k-1}$ and $\mathbf{v}_k$ are assumed to be white, mutually independent and independent of the initial state $\mathbf{x}_0$. The possibly nonlinear functions $\mathbf{f}_{k-1}(.)$ and $\mathbf{h}_k(.)$ are assumed to be known. In the following the pdf's of $\mathbf{w}_k$ and $\mathbf{v}_k$ are denoted $p(\mathbf{w}_k)$ and $p(\mathbf{v}_k)$, respectively. The aim of filtering is to find the conditional probability density function (posterior)

$$p(\mathbf{x}_k|\mathbf{y}_{1:k}),$$

where $\mathbf{y}_{1:k} \triangleq (\mathbf{y}_1,\ldots,\mathbf{y}_k)$. The posterior can be determined recursively according to the following relations [34].

Prediction (prior):

$$p(\mathbf{x}_k|\mathbf{y}_{k-1}) = \int p(\mathrm{x}_k|\mathbf{x}_{k-1})p(\mathbf{x}_{k-1}|\mathbf{y}_{1:k-1})\,\mathrm{d}\mathbf{x}_{k-1}, \quad (3)$$

Update (posterior):

$$p(\mathbf{x}_k|\mathbf{y}_{1:k}) = \frac{p(\mathbf{y}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{y}_{1:k-1})}{\int p(\mathbf{y}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{y}_{1:k-1})\,\mathrm{d}\mathbf{x}_{k-1}}, \quad (4)$$

where the transition pdf is $p(\mathbf{x}_k|\mathbf{x}_{k-1}) = p_{\mathbf{w}_k}(\mathbf{x}_k - \mathbf{f}_{k-1}(\mathbf{x}_{k-1}))$ and the likelihood is

$$p(\mathbf{y}_k|\mathbf{x}_k) = p_{\mathbf{v}_k}(\mathbf{y}_k - \mathbf{h}_k(\mathbf{x}_k)). \quad (5)$$

The initial condition for the recursion is given by the pdf of the initial state $p(\mathbf{x}_0|\mathbf{y}_{1:0}) = p(\mathbf{x}_0)$. Point estimates can be computed from the posterior distribution, e.g. posterior mean.

In general and in the cases analyzed in this paper, the conditional probability density function cannot be determined analytically. Because of this, there are many approximative methods to compute the posterior mean (see e.g. [35]). Besides the posterior mean these methods generally yield also a posterior covariance matrix. In case the true state is known we can check whether the posterior is consistent with respect to a consistency test, such as the Gaussian consistency test [36, p. 235 ff.]. The idea of those tests is to assess the accuracy of the state estimate's covariance matrix. For example, in the Gaussian consistency test, for a risk level of 5%, the posterior is said to be consistent if the posterior covariance matrix $\mathbf{P}_k$ and the posterior mean $\hat{\mathbf{x}}_k \in \mathbb{R}^2$ fulfill the inequality $(\hat{\mathbf{x}}_k - \mathbf{x}_k)^T \mathbf{P}_k^{-1}(\hat{\mathbf{x}}_k - \mathbf{x}_k) \le \chi_2^2 = 5.9915$, where $\mathbf{x}_k$ is the true state.

IV. COVERAGE AREA MODELS

A. Coverage Area Estimation

A computationally light method for parametric fingerprinting is proposed in [20, 29]. In order to reduce the size of the FP radio map the authors represent the coverage area (CA; aka reception region) of any communication node (CN; called AP in WLANs or base station in cellular radio networks) by an elliptical probability distribution, for which the distribution's parameters and the location estimates are solvable in closed form. The probability distribution represents only the region in which a signal from the AP can be received; other than an implied reception strength threshold, it gives no information about the RSS. Although real CAs are often irregularly shaped, modeling them with simple shapes makes it possible to keep the database for storing CAs compact while providing acceptable accuracy. Thus, it enables fast transmission to a UE [20, 29] and fast computation of the UE's position, since any ellipse can be represented by five parameters: three parameters for the origin-centred ellipse (2-by-2 symmetric positive definite matrix), and two more parameters to specify the location of the ellipse centre [37]. Koski et al. [29] and Piché [37] point out that also other shapes, such as circles (three parameters) or polygons (at least six parameters) for modeling CAs of CNs could be used. Furthermore, the 3GPP TS 23.032 standard supports the use of geometrical shapes such as ellipses, polygons and ellipsoids [4, p. 98].

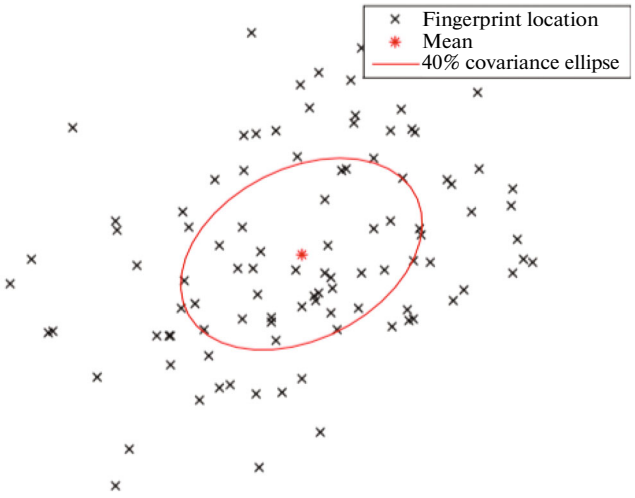

Fig. 1. Fingerprint locations $\mathbf{z}$ in which an AP's signal is received and the mean and covariance of a fitted bivariate Gaussian.

The coverage area is modeled in [20, 29] by a posterior distribution for the ellipse parameters $\boldsymbol{\theta}$ given the FP locations $\mathbf{z} = \{\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_n\}$ where the CN was heard. The distribution is given by Bayes' rule

$$p(\boldsymbol{\theta}|\mathbf{z}) \propto p(\mathbf{z}|\boldsymbol{\theta})p(\boldsymbol{\theta}), \quad (6)$$

where the likelihood and prior pdf are Gaussian. In other words, the CA is modeled by fitting the mean and covariance of a multivariate Gaussian to the data. Figure 1 shows the FP locations $\mathbf{z}$ as well as the mean and covariance of a fitted multivariate Gaussian, whose parameters are contained in $\boldsymbol{\theta}$ (for the mean and for the symmetric covariance matrix). The ellipse

contains approximately 40% of the bivariate Gaussian's probability.

Using a Bayesian formulation of the regression problem has two advantages. Firstly, the use of the Bayesian prior pdf $p(\boldsymbol{\theta})$ allows one to exploit information about "typical" coverage areas, which is crucial when only a few FPs are available [20, 29]. Such information is available through experimental studies. For example, the typical reception range for WLAN in indoor environments is 20–50 m [6, p. 9]. According to Trevisani and Vitaletti [3] and Molisch [14, p. 17 ff.] the size of a CA in a cellular network depends strongly on the cell type; it may range from 1 m for nanocells to 30 km for macrocells. Furthermore, Trevisani and Vitaletti [3] point out that the size is influenced by a variety of factors, such as interferences, local expected traffic and sensitivity of the UE's antenna. This variability can be also modeled within the Bayesian prior pdf, by using a distribution with a larger variance.

Secondly, using Bayes' rule with the independence assumptions given for the SSM (2) for finding the ellipse's parameters allows recursive estimation and updating of estimates [29]. Updating the posterior $p(\boldsymbol{\theta}|\mathbf{z})$ as new FPs become available can be done by either using Bayes' rule [38, p. 14 ff.] or by computing it as time series [38, p. 29 ff.]. Koski [38] points out that the latter approach enables one to take into account that the parameters can change in time. Such changes are common and can be caused, for example, by constructing new buildings, changing floor plans, or modifying radio network topologies [4, 6, 14, 37, 38].

One possible point of criticism for the method above is the assumption that [20, 29] model location reports having a Gaussian (Normal) distribution. It is common knowledge that the Normal regression model lacks robustness, in the sense that outliers can cause CAs being estimated too large (see e.g. [39]). Reasons for outlier location reports include unusual reception conditions, software or hardware malfunctions in GNSS (when using it for determining the coordinates of the location reports) or radio signal reception [37]. Piché [37] argues that, while gross outliers can easily be detected by heuristics, "moderate" outliers are hard to discover, especially if the list of FPs used for CA determination contains a large amount of them. Thus, he recommends to rather model location reports as having Student-t distribution so that outliers are automatically accommodated by the distribution's heavier tails. In [37] he shows how the Student regression can be computed by Gibbs sampling [40] or Expectation Maximization (EM) algorithm [41]. For both algorithms it is shown in [37] that it is also for Student-t distributed location reports possible to include information on "typical" CAs via an informative prior pdf, as in the case of Gaussian distributed location reports. Introducing such a prior pdf requires only minor changes in EM or Gibbs sampler algorithm.

The methods considered above ignore completely the RSS/RSSI values corresponding to IDs of heard CNs. Hence they are less sensitive to changes in the radio environment than fingerprinting methods that use these values. This gain in robustness, however, comes generally at the cost of lower accuracy compared with nonparametric fingerprinting methods (e.g. WKNN), which besides FP locations and IDs of CNs observed in each FP often also store corresponding RSS or RSSI values. The RSS values, and therefore also the RSSI values, depend on the distance between CN (emitter aka transmitter) and UE (receiver) and are commonly modeled as function of this distance using path loss (PL) models, which will be discussed in section V.

A coverage area method that uses RSS information is proposed in [20, 39]. Instead of storing only one CA per CN in the database, several CAs per CN are stored, which are modeled from FP data that is grouped according to RSS. In [39] the authors examine the use of one, two and three CAs per CN assuming both Gaussian and Student-t distribution for location reports. FPs are grouped based on their RSS values and different CAs are generated using only location reports of their corresponding group. Three different grouping rules are considered: RSS-level, n strongest CNs of each FP and x% strongest CNs of each FP.

B. Positioning Using Coverage Areas

A position estimate for a UE using coverage areas [20, 29, 39] can be obtained by applying Bayes' rule. The position estimate and an uncertainty measure of the estimate can be extracted from a Gaussian posterior probability density function $p(\mathbf{x}|\mathbf{c})$ of the UE position $\mathbf{x}$ given a list $\mathbf{c} = (c_1, c_2, ..., c_N)$ of CNs observed by the UE in the current position. For the conjugate (i.e. Gaussian) prior pdf of this position, a suitable mean and covariance, which represent prior knowledge on UE's position, should be chosen. In case such information is unavailable, setting the covariance very large is justified [20]. For computing the likelihood $p(\mathbf{c}|\mathbf{x})$ [20, 29, 39] it is assumed that prior probabilities of observing c_n are equal for all $n = 1, ..., N$ and that observations are conditionally independent given $\mathbf{x}$.

