Kaltiokallio, Ossi; Hostettler, Roland; Patwari, Neal

A Novel Bayesian Filter for RSS-Based Device-Free Localization and Tracking

Published in:
IEEE Transactions on Mobile Computing

DOI:
10.1109/TMC.2019.2953474

Published: 01/03/2021

Document Version
Publisher's PDF, also known as Version of record

A Novel Bayesian Filter for RSS-Based Device-Free Localization and Tracking

Ossi Kaltiokallio, Roland Hostettler, *Member, IEEE*, and Neal Patwari

Abstract—Received signal strength based device-free localization applications utilize a model that relates the measurements to position of the wireless sensors and person, and the underlying inverse problem is solved either using an imaging method or a nonlinear Bayesian filter. In this paper, it is shown that the Bayesian filters nearly reach the posterior Cramér-Rao bound and they are superior with respect to imaging approaches in terms of localization accuracy because the measurements are directly related to position of the person. However, Bayesian filters are known to suffer from divergence issues and in this paper, the problem is addressed by introducing a novel Bayesian filter. The developed filter augments the measurement model of a Bayesian filter with position estimates from an imaging approach. This bounds the filter's measurement residuals by the position errors of the imaging approach and as an outcome, the developed filter has robustness of an imaging method and tracking accuracy of a Bayesian filter. The filter is demonstrated to achieve a localization error of 0.11 m in a 75 m^2 open indoor deployment and an error of 0.29 m in a 82 m^2 apartment experiment, decreasing the localization error by 30-48 percent with respect to a state-of-the-art imaging method.

Index Terms—Received signal strength, wireless sensor networks, Bayesian filtering, posterior Cramér-Rao bound, positioning and tracking

1 Introduction

Ubiquitous radio frequency (RF) sensing technologies have experienced a surge of interest over the past years and are considered as a potential candidate to be used in smart homes. Smart homes control heating, ventilation and air conditioning systems to improve environmental sustainability and the comfort of their residents [1]. Our vision is that future smart homes would not only monitor the homes we live in, but also its inhabitants. Such system capabilities are enabled by RF signals and recently, various radio signal measurements have been demonstrated for vital sign monitoring [2], activity and gesture recognition [3], and localization and tracking [4]. With such information, the smart home could be controlled using our gestures, the vital sign information could be used to enhance our health-awareness, and heating and lighting could be automatically adjusted based on our location.

Non-invasive RF sensing technologies are built upon the fact that humans alter the propagation characteristics of radio signals and at the receiver, these changes can be quantified using the radio's channel measurements. Research has demonstrated the use of various radio signal measurements for inference, including time delay [2], phase [3], and signal strength [4]; and these have been used for various purposes as mentioned above. Most notably, the technology is non-invasive and does not require the person to carry any electronic device. Moreover, the technology can be realized with received signal strength (RSS) measurements that are ubiquitously available in nearly all receivers. In this paper, we consider narrowband wireless devices that measure the RSS and we utilize the channel measurements for locating and tracking people in indoor environments. It is to be noted that the technology is not limited to localizing people and the proposed method could also be used to locate large animals [5] and vehicles [6]. The readers are referred to [7], [8] for a comprehensive overview of RF-based passive localization technologies.

In RSS-based device-free localization and tracking (DFLT), the algorithmic approaches can be divided into two categories. In the first, the person is located using an imaging approach [9], [10] and a Kalman filter (KF) is used for tracking [11], [12]. In the second, a propagation model together with a nonlinear Bayesian filter such as a particle filter (PF) [13], [14], [15] or an extended Kalman filter (EKF) [16] is used to track the kinematic state of the target. The considered problem can be solved more accurately using a nonlinear Bayesian filter, however, these filters are known to suffer from divergence issues if the modeling errors are significant [17, p.128]. To address this problem, we introduce a novel Bayesian filter in which the measurement update recursion is augmented with position estimates from an imaging approach. The benefit of the proposed approach is that the filter's measurement residuals are bounded by the position errors of the imaging approach, and as a result, the filter has the robustness of an imaging method and almost the tracking accuracy of a nonlinear Bayesian filter. Moreover, the implemented filter is computationally efficient. We refer to the developed Bayesian filter as *Fusion Filter* (FF) and it practically merges an EKF approach with an imaging approach.

- *O. Kaltiokallio is with the Department of Communications and Networking, Aalto University, 02150 Espoo, Finland. E-mail: ossi.kaltiokallio@aalto.fi.*
- *R. Hostettler is with the Department of Engineering Sciences, Uppsala University, 752 36 Uppsala, Sweden. E-mail: roland.hostettler@angstrom.uu.se.*
- *N. Patwari is with the McKelvey School of Engineering, Washington University in St. Louis, St. Louis, MO 63130 USA, and also with Xandem Technology LLC, Salt Lake City, UT 84111 USA. E-mail: npatwari@wustl.edu.*

Manuscript received 24 Jan. 2019; revised 16 Oct. 2019; accepted 3 Nov. 2019. Date of publication 11 Dec. 2019; date of current version 3 Feb. 2021. (Corresponding author: Ossi Kaltiokallio.)
Digital Object Identifier no. 10