The latter assumption is a weakness of the model [20]. For neighboring CNs c_i and c_j, for example, the independence assumption is often violated. If c_i is observed in $\mathbf{x}$ then this clearly affects the probability of c_j being observed, since CAs generally overlap, meaning that both CNs can be observed in the same area. Koski et al. [20] speculate that using information on CNs not being observed in $\mathbf{x}$ might improve positioning accuracy significantly. However, because closed-form solutions are unavailable for such data, such an approach could not compete with the proposed algorithm in terms of computational complexity.

Koski et al. [29] point out that WLAN APs (i.e. CNs) are often hearable in an entire building, which results in their coverage areas being useless for positioning. Therefore, eliminating some CNs in both learning and positioning phase may actually increase position accuracy [29], [42, p. 71 f.]. A variety of techniques have been examined for selecting CNs to be pruned, including forward selection and backward elimination, weighting CNs using Generalized Cross Validation, selection of CNs based on information gain, divergence measures or discrimination score. Readers are referred to [43] for a more extensive overview.

In [29] and [44] the authors suggest usage of a signal strength threshold value for reducing the number of observations. CNs heard with an RSSI below this threshold are eliminated from the FP, i.e. are not used for CA determination or positioning. To ensure comparability of RSSI values from different devices the authors in [29] apply the RSSI histogram approach mentioned in Section I. For the positioning using WLAN measurements, positioning mean errors and root mean square errors were reduced significantly by applying this technique. Instead of the histogram approach other calibration methods, which were explained earlier, could be used. However, when using signal strength based elimination, it has to be ensured that CN elimination does not decrease consistency of position estimates and/or overpruning, in which (almost) all CNs are eliminated from the observation list. The latter problem might occur, for example, in outdoor environments for WLAN measurements, where signal strength values and their dynamics are considerably weaker than in indoor environments since the signals generally must pass through thick walls.

This behavior might also explain why using multiple CAs per CN for WLAN-based positioning in outdoor environments provides only small improvements compared with using one CA per CN [39], whereas in indoor environments it yields significantly better accuracy [20, 39]. Alternatively, overpruning might be avoided by using the n strongest rule or the $x\%$ strongest rule, neither of which require RSS calibration.

V. PATH LOSS MODELS

Path loss (PL) models refer to models of the signal power loss L_P or the received signal strength P_{RSS} along a radio link, averaged over large-scale and small-scale fading [14, p. 127]. In the simplest models the PL depends only on the transmit power and the distance d a radio wave travels; more complex models take further factors into account. For an overview of propagation mechanisms and PL models we refer the reader to [6, 14, 45] and references therein.

The relation between the RSS and the radio wave's traveled distance can be used for positioning. From RSS measurements and PL models the distances between a set of reference nodes and the target node are estimated, which then enables estimation of the target node's position. However, the position estimate is sensitive to signal noise and PL model parameter uncertainties because the distance-power gradient is relatively small [1]. Consequently, these estimates are generally less accurate than radio-signal based estimates that are derived using AOA (angle-of-arrival) or time delay measurements. However, Patwari et al. [46] show that for sufficiently high CN-density positioning algorithms relying on PL models (and thus on RSS) can achieve similar performance as time delay based algorithms.

A. Parameter Estimation for PL Models

Earlier studies assumed the parameters of the PL model to be known a-priori, which is an oversimplification for several real-world applications and therefore ill-suited [47]. Thus, the model's parameters should be estimated based on FP data consisting of CN-IDs and corresponding RSS values.

There are various approaches on how to estimate the model parameter(s). Some methods first estimate the CN's position or assume it to be known, and then estimate the parameter(s) using the CN position (e.g. [48, 49]); others estimate the CN's position and the PL model parameter(s) simultaneously (e.g. [47, 50]).

In [48] a statistical modeling approach is introduced, in which the UE's position is estimated using a statistical signal power model, and the CN position is assumed to be known. The parameters of the model are estimated using the EM algorithm [41] to find their maximum likelihood values. The authors point out that the main difference of their algorithm, compared with the geometric approach, is that it infers the signal properties from the location. However, often the positions of CNs are unknown and have to be estimated as well. Li [47] found that estimating the CN location alone for fixed PL model parameters can result in large errors when the values for the parameters are chosen inaccurately, and he recommends simultaneous estimation. Furthermore, he points out that joint estimation removes the necessity of extensive channel measurements.

In [47] Li estimates CN position and the PL exponent n (aka distance-power gradient) of the classic narrowband radio propagation PL model[3]

$$L_P(d) = L_P(1) + 10n\log_{10}(d) + w, \qquad (7)$$

where $L_P(d)$ is the signal power loss in dB at a distance of d meters from the CN. The zero-mean Gaussian random variable w with variance σ_w^2 is used for modeling the shadow fading (aka slow fading). The approxi-

[3] Roos et al. [48] use a more complex PL model that includes a parameter for the transmission direction. Since their focus is on cellular networks in which the CNs (aka base stations), in general, have directional antennas this should provide more accurate PL estimates. For isotropic WLAN APs, as used by Li [47], the PL should (theoretically) be the same in all directions.

mately lognormal distribution of shadow fading, which implies Gaussian distribution of w in (7), has been empirically observed, for example, in [51–54]. For the estimation Li [47] applies the Levenberg-Marquardt method, a modified Gauss-Newton (GN) algorithm, on a system of nonlinear equations.

Nurminen et al. [50] go a step further and estimate in addition the apparent transmission power $A = P_{\mathrm{RSS}}(1)$ for a version of the log-distance model, namely

$$P_{\mathrm{RSS}}(d) = A - 10n\log_{10}(d) + w, \qquad (8)$$

using the Iterative Reweighted Least Squares (IRLS) method, which is also a GN algorithm. According to Dil and Havinga [55] (8) can be used for describing P_{RSS} dependency of distance d in any indoor environment. The Bayesian approach used in [50] furthermore lends itself well to update the estimate of CN position and PL model parameters as new fingerprint data becomes available.

The algorithm uses uninformative Gaussian prior pdf. Given enough fingerprints, according to Nurminen et al. [50], one can choose the valid prior mean values for the PL model parameters arbitrarily, since for large numbers of FPs the posterior distribution is typically unimodal. This is supported by Li's [47] finding that the effect of an inaccurate prior for the PL exponent n on the estimation results is negligibly small. However, he stresses that, especially for cases with limited data, a well-chosen informative prior would be beneficial. Various studies yielded values for the PL exponent for different environments and networks (e.g. [45, 56]); for the apparent transmission power fewer studies are available (e.g. [56]).

For the prior CN position more care should be taken in order to prevent IRLS placing the CN in an area of weak RSS values [50]. Nevertheless, even with such measures it cannot be guaranteed that the algorithms find the correct CN position. For example, the RSS map might contain several peaks or the true CN position could be outside the RSS map or it may be that too few measurements are available for determining a clear peak [50]. Both the (Bayesian) IRLS method in [50] and the Levenberg-Marquardt method in [47] give the user a tool to distinguish between reliable and unreliable position estimates and PL models in the form of a covariance matrix. For the latter approach the covariance can be computed once the optimal estimates are found, in the former it is automatically available as posterior covariance matrix. Furthermore, the approach in [50] accounts for correlation in measurement errors by adding a small constant diagonal matrix for the CN position's covariance matrix (Li [47] assumes all observations to be statistically independent). The cross-correlation between CN position and PL model parameters is, however, neglected, mainly to limit the number of parameters.

One possible point of criticism of the methods in [47, 50] is that the authors assume the standard deviation σ_w of the shadow fading component to be fixed, although [47, 50] stress that it is highly dependent on the propagation environment, which is confirmed, e.g., by Ghassemzadeh et al. for UWB networks [54]. Typical values vary between 1 dB and 6 dB [56, 57] for WLAN. Larger values have been observed, especially in larger buildings (see e.g. [57, p. 139 ff.] for more values). In cellular networks values between 5 dB and 16 dB have been observed [52, 56]. Nurminen et al. [50] use fixed $\sigma_w = 6$ dB, whereas Li [47] studies the influence of varying values that are fixed during the estimation on the errors in CN position estimates. His tests show that the value of σ_w can influence the bias and the efficiency of location estimators significantly depending on the used estimation method.

In [18] Han et al. ignore the shadow fading component's standard deviation, and compute point estimates for the parameters of their PL model. However, their approach is worth consideration because it allows generating a radio map using significantly less FP data. Furthermore, their PL model, which is an extension of Seidel's model [58], takes into account the angles between signal path and obstacles. This means, the modeled PL depends on the angle in which the signal hits the obstacle. For example, if the signal hits a wall in a 90 degree angle the distance it travels through the wall is significantly shorter than if it hits the same wall in a 60 degree angle. The PL model in [18] accounts for this.

For generating the radio map, the authors collect FP data only for a small fraction of the area that should be covered by the radio map. Those FPs are then used to establish a linear equation that can be solved using a least-squares method to obtain the estimates for the PL model parameters. They then use the PL model to generate FPs for those locations in which no measurements were taken. However, this second step will be unnecessary if parametric methods are used for positioning, as will be seen in the next subsection. In their paper Han et al. show that it suffices to collect FPs from 20% of the grid points for which FPs will be available in the radio map. In their test the radio map using those points and their method for generating FPs for the other 80% grid points has a similar cumulative distribution function (cdf) for the predicting error as the radio map for which in all grid points FPs were collected.

Shrestha et al. [19] use deconvolution-based methods to reduce the radio map size by a factor of ten compared with the map for nonparametric approaches. Besides PL model (8), they study also a multi-slope PL model, which takes different values for n depending on the distance between transmitter and receiver. Furthermore, the authors extend both models with an additive floor loss parameter, and consider 3-dimensional positions.

The parameters of the PL models are estimated in [19] as follows. For each FP measurement parameter estimates are computed using either least square (LS), weighted LS or minimum mean square error (MMSE) assuming the FP's location as AP location. Then the

expected RSS values in that location given the current parameter estimates is computed. The FP's location for which the mean square error (MSE) of measured and expected RSS values is smallest is used as AP position estimate. Alternatively, the average of k FPs that give lowest MSE could be chosen. In this case, the PL model parameters have to be recomputed using the new AP position estimate.

B. Positioning Using PL Models

Once the parameters of the PL model and the positions for all CNs are estimated, range estimates can be derived using the PL model and measured RSS values. For computing the UE position estimate subsequently trilateration or some other nonlinear estimation technique can be used.

Nurminen et al. [50] test three different methods that use the PL model (8) with real WLAN data in an indoor office environment: a grid method where Monte Carlo integration is used for computing the likelihood in each point of a spatial grid, the Metropolis-Hastings algorithm, and the IRLS. For comparison the authors apply the CA method presented in [20, 29], filtered using standard Kalman filter, and WKNN with $k = 3$ and unfiltered measurements. Within the tests the floor is assumed to be known. Furthermore, grid method, Metropolis-Hastings algorithm and IRLS are analyzed using both point estimates and Gaussian distributions for the PL model parameter values.

When using extensive FP data for estimation of CN position and parameter estimates, WKNN provides the best accuracy, followed by grid method, Metropolis-Hastings algorithm and IRLS, which all provide similar accuracy. The CA method performs worst. However, when the FP data for some of the CNs is limited, WKNN drops to the same accuracy level as grid method, Metropolis-Hastings algorithm and IRLS. This is in accordance with earlier findings. For example, Dil and Havinga [59] find that in the case of limited FP data nonparametric FP positioning algorithms, such as WKNN, are outperformed by range-based algorithms.

The tests in [50] show that assuming Gaussian distributions for the parameters rather than point estimates is, in general, beneficial; the advantage of using a distribution becomes clearer in the tests with limited FP data. These results do not come as a surprise, since the PL model contains approximation errors [60]. If less FP data is available for estimating the PL model those errors, in general, are larger. Therefore, in such situations it should be beneficial, from a theoretical point of view, to assume more uncertainty in the parameter estimates.

Another possible explanation why assuming a distribution for the PL exponent gives better results than assuming point estimates is that the PL exponent n can be assumed constant only for a limited time in an environment [47]. However, even if the value changes it should still be close to the previous value, as long as the environment stays the same. This can be captured to some extent by assuming some uncertainty in the PL exponent estimate. In addition, storing the uncertainties also enables updating the parameters recursively and using time evolution models when new FPs become available.

In terms of computational demand, the grid method and the Metropolis-Hastings algorithm have no edge compared with the WKNN, whereas the IRLS is significantly faster and achieves running times close to those of the CA method.

As mentioned in the previous subsection, Han et al. [18] use their PL model and some FP data for generating the full radio map of a certain environment. For positioning they then use the (non-parametric) WKNN. Their test shows that positioning accuracy of their approach is slightly worse, but generating the radio map is significantly faster. It could be even faster if the authors would simply store the parameters of their PL model for each CN in their radio map and derive in the positioning phase ranging estimates from the PL model and the RSS values, and then use some nonlinear estimation technique to obtain a UE position estimate.

Shrestha et al. [19] go the same way as Han et al. [18], generating a full radio map from the PL parameters in the positioning phase. Their idea is to create an artificial grid and compute for each grid point signal strength differences between the UE's measured and the expected RSS values in that grid point. The position estimate can then obtained by using a version of the NN method. In their paper the authors use the KNN with $k = 4$. For three of four test buildings the 4NN using the full radio map (i.e. original FP data stored in radio map) outperforms the 4NN using artificial radio map generated from the radio map containing only PL model parameters, but differences are rather small. The best option for computing PL model parameter estimates in the offline-stage is the MMSE, according to the tests. Furthermore, PL model (8) gives the best trade-off between positioning accuracy and computational complexity.

Again, it would be worth to skip the radio map reconstruction step and simply use the PL parameters and measured RSS values to compute range estimates and then apply some nonlinear estimation technique. However, when using the multi-slope PL model using the correct estimate for the PL exponent could be tricky.

Accuracy levels of PL model-based positioning could be improved further by replacing the isotropic PL models that were considered so far with anisotropic PL models. It is well known that channel characteristics in different directions from the CN differ even for omni-directional antennas due to varying environments (e.g., a WLAN AP mounted in corner of a room). Furthermore, in practice directional antennas are widely used since they decrease interference with other systems [14], and allow significantly higher data

throughput and range extension [61]. Thus, using different PL model parameters for different directions might be beneficial in terms of accuracy. However, once more this comes at the cost of a larger database. Storing PL model parameters for two directions already doubles the size of the radio map. In addition, it complicates the positioning phase since one either has to decide which of the parameters to use or has to compute an (weighted) average of the parameters.

VI. GAUSSIAN MIXTURES AND RELATED APPROACHES

A known disadvantage of the CA approach discussed in Section IV is that most of the probability mass is located near the center of the ellipse that is used for describing a CN's coverage area. However, for weak signals the UE is more likely to be close to the edge of the CA. Therefore, CAs yield in such cases rather poor estimates in the positioning phase [20]. In the previous section we looked at approaches that address such issues by taking into account the RSS in addition to the CN-ID by using PL models. Alternatively, we could apply Gaussian mixture (GM) models (aka Gaussian sum models).

A Gaussian mixture is a convex combination of Gaussian density functions $\{\mathrm{N}\}(\mathbf{x};\boldsymbol{\mu},\boldsymbol{\Sigma})$, namely

$$p(\mathbf{x}) = \sum_{n=1}^{N} \omega_n \{\mathrm{N}\}(\mathbf{x};\boldsymbol{\mu}_n,\boldsymbol{\Sigma}_n), \tag{9}$$

where weights ω_n are nonnegative and sum to one. The main motivation behind GM and filters based on it is that any density function can be approximated, except at discontinuities, by a convex combination of Gaussian densities arbitrarily closely [62–64]. That is, as the number of Gaussian components within the GM increases and the norm of all covariance matrices approaches zero the approximative density function converges uniformly towards the desired density function ([62], Lo (1969) and Alspach (1970) cf. [65]). Unlike other approximation techniques such as Gram-Charlier and Edgeworth expansions, a GM is a valid density function itself [62, 65].

Sorenson and Alspach [62] point out that the approximation quality depends not only on the number of components in the GM but also on their placement. Furthermore, they stress that there is no obvious way to choose the parameters of the components due to the GM's lack of orthogonalizability. Thus, they suggest choosing them so that either the L^k-norm of the approximation error is minimized or the GM approximation matches some of the moments of the true density exactly. In addition, they point out that assigning the same covariance to all components eases the computational demand significantly. However, for many cases the latter idea yields poor approximations.

A. Representing FP Data Using GMs

For reducing the size of the FP radio map a more sophisticated method than the single Gaussian coverage area approach presented in [20, 29] that uses signal strength values is proposed by Kaji and Kawaguchi [44]. They suggest representing a CN's RSS distribution as a GM model. Although this approach generally will require more data to be stored in the radio map than the CA approach of Section IV, it should still require considerably less storage compared with traditional FP databases.

In their algorithm the collected FP data is first transformed into a point distribution, where the point density depends on the signal strength received in a FP (the higher the RSS or RSSI value the higher the point density). Then the parameters of the GM model, namely mean values $\{\boldsymbol{\mu}_n\}_{n=1}^{N}$, covariance matrices $\{\boldsymbol{\Sigma}_n\}_{n=1}^{N}$ and component weights $\{\omega_n\}_{n=1}^{N}$, are optimized by EM [41]. Kaji and Kawaguchi point out that their approach allows updating the GM models as new FP data becomes available. They do not provide an equation or rule for determining the number of components N. In our tests in Section VII we use $N = \max([K/40],8)$, where K is the number of FPs in which the specific CN is heard.

A different approach, developed for robot localization, is introduced by Koshizen [66]. The approach is called GM Bayes' with regularized expectation maximization (GMB-REM). In the offline phase of this algorithm fingerprints containing measurements from sensors are collected at various positions within a grid. Then for each position $\mathbf{x}$ in the grid the conditional density function of sensor measurements $\mathbf{y}$ given $\mathbf{x}$ is computed as a regularized GM model. The parameters of the components are chosen such that they maximize the log-likelihood function given the (offline) training data using EM. In the positioning phase the likelihood of each position $\mathbf{x}$ is multiplied with the prior pdf and then renormalized in order to obtain the posterior. In [66–68] the authors introduce further techniques for sensor selection and a new sensor fusion system for the GMB-REM. However, the algorithm suffers from the common drawbacks of FP databases, and its application on large scale is unpractical.

B. Positioning Using GM Models

In positioning tasks the system (2) is often non-Gaussian and/or significantly nonlinear (see e.g. [69] for a criterion for significant nonlinearity). Therefore the Bayesian recursion is generally unsolvable in closed form [62]. Applying an EKF to solve such generally multimodal systems has the disadvantage that it follows a single peak of the pdf, meaning that it gives rather a maximum likelihood estimate than a minimum variance estimator [65]. Besides computationally demanding methods such as the particle filter (see e.g. [70]), also GM-based filters can be used to achieve

excellent performance for significantly nonlinear and/or non-Gaussian systems.

In [65] it is stated that already in 1965 Aoki suggested to approximate the posterior as GM. Ali-Löytty [71] introduces such an approximation that he calls efficient GM filter (EGMF). His method uses parallel planes to split the state space into pieces and then approximates the posterior in every piece by one Gaussian. Ali-Löytty shows that, unlike most other GM filters, the EGMF yields optimal results in the sense of mean and covariance in the linear case. Furthermore, he finds that the EGMF provides better accuracy than traditional Kalman-type filters (e.g. EKF, UKF) and the sigma point GMF (SPGMF) [72], while requiring fewer components than the SPGMF. If the prior pdf follows a Gaussian distribution, then the EGMF's number of components is equal to the dimension of state variable **x**. One disadvantage of the SPGMF is its need for analytical differentiation. In order to avoid such differentiation, Raitoharju and Ali-Löytty [73] propose the adaptive splitting (AS) method. This method first finds the direction of maximal nonlinearity within a Gaussian prior pdf. If the measurement's nonlinearity is significant with respect to the criterion proposed by Ali-Löytty and Sirola [72], then the Gaussian component describing the measurement is split into a mixture of two Gaussians. The splitting is repeated until none of the components shows a significant nonlinearity anymore. The results in [73] suggest that the proposed method requires fewer components than SPGMF while providing better approximation of the reference pdf.

A further important fact is that assuming Gaussian distribution of the prior pdf is not always feasible. It is obvious that, when looking into filtering, for a posterior at time k that is described by a GM, the prior pdf at time $k + 1$ should be also a GM. In general, if the initial state is not Gaussian and/or in highly nonlinear situations we should apply a bank of Gaussian filters, namely a GM filter (GMF), for solving the problem [72]. Anderson and Moore [74, p. 212] and Ali-Löytty [75] show that if the GM approximation of the prior pdf converges towards the true prior pdf as the number of components increases while their covariances decrease, then the GM approximation of the posterior converges towards the true posterior as well. Lo [64] presents an application of GMs for filtering a system with linear dynamics and arbitrarily distributed prior pdf and some examples, which provide an excellent introduction to the concept.

Care has to be taken to limit the number of components in the GM; this has been mentioned already when GM was introduced [62, 65]. Sorenson and Alspach [62] suggest to either merge components with approximately equal means and covariances or drop components with sufficiently small weights on the GM (called forgetting). Alternative methods are introduced, for example, in [69, 76, 77]. For a more detailed and broader overview on component reduction methods we refer the reader to [78, 79] and references therein.

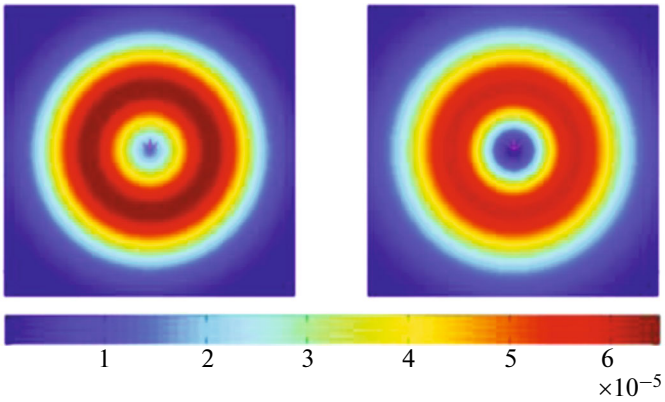

Fig. 2. Normalized exact likelihood (left side) for measurement of an isotropic CN (magenta asterisk inside the ring) and its approximation yielded by GGM (right side).

In addition to reducing the number of Gaussian components it would be beneficial to already keep the number of components small in the approximation phase. In [80] Müller et al. therefore propose a generalized version of GM (GGM) that relaxes the nonnegativity restriction on component weights, and call it Gaussian mixture allowing negative weights (GMA). For the isotropic ranging model, which is considered in [80], the measurement likelihood has a ring-shaped pattern as shown in the left plot of Fig. 2. While a traditional GM would require a large number of components, due to the infinite number of peaks of the likelihood, the GGM yields a satisfying approximation with only two components (see Fig. 2 in the right plot), one having positive and one having negative weight. Special care has to be taken when assigning the weights of the components to ensure that the resulting likelihood is everywhere nonnegative and thus a valid likelihood. In the filter based on GGM the authors collapse the GM posterior at each time step using moment matching, since the reduction methods mentioned previously generally are only suitable for GMs with nonnegative component weights.

Müller et al. apply the GGM in [80] for positioning in cellular networks, and their results indicate that the GGM outperforms both single time-step EKF and the Gaussian CA approach [20, 29] in terms of accuracy and consistency. The filtered version of the GGM (GGMF) also outperforms the EKF and the CA-based filtered approach. In [81]) those findings are confirmed for positioning in an UWB network.

Ali-Löytty and Sirola [69] perform extensive simulations of GM using both cellular measurements and measurements from a Global Navigation Satellite System (GNSS), i.e. hybrid positioning. Their results suggest that only multimodal likelihoods should be approximated by a GM.

Table 1. Data set sizes. Some APs could be heard on several floors and/or in both buildings

Building	Floor	APs	FPs	TPs
1	1	200	889	19
1	2	289	243	47
1	3	212	160	22
2	1	154	1530	168
2	2	186	1582	33
2	3	148	333	19

C. Alternatives to Gaussian Mixtures

All the methods considered in this section so far could be significantly degraded by the previously mentioned sensitivity of Gaussian regression to outliers. Bishop and Svensen [82] point out that this sensitivity might result in an overestimation of the number of required Gaussian components. They propose a Bayesian approach for mixture modeling based on Student-t distributions, which is more robust to non-Gaussianity in the data (McLachlan and Peel [83] make the same proposition). The major drawback of using Student-t distributed components is that, contrary to using Gaussian distributed components, no closed-form solution for the likelihood maximization exists [82, 83]. However, as [37, 82] show, any Student-t distribution can be represented as an infinite mixture of scaled Gaussians. Therefore, EM can be used to find the maximum likelihood, while the computational load of the proposed algorithm [82] is only slightly larger than using the ML technique for finding parameters of GM models. However, t-mixtures have so far not been used in FP positioning.

VII. COMPARATIVE TESTING

In this section we compare the performance of several parametric fingerprinting and positioning methods described in the previous sections. We evaluated these methods by analyzing their WLAN based positioning accuracy for six test tracks located within two buildings of Tampere University of Technology. Building 1 has an area of approximately 10000 m^2 and building 2 has an area of approximately 6600 m^2; both buildings are three-story. The total number of detected APs within both buildings is 506. For two of the tracks measurements were collected several months later than for the other four tracks, which were collected at the same time as the data used for generating the radio maps. Some of the test tracks had floor changes, which were assumed to be known. The radio maps were built separately for each floor. Table 1 shows for each floor of the two buildings the number of detected APs, the number of FPs, and the number of test points (TP) for the four tracks collected at the same time as the data used for the radio maps. TPs are points on the test tracks that we positioned in our evaluation.

For comparison we implemented CA-based positioning with single CA [29] and 2-level CAs with limit –70 dBm [39], PL model [50], GGM approximation of the PL model [80] and the signal strength estimation model from [44] (denoted GMEM). In addition to these parametric methods we used a weighted k-nearest neighbors method (WKNN) with $k = 5$ as a reference. Figure 3 shows how much data

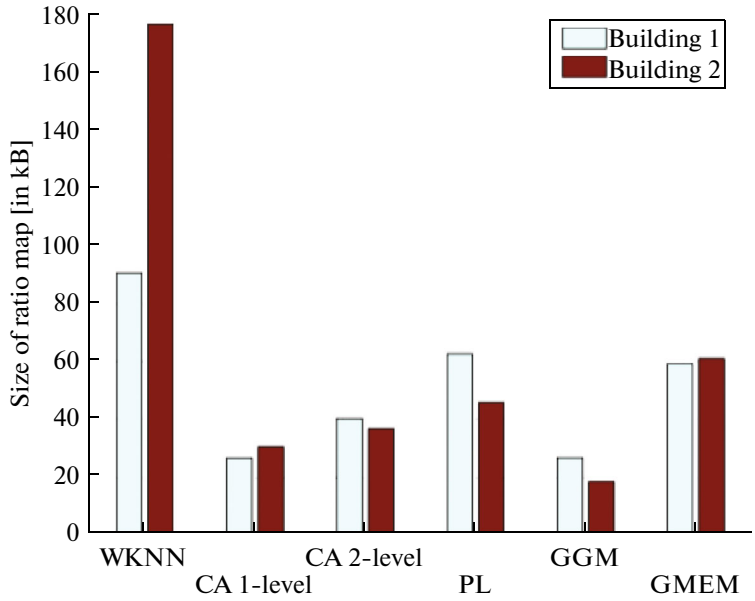

Fig. 3. Data storage requirements for radio maps for tested methods in our two test buildings.

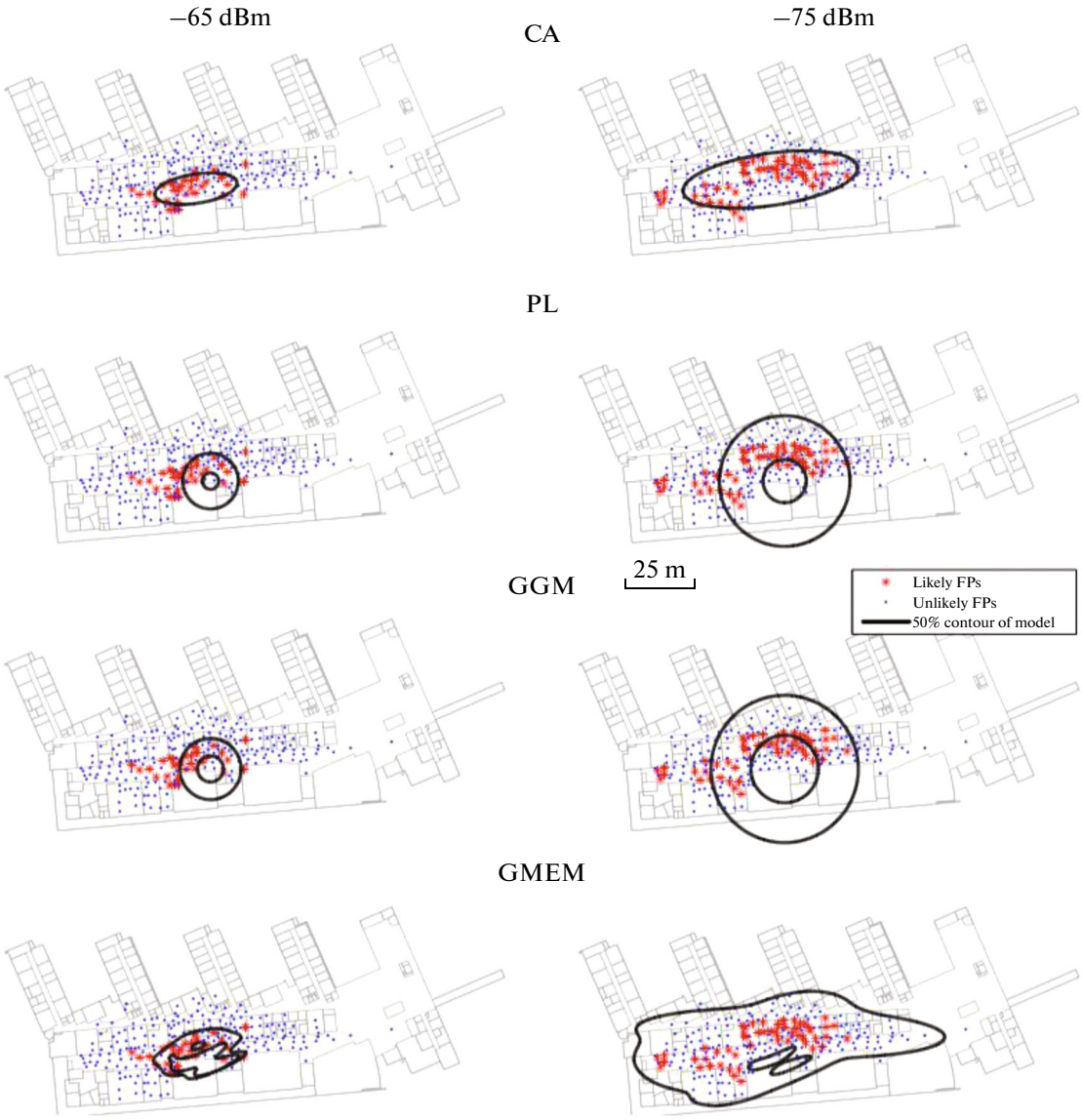

Fig. 4. Likelihoods for different models and different RSS values. Each likely FP has a higher probability mass than any unlikely FP.

storage each method requires for its radio map. The WKNN method does not summarize the FPs in any way and therefore requires the largest radio maps in both buildings. In our tests the parametric methods reduce the radio map size between 30% and 90%. However, because the size of the radio map used by WKNN depends on the number of FPs and the size of the radio map used by the other methods depend on the number of APs these numbers cannot be generalized.

Figure 4 shows the contour curves containing 50% of probability mass for all tested approaches (1-level CA is same as the 2-level model with measurement −75 dBm), except WKNN. For computing the probabilities from likelihoods we used a rectangular uniform prior pdf that covers the whole building. The FPs are similarly divided into groups *likely* and *unlikely* with likely FPs containing 50% of the probability mass. The standard deviation for RSS based methods is set to 6 dB. By visual inspection the shapes of contours are quite different except for PL and GGMF, but if we consider the numbers of likely FPs inside the contours they are similar, except for the GMEM with weak signal strength.

In Figure 5 for all APs the PL exponents estimated by Nurminen's approach [50] as a function of the number of FPs in which the specific AP was observed are displayed. The figure shows that for small numbers of FPs the PL exponents often take values less than 2 (68% of all APs that were received in fewer than 100 FPs have $n < 2$ but only 27% of all APs that were heard in more than 100 FPs have $n < 2$). A PL exponent of 2 means that the signal propagates in free space; values smaller than 2 in our tests can be

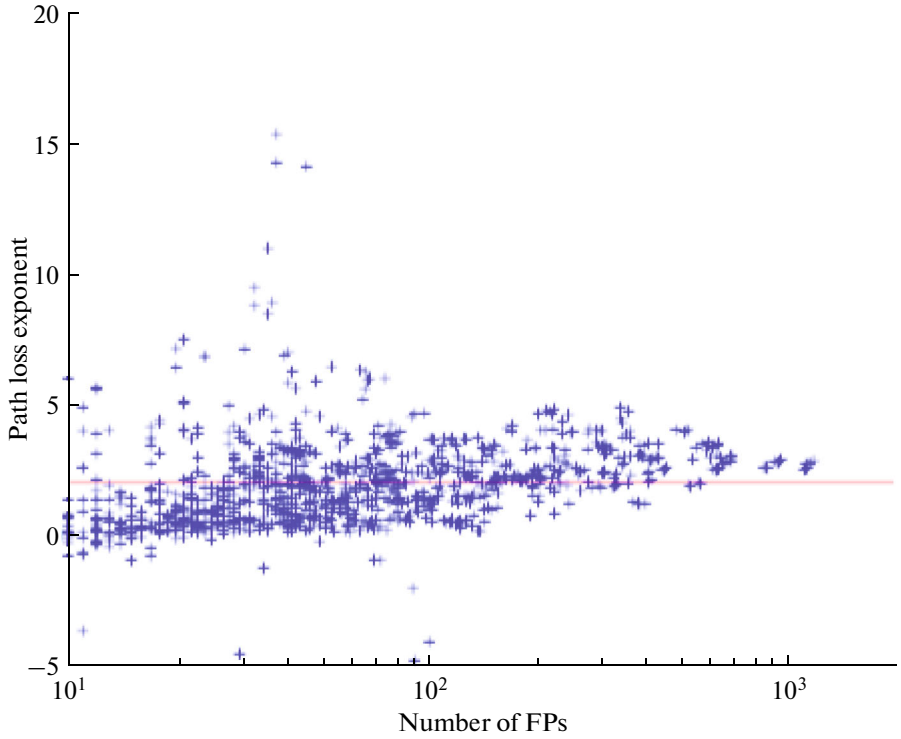

Fig. 5. PL exponent as function of FPs used in learning.

explained by the fact that the corridors in which the FPs were collected acted as waveguides [14, p. 66].

The true routes for all six test tracks were measured by clicking a map plot on a touch screen while walking and interpolating between the clicks, and were estimated for both static case and time series (i.e. filtered case). For the filtering we considered the state vector $\mathbf{x}_k$ containing location and velocity of the UE. Both CA-models and GGMF were updated using a plain Kalman filter. In addition, we collapsed the GGMF to a single component after 5 measurements and after each time step. The GMEM used a grid for static position estimation and a particle filter with 300 particles for the time series estimation; the PL model method used Gauss-Newton for static positioning and a GM filter [73] for time series. In time series the effect of parameter uncertainties varied depending on the location, and therefore was computed in the prior mean of the estimate. The WKNN was given a standard deviation of 10 meters for filtering with a Kalman filter.

For filtering we chose a linear state transition equation (2a) with additive zero-mean noise, i.e. $\mathbf{x}_k = \mathbf{F}_{k-1}\mathbf{x}_{k-1} + \mathbf{w}_{k-1}$ with

$$\mathbf{F}_{k-1} = \begin{bmatrix} \mathbf{I} & \Delta t\mathbf{I} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}, \tag{10}$$

where Δt is the measurement interval in seconds, and $\mathbf{w}_{k-1} \sim N(0, \mathbf{Q})$ with

$$\mathbf{Q} = 0.1\mathbf{m}^2 \begin{bmatrix} \frac{(\Delta t)^3}{3}\mathbf{I} & \frac{(\Delta t)^2}{2}\mathbf{I} \\ \frac{(\Delta t)^2}{2}\mathbf{I} & \Delta t\mathbf{I} \end{bmatrix}. \tag{11}$$

The measurement equation (2b) depended on the used positioning approach. For the static WKNN $\mathbf{y}_k$ contained AP-IDs and corresponding RSS values. The k strongest AP-IDs were picked and the weighted average of their locations was used as position estimate. In filtering the weighted average was used as posterior mean for the Kalman filter and the posterior covariance matrix was set to $10^2\mathbf{m}^2\mathbf{I}$.

For the CA approaches, the jth measurement $y_{k,j}$ was modeled as $y_{k,j} = \mathbf{m}_{\mathrm{ID}_{k,j}} + \in$, where $\mathbf{m}_{\mathrm{ID}_{k,j}}$ is the mean of the CA for the AP with identifier $\mathrm{ID}_{k,j}$ and $\in$ is zero-mean Gaussian with the same covariance as the AP's CA model.

In the PL method y_k contained RSS measurements. The RSS for the AP $\mathrm{ID}_{k,j}$ was modeled as

$$y_{k,j} = A_{\mathrm{ID}_{k,j}} - 10n_{\mathrm{ID}_{k,j}} \log_{10}\left(\left\|\mathbf{x}_k - \mathbf{a}_{\mathrm{ID}_{k,j}}\right\|\right) + \nu_{k,j}, \tag{12}$$

where $\|.\|$ is the Euclidean distance between UE at $\mathbf{x}_k$ and AP $\mathrm{ID}_{k,j}$ located at $\mathbf{a}_{\mathrm{ID}_{k,j}}$. This measurement model was used for both static and filtered positioning.

For the GGM approach the PL model (8) was used to derive the covariance matrices of the Gaussian components. Each measured RSS value was modeled

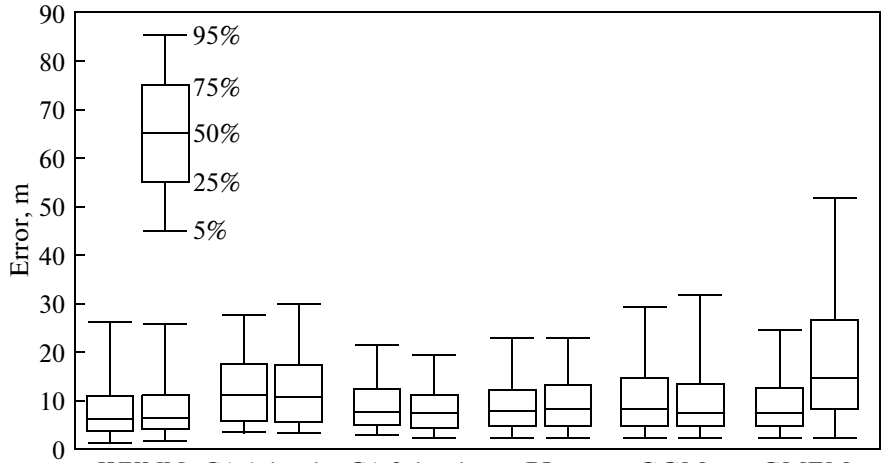

Fig. 6. Method performances with all data.

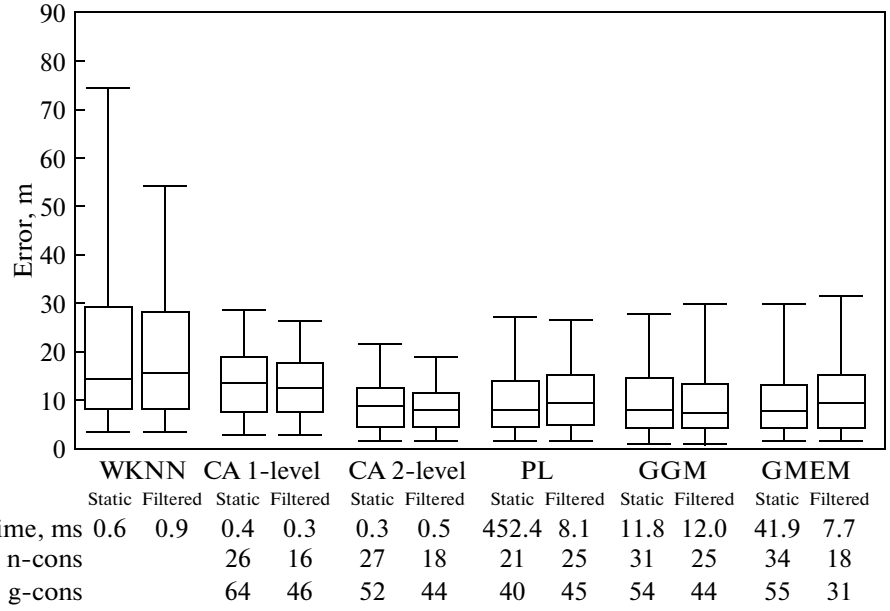

Fig. 7. Method performances with five strongest measurements.

by a GGM with two components. The GMEM modeled the likelihood of the j measurement as

$$p(y_{k,j}|\mathbf{x}_k) = p_{\{\mathrm{N}\}}(f(\mathbf{x}_k)|y_{k,j},6^2), \quad (13)$$

where $p_{\{\mathrm{N}\}}(f(\mathbf{x}_k)|y_{k,j},6^2)$ is the pdf of $\{\mathrm{N}\}(y_{k,j},6^2)$ evaluated at $f(\mathbf{x}_k)$. Function $f(\mathbf{x}_k)$ yields the RSS and is defined as

$$f(\mathbf{x}_k) = \sum_n \omega_{n,\mathrm{ID}_{k,j}} p_{\{\mathrm{N}\}}(\mathbf{x}_k|\boldsymbol{\mu}_{n,\mathrm{ID}_{k,j}},\boldsymbol{\Sigma}_{n,\mathrm{ID}_{k,j}}) - 90\,\mathrm{dBm}, \quad (14)$$

where $\boldsymbol{\mu}_{n,\mathrm{ID}_{k,j}}$ and $\boldsymbol{\Sigma}_{n,\mathrm{ID}_{k,j}}$, are the mean and the covariance matrix of the nth Gaussian component of the GM for the AP with identifier $\mathrm{ID}_{k,j}$ and $\omega_{n,\mathrm{ID}_{k,j}}$ is the component weight.

The methods were tested in four different scenarios:

- Fig. 6: full data
- Fig. 7: only the APs with five strongest signals were used for positioning

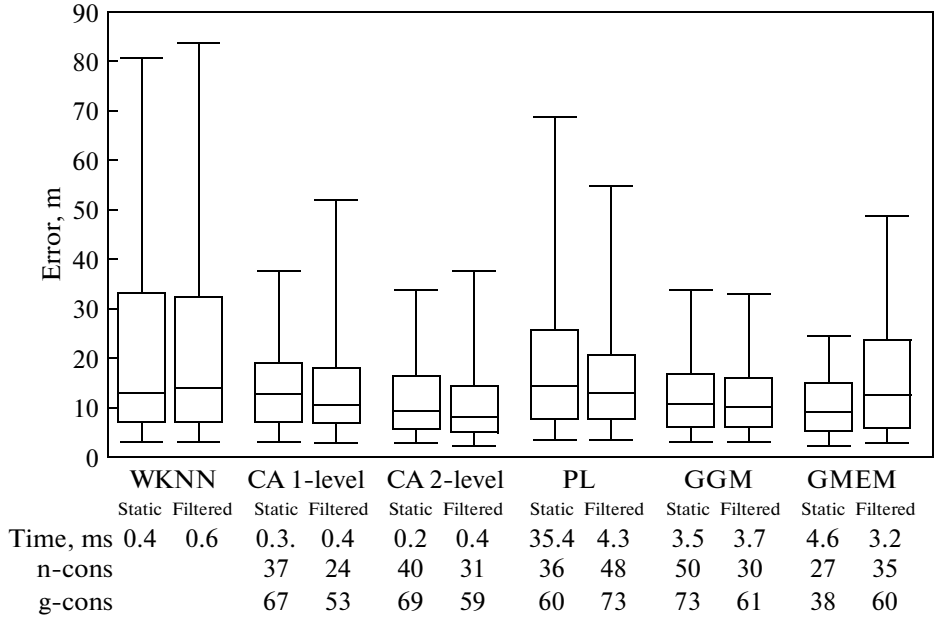

Fig. 8. Method performances with 90% of APs dropped.

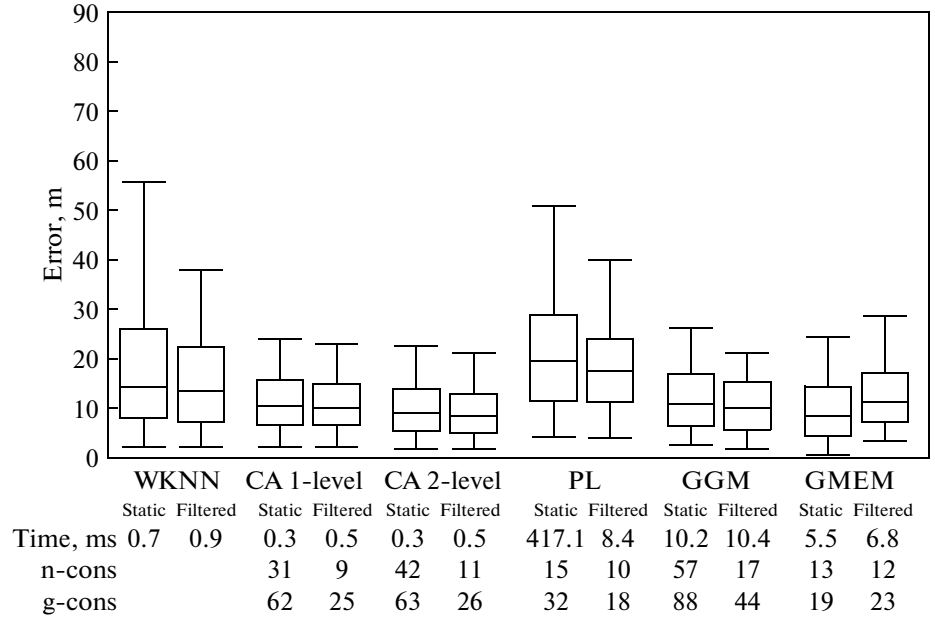

Fig. 9. Method performances when positioning done several months after data for radio map generation was collected.

• Fig. 8: 90% of APs were dropped pseudorandomly to check how the methods perform in situations with low AP density

• Fig. 9: data for generating the radio maps and data for positioning where collected with a time gap of several months to evaluate the methods' performance degradation over time

In Figs. 6–9 we present quantiles with box plots for the positioning errors, absolute time for one position estimate, and consistency values that can be used to evaluate the accuracy of the estimated position's covariance matrix that is reported by a method. The boxes show the 5%, 25%, 50% (median), 75%, and 95% quantiles of the 2D position errors. For the n-cons (normal consistency [36, p. 235 ff.]) values we assumed Gaussian distributed positioning errors, and computed how often the errors were within the 50% ellipse, i.e.

$$(\hat{\mathbf{x}}_u - \mathbf{x}_u)^T \mathbf{P}_u^{-1} (\hat{\mathbf{x}}_u - \mathbf{x}_u) \leq \chi_2^2(0.5) = 1.3863, \qquad (15)$$

Table 2. Summary of parametric fingerprint methods analyzed in this paper

Method	Radio map entries per CN	Radio map generation method	Positioning method
Coverage area (Gaussian) (1-level) [20, 29]	1 bivariate Gaussian (2 parameters for mean, 3 for covariance)	Bayes' rule using locations where CN's signal is received	Bayes' rule using CA-centres of observed CNs
Coverage area (Gaussian) (2-level) [39]	2 bivariate Gaussians (4 parameters for means, 6 for covariances)	Bayes' rule using locations where CN's signal is received	Bayes' rule using CA-centres of observed CNs
Coverage area (Student-t) (1-level) [37]	1 bivariate Student-t (2 parameters for mean, 3 for shape, 1 for dof)	EM (for MAP) or Gibbs Sampler (posterior) using locations where CN's signal is received	Bayes' rule using CA-centres of observed CNs
Coverage area (Student-t) (2-level) [39]	2 bivariate Student-t:s (4 parameters for mean, 6 for shape, 2 for dof)	EM (for MAP) or Gibbs Sampler (posterior) using locations where CN's signal is received	Bayes' rule using CA-centres of observed CNs
Path loss model Nurminen et al. [50]	2 bivariate Gaussians (1 for CN position, 1 for PLM parameters)	Iterative Reweighted Least Squares (IRLS)	Grid method using standard Monte Carlo, Metropolis-Hastings or IRLS
Path loss model Han et al. [18]	8 parameters (2 for CN position, 6 for PL model)	Solving linear equation using least-squares method	Recreate full FP radio map and WKNN
Path loss model Shrestha et al. [19]	3 parameters for CN position (3-dim.) plus (i) 2 for PL model (8) (ii) 3 for PL model (8) with floor parameter (iii) M for Mth order multi-slope PL model (iv) $M + 1$ for multi-slope with floor parameter	MSE minimisation with (a) least-square or (b) weighted least-square or (c) Minimum Mean Square Error	Recreate full FP radio map and KNN ($k = 4$)
Generalised Gaussian Mixture [80, 81]	5 parameters (2 for CN, 2 for PL model plus 1 additional)	Iterative Reweighted Least Squares (IRLS)	Bayes' rule using CA-centres of observed CNs and RSS values
Gaussian Mixture by Kaji & Kawaguchi [44]	N bivariate Gaussians for GM components	FP data transformation, parameters GM components fitted by EM	Static problem: grid approach filtering problem: particle filter
GMB-REM [66–68]	1 regularised GM for each grid point	EM for finding maximum log-likelihood of FP data	Multiply likelihood of each grid point with prior pdf to get posterior pdf

where $\hat{\mathbf{x}}_u$ is the estimated UE position, $\mathbf{P}_u^{-1}$ its covariance matrix and $\mathbf{x}_u$ the true UE position. This measure may be used for checking the error estimate in both ways (if it is too small or large) as long as the distribution is close to the normal distribution. In g-cons (general consistency [84]) we computed how often the errors were within 50% for any distribution using the modified Chebyshev inequality, namely

$$(\hat{\mathbf{x}}_u - \mathbf{x}_u)^T \mathbf{P}_u^{-1} (\hat{\mathbf{x}}_u - \mathbf{x}_u) \le \frac{2}{0.5} = 4. \qquad (16)$$

When using all of the data all parametric methods were inconsistent, with n-cons values far from the desirable 50% and g-cons values far from the 50% that can be interpreted as minimum requirement (a g-cons of 60% is not necessarily better/worse than 55%), and there are no significant differences between the accuracies of the different methods, except for filtered GMEM. The computation time for static GMEM is higher than for filtered GMEM because it is computed on a grid, whereas the filtered GMEM uses a particle filter. The results suggest that the 300 particles proposed in [44] was too few. Using only the five strongest measurements improved the consistency and reduced the relative computing time[4] for all methods. Since the GGM's computational demand depends exponentially on the number of measurements [81] the reduction in computation time for static and filtered GGM could be expected, although in our tests the dependence is not exponential due to collapsing a GGM

[4] The large time value for static PL can be explained by the facts that in two (of 308) points the convergence was extremely slow and that our implementation did not restrict the number of iterations.

after five measurements to a single Gaussian. At the same time the positing accuracy degraded significantly only for WKNN and the CA 1-level approach. This is evidence for dependency of the measurements. In the test building there were some MIMO (Multiple Input Multiple Output) APs that produced dependent measurements.

Figure 8 reveals that the more sophisticated approaches (PL, GGM and WKNN) perform worse or similar than the relatively simple CA methods for scenarios with low AP density. The same holds for the scenario in which the radio map was outdated (compare Fig. 9).

One possible reason for the static and the filtered GGM's poor performance in all four scenarios (compared with their performance in [80] and in [81]) might be that we used a different approach for determining the covariance matrices of the GGM's two Gaussian components, since our tests where carried out in a WLAN rather than in a cellular network. A deeper analysis of the GGM can be found in [81].

VIII. CONCLUSION

In this paper we considered different parametric fingerprinting and positioning methods, analyzing weaknesses and strengths. We tested several of those methods with real WLAN data for different test tracks and scenarios; we furthermore compared their positioning accuracy and consistency with each other and a WKNN (as an example of a nonparametric FP method).

Table 2 summarizes the considered parametric methods; what parameters are stored in the radio map instead of FP data, how these parameters are obtained and how they are used for positioning.

All the proposed methods help to significantly reduce the size of the radio map used in the positioning phase, compared with nonparametric methods. In addition, it is possible to update the radio maps used by the CA, PL model, GGM and GMEM approaches we tested as new FPs become available.

Our tests show that all parametric methods, except the CA 1-level and the filtered GMEM method, provide similar positioning accuracy than the nonparametric WKNN in case of a high CN density and when using all available measurements. However, this comes at the cost of significantly higher computation time for the PL model, GGM and GMEM methods. When using only the five strongest measurements their computation time drops significantly. Furthermore, all parametric methods still show similar positioning performances, while the WKNN's performance degrades considerably. This means that the parametric methods need fewer observable CNs than the nonparametric method to achieve satisfying positioning accuracy. When the CN density is low or the mapping data is outdated then the simpler CA 2-level technique achieves at least similar positioning accuracy than the more sophisticated parametric techniques and the WKNN. Thus, the CA technique gives the best trade-off between accuracy and computational demand. The other parametric methods are, like the WKNN, more vulnerable to harsh environments. However, we believe, and studies presented in this paper have shown, that both PL model and GGM approach can outperform the CA methods when thoroughly trained for their specific application, which we excluded from our tests.

It is important to notice that the achieved positioning accuracy of all methods is sufficient for many real-world applications (e.g. weather forecast, advertising), but insufficient for navigation. To improve the methods' performances map information and additional measurements (e.g. from an IMU, AOA or time delay measurements) could be used. We believe that using floor maps would improve their positioning accuracy since, for example, crossing walls could be prohibited. For example, Nurminen et al. [85] show that using the floor map improves their particle filter's positioning accuracy significantly. However, how to combine them with some of the other methods presented in this paper is still an open question. Furthermore, map information can also be used in the offline phase to generate a more accurate radio map.

ACKNOWLEDGMENTS

Philipp Müller acknowledges the financial support of the TUT Doctoral Programme. Matti Raitoharju has been supported by Tampere Doctoral Programme in Information Science and Engineering, Nokia Inc., Nokia Foundation and Jenny and Antti Wihuri Foundation. This research was partly funded by HERE, a Nokia Business.

The authors thank Henri Nurminen for his advice and for provision of Matlab code for his approach [50].

The authors declare that there is no conflict of interests regarding the publication of this article.

REFERENCES

1. Gezici, S., Tian, Z., Biannakis, G.B., Kobayashi, H., Molisch, A.F., Poor, H.V., and Sahinoglu, Z., Localization via ultra-wideband radios: a look at positioning aspects for future sensor networks, *IEEE Signal Processing Magazine*, July 2005, vol. 22, no. 4, pp. 70–84.
2. Gustafsson, F. and Gunnarsson, F., Mobile positioning using wireless networks: possibilities and fundamental limitations based on available wireless network measurements, *IEEE Signal Processing Magazine*, July 2005, vol. 22, no. 4, pp. 41–53.
3. Trevisani, E. and Vitaletti, A., Cell-id location technique, limits and benefits: an experimental study, in *Proceedings of Sixth IEEE Workshop on Mobile Computing Systems and Applications (WMCSA 2004)*, Windermere, Cumbria, United Kingdom, December 2004.
4. Wirola, L., Studies on location technology standards evolution in wireless networks, Ph.D. dissertation, Tampere University of Technology, March 2010,

[Online], Available: http://URN.fi/URN:NBN:fi:tty-201002121065
5. Dardari, D., Conti, A., Ferner, U., Giorgetti, A., and Win, M.Z., Ranging with ultrawide bandwidth signals in multipath environments, *Proceedings of the IEEE*, February 2009, vol. 97, no. 2, pp. 404–426.
6. Mautz, R., Indoor positioning technologies, Habilitation Thesis, http://e-collection.library.ethz.ch/eserv/eth:5659/eth-5659-01.pdf, February 2012.
7. Liu, H., Darabi, H., Banerjee, P., and Liu, J., Survey of wireless indoor positioning techniques and systems, *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, November 2007, vol. 37, no. 6, pp. 1067–1080.
8. Silventoinen, M.I. and Rantalainen, T., Mobile station emergency locating in GSM, in *Proceedings of IEEE International Conference on Personal Wireless Communications*, New Delhi, India, February 1996, pp. 232–238.
9. Constandache, I., Gaonkar, S., Sayler, M., Choudhury, R.R., and Cox, L., EnLoc: Energy-efficient localization for mobile phones, in *2009 Proceedings IEEE INFOCOM*, Rio de Janeiro, Brazil, April 2009, pp. 2716–2720.
10. Spirito, M., Pöykkö, S., and Knuuttila, O., Experimental performance of methods to estimate the location of legacy handsets in GSM, in *IEEE VTS 54th Vehicular Technology Conference (VTC), 2001*, Atlantic City, NJ, USA, October 2001, pp. 2716–2720.
11. el Peral-Rosado, J.A., López-Salcedo, J.A., Seco-Granados, G., Zanier, F., and Crisci, M., Analysis of positioning capabilities of 3GPP LTE, in *Proceedings of 25th International Technical Meeting of the Satellite Division of ION*, Nashville, TN, USA, September 2012.
12. Wirola, L., Laine, T.A., and Syrjärinne, J., Mass-market requirements for indoor positioning and indoor navigation, in *2010 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Zurich, Switzerland, September 2010.
13. Park, J., Curtis, D., Teller, S., and Ledlie, J., Implications of device diversity for organic localization, in *2011 Proceedings IEEE INFOCOM*, April 2011, pp. 3182–3190.
14. Molisch, A.F., *Wireless Communications*, Wiley - IEEE, January 2011, second edition.
15. Honkavirta, V., Perälä, T., Ali-Löytty, S., and Piché, R., A comparative survey of WLAN location fingerprinting methods, in *Proceedings of the 6th Workshop on Positioning, Navigation and Communication 2009 (WPNC'09)*, Hannover, Germany, March 2009, pp. 243–251.
16. Figuera, C., Mora-Jiménez, I., Guerrero-Curieses, A., Rojo-Álvarez, J.L., Everss, E., Wilby, M., and Ramos-López, J., Nonparametric model comparison and uncertainty evaluation for signal strength indoor location, *IEEE Transactions on Mobile Computing*, September 2009, vol. 8, no. 9, pp. 1250–1264.
17. Müller, P., Raitoharju, M., and Piché, R., A field test of parametric WLAN-fingerprint-positioning methods, in *Proceedings of the 17th International Conference on Information Fusion (Fusion 2014)*, Salamanca, Spain, July 2014, pp. 1–8.
18. Han, S., Gong, Z., Meng, W., and Li, C., An indoor radio propagation model considering angles for WLAN infrastructures, *Wireless Communications and Mobile Computing*, 2015 (in press).
19. Shrestha, S., Talvitie, J., and Lohan, E.S., Deconvolution-based indoor localization with WLAN signals and unknown access point locations, in *2013 International Conference on Localization and GNSS (ICL-GNSS)*, Torino, Italy, June 2013, pp. 1–6.
20. Koski, L., Piché, R., Kaseva, V., Ali-Löytty, S., and Hännikäinen, M., Positioning with coverage area estimates generated from location fingerprints, in *Proceedings of the 7th Workshop on Positioning, Navigation and Communication 2010 (WPNC'10)*, Dresden, Germany, March 2010, pp. 99–106.
21. Machaj, J., Piché, R., and Brida, P., Rank based fingerprinting algorithm for indoor positioning, in *2011 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Guimaraes, Portugal, September 2011, pp. 6–11.
22. Wirola, L., Wirola, L., and Piché, R., Bandwidth and storage reduction of radio maps for offline WLAN positioning, in *2013 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Montbeliard-Belfort, France, October 2013.
23. Eisa, S., Peixoto, J., Meneses, F., and Moreira, A., Removing useless APs and fingerprints from WiFi indoor positioning radio maps, in *2013 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Montbeliard-Belfort, France, 2013, pp. 739–745.
24. Raitoharju, M., Fadjukoff, T., Ali-Löytty, S., and Piché, R., Using unlocated fingerprints in generation of WLAN maps for indoor positioning, in *Proceedings of PLANS 2012 IEEE/ION Position Location and Navigation Symposium*, Myrtle Beach, SC, USA, April 2012, pp. 576–583.
25. Wirola, L., Halivaara, I., and Syrjärinne, J., Requirements for the next generation standardized location technology protocol for location-based services, *Journal of Global Positioning Systems*, 2008, vol. 7, no. 2, pp. 91–103.
26. Kaemarungsi, K and Krishnamurthy, P., Properties of indoor received signal strength for WLAN location fingerprinting, in *The First Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services (MOBIQUITIOUS 2004)*, Boston, MA, USA, August 2004, pp. 14–23.
27. Vaupel, T., Seitz, J., Kiefer, F., Haimerl, S., and Thielecke, J., Wi-Fi positioning: System considerations and device calibration, in *2010 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Zurich, Switzerland, September 2010.
28. Laoudias, C., Piché, R., and Panayiotou, C.G., Device signal strength self-calibration using histograms, in *2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Sydney, NSW, Australia, November 2012.
29. Koski, L., Perälä, T., and Piché, R., Indoor positioning using WLAN coverage area estimates, in *2010 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Zurich, Switzerland, September 2010.
30. Haeberlen, A., Flannery, E., Ladd, A.M., Rudys, A., Wallach, D.S., and Kavraki, L.E., Practical robust localization over large-scale 802.11 wireless networks, in *MobileCom'04*, Philadelphia, PA, USA, September-October 2004.
31. Kjaergaard, M.B., Indoor location fingerprinting with heterogeneous clients, *Pervasive and Mobile Computing*, 2011, vol. 7, no. 1, pp. 31–43.

32. Cheng, Y.-C., Chawathe, C., LaMarca, A., and Krumm, A., Accuracy characterization for metropolitan-scale Wi-Fi localization, in *Proceedings of the 3rd International Conference on Mobile Systems, Applications and Services (MobiSys 2005)*, Seattle, WA, USA, June 2005.
33. Hossain, A.M. and Soh, W.-S., Cramér-Rao bound analysis of localization using signal strength difference as location fingerprint, in *2010 Proceedings IEEE INFOCOM*, San Diego, CA, USA, March 2010.
34. Jazwinski, A.H., *Stochastic Processes and Filtering Theory*, ser. Mathematics in Science and Engineering, Academic Press, 1970, vol. 64.
35. Särkkä, S., *Bayesian Filtering and Smoothing*, Cambridge University Press, 2013.
36. Bar-Shalom, Y., Li, R.X., and Kirubarajan, T., *Estimation with Applications to Tracking and Navigation, Theory Algorithms and Software*, John Wiley & Sons, 2001.
37. Piché, R., Robust estimation of a reception region from location fingerprints, in *2011 International Conference on Localization and GNSS (ICL-GNSS)*, Tampere, Finland, June 2011, pp. 31–35.
38. Koski, L., Positioning with Bayesian coverage area estimates and location fingerprints, Master's thesis, Tampere University of Technology, http://math.tut.fi/posgroup/koski_mscth.pdf, March 2010.
39. Raitoharju, M., Ali-Löytty, S., Piché, R., and Dashti, M., Positioning with multilevel coverage area models, in *2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Sydney, NSW, Australia, November 2012.
40. Geman, S and Geman, D., Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1984, vol. PAMI-6, no. 6, pp. 721–741.
41. Dempster, A.P., Laird, N.M., and Rubin, D.B., Maximum likelihood from incomplete data via the EM algorithm, *Journal of the Royal Statistical Society, Series B (Methodological)*, vol. 39, no. 1, pp. 1–38, 1977.
42. Honkavirta, V., Location fingerprinting methods in wireless local area networks, Master's thesis, Tampere University of Technology, http://urn.fi/URN:NBN:fi:tty-200911137109, November 2008.
43. Laoudias, C., Panayiotou, C.G., and Kemppi, P., On the RBF-based positioning using WLAN signal strength fingerprints, in *Proceedings of the 7th Workshop on Positioning, Navigation and Communication 2010 (WPNC'10)*, Dresden, Germany, March 2010, pp. 93–98.
44. Kaji, K. and Kawaguchi, N., Design and implementation of WiFi indoor localization based on Gaussian mixture model and particle filter, in *2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Sydney, NSW, Australia, November 2012.
45. Hashemi, H., The indoor radio propagation channel, *Proceedings of the IEEE*, July 1993, vol. 81, no. 7, pp. 943–968.
46. Patwari, N., Alfred O. Hero, I., Perkins, M., Correal, N.S., and O'Dea, R.J., Relative location estimation in wireless sensor networks, *IEEE Transactions on Signal Processing*, August 2003, vol. 51, no. 8, pp. 2137–2148.
47. Li, X., RSS-based location estimation with unknown pathloss model, *IEEE Transactions on Wireless Communications*, December 2006, vol. 5, no. 12, pp. 3626–3633.
48. Roos, T., Myllymäki, P., and Tirri, H., A statistical modeling approach to location estimation, *IEEE Transactions on Mobile Computing*, Jan-Mar 2002, vol. 1, no. 1, pp. 59–69.
49. Bo-Chieh Liu, S.M., Lin, K.H., and Wu, J.-C., Analysis of hyperbolic and circular positioning algorithms using stationary signal-strength-difference measurements in wireless communications, *IEEE Transactions on Vehicular Technology*, March 2006, vol. 55, no. 2, pp. 499–509.
50. Nurminen, H., Talvitie, J., Ali-Löytty, S., Müller, P., Lohan, E.S., Piché, R., and Renfors, M., Statistical path loss parameter estimation and positioning using RSS measurements in indoor wireless networks, in *2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Sydney, NSW, Australia, November 2012.
51. Salo, J., Vuokko, L., El-Sallabi, H.M., and Vainikainen, P., An additive model as a physical basis for shadow fading, *IEEE Transactions on Vehicular Technology*, January 2007, vol. 56, no. 1, pp. 13–26.
52. Erceg, V., Greenstein, L.J., Tjandra, S.Y., Parkoff, S.R., Gupta, A., Kulic, B, Julius, A.A., and Bianchi, R., An empirically based path loss model for wireless channels in suburban environments, *IEEE Journal on Selected Areas in Communications*, 1999, vol. 17, no. 7, pp. 1205–1211.
53. Rappaport, T.S., Seidel, S.Y., and Takamizawa, K., Statistical channel impulse response models for factory and open plan building radio communication system design, *IEEE Transactions on Communications*, 1991, vol. 39, no. 5, pp. 794–807.
54. Ghassenizadeh, S.S., Jana, R., Rice, C.W., Turin, W., and Tarokh, V., A statistical path loss model for in-home UWB channels, in *2002 IEEE Conference on Ultra Wideband Systems and Technologies*, Baltimore, MD, USA, May 2002, pp. 59–64.
55. Dil, B. and Havinga, P., RSS-based localization with different antenna orientations, in *2010 Australasian Telecommunication Networks and Applications Conference (ATNAC)*, Auckland, New Zealand, 2010, pp. 13–18.
56. Shrestha, S., Laitinen, E., Talvitie, J., and Lohan, E.S., RSSI channel effects in cellular and WLAN positioning, in *Proceedings of the 9th Workshop on Positioning, Navigation and Communication 2012 (WPNC'12)*, Dresden, Germany, March 2012, pp. 187–192.
57. Rappaport, T.S., *Wireless Communications - Principles and Practice*, Prentice-Hall, 2010, second edition.
58. Seidel, S.Y. and Rappaport, T.S., 914 MHz path loss prediction models for indoor wireless communications in multifloored buildings, *IEEE Transactions on Antennas and Propagation*, February 1992, vol. 40, no. 2, pp. 207–217.
59. Dil, B and Havinga, P., On the calibration and performance of RSS-based localization methods, in *Internet of Things 2010 Conference*, Tokyo, Japan, November/December 2010.
60. Hata, M., Empirical formula for propagation loss in land mobile radio services, *IEEE Transactions on Vehicular Technology*, August 1980, vol. 29, no. 3, pp. 317–325.
61. Shujaee, K., Ebaid, A., George, R., and Sazegarnejad, M.A., Channel modeling and range

extension for UWB communications using directional antenna in LOS and NLOS path loss models, in *World Automation Congress (WAC 2008)*, Hawaii, HI, USA, September-October 2008.

62. Sorenson, H.W. and Alspach, D.L., Recursive Bayesian estimation using Gaussian sums, *Automatica*, July 1971, vol. 7, no. 4, pp. 465–479.
63. Ali-Löytty, S., Gaussian mixture filters in hybrid positioning, Ph.D. dissertation, Tampere University of Technology, August 2009, Online, Available: http://URN.fi/URN:NBN:fi:tty-200905191055
64. Lo, J. T.-H., Finite-dimensional sensor orbits and optimal nonlinear filtering, *IEEE Transactions on Information Theory*, September 1972, vol. 18, no. 5, pp. 583–588.
65. Alspach, D.L. and Sorenson, H.W., Nonlinear Bayesian estimation using Gaussian sum approximations, *IEEE Transactions on Automatic Control*, August 1972, vol. 17, no. 4, pp. 439–448.
66. Koshizen, T., The sensor selection task of the Gaussians mixture Bayes' with regularized EM (GMB-REM) technique in robot position estimation, in *Proceedings of the 1999 IEEE International Conference on Robotics & Automation*, Detroit, MI Detroit, MI Detroit, MI, USA, May 1999, pp. 2620–2625.
67. Koshizen, T., The evolved Gaussian mixture Bayes' technique using sensor selection task integrated with sensor fusion scheme in mobile robot position estimation, in *Proceedings of 1999 IEEE International Symposium on Computational Intelligence in Robotics and Automation (CIRA'99)*, Monterey, CA, USA, November 1999, pp. 202–207.
68. Koshizen, T., Bartlett, P., and Zelinsky, A., Sensor fusion of odometry and sonar sensors by the Gaussian mixture Bayes' technique in mobile robot position estimation, in *IEEE SMC '99 Conference Proceedings*, vol. 4, Tokyo, Japan, 1999, pp. 742–747.
69. Ali-Löytty, S. and Sirola, N., Gaussian mixture filter in hybrid navigation, in *Proceedings of The European Navigation Conference GNSS 2007*, Geneva, Switzerland, May 2007, pp. 831–837.
70. Arulampalam, M.S., Maskell, S., Gordon, N., and Clapp, T., A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking, *IEEE Transactions on Signal Processing*, 2002, vol. 50, no. 2, pp. 174–188.
71. Ali-Löytty, S., Efficient Gaussian mixture filter for hybrid positioning, in *2008 IEEE/ION Position Location and Navigation Symposium*, Monterey, California, May 2008, pp. 60–66.
72. Ali-Löytty, S. and Sirola, N., Gaussian mixture filter and hybrid positioning, in *Proceedings of ION GNSS 2007*, Fort Worth, Texas, September 2007, pp. 562–570.
73. Raitoharju, M. and Ali-Löytty, S., An adaptive derivative free method for Bayesian posterior approximation, *IEEE Signal Processing Letters*, February 2012, vol. 19, no. 2, pp. 87–90.
74. Anderson, B.D.O. and Moore, J.B., *Optimal Filtering*, ser. Prentice-Hall information and system sciences, Prentice-Hall, 1979.
75. Ali-Löytty, S., On the convergence of the Gaussian mixture filter, Tampere University of Technology, Institute of Mathematics, http://urn.fi/URN:NBN:fi:tty-2011041510704, Tech. Rep. Research report 89, 2008.
76. Salmond, D.J., Mixture reduction algorithms for target tracking, in *IEE Colloquium on State Estimation in Aerospace and Tracking Applications*, London, England, 1989, pp. 7/1–7/4.
77. Schieferdecker, D. and Huber, M.F., Gaussian mixture reduction via clustering, in *Proceedings of the 12th International Conference on Information Fusion (FUSION '09)*, Seattle, WA, USA, 2009, pp. 1536–1543.
78. Crouse, D.F., Willett, P., Pattipati, K., and Svensson, L., A look at Gaussian mixture reduction algorithms, in *Proceedings of the 14th International Conference on Information Fusion (FUSION '11)*, Chicago, IL, USA, 2011.
79. Kristan, M., Skočaj, D., and Leonardis, A., Incremental learning with Gaussian mixture models, in *Computer Vision Winter Workshop 2008*, Moravske Toplice, Slovenia, February 2008, pp. 25–32.
80. Müller, P., Ali-Löytty, S., Dashti, M., Nurminen, H., and Piché, R., Gaussian mixture filter allowing negative weights and its application to positioning using signal strength measurements, in *Proceedings of the 9th Workshop on Positioning, Navigation and Communication 2012 (WPNC'12)*, Dresden, Germany, March 2012, pp. 71–76.
81. Müller, P., Wymeersch, H., and Piché, R., UWB positioning with generalized Gaussian mixture filters, *IEEE Transactions on Mobile Computing*, October 2014, vol. 13, no. 10, pp. 2406–2414.
82. Bishop, C.M. and Svensén, M., Robust Bayesian mixture modeling, in *Proceedings of 12th European Symposium on Artificial Neural Networks (ESANN 2004)*, Bruges, Belgium, April 2004, pp. 69–74.
83. McLachlan, G.J. and Peel, D., Robust cluster analysis via mixtures of multivariate t-distributions, in *Advances in Pattern Recognition*, ser. Lecture Notes in Computer Science, Amin, A., Dori, D., Pudil, P., and Freeman, H., Eds, Springer Berlin Heidelberg, 1998, vol. 1451, pp. 658–666.
84. Ali-Löytty, S., Sirola, N., and Piché, R., Consistency of three Kalman filter extensions in hybrid navigation, in *Proceedings of The European Navigation Conference GNSS 2005*, Munich, Germany, Jul. 2005.
85. Nurminen, H., Ristimäki, A., Ali-Löytty, S., and Piché, R., Particle filter and smoother for indoor localization, in *2013 International Conference on Indoor Positioning and Indoor Navigation (IPIN2013)*, Montbeliard-Belfort, France, October 2013, pp. 137–146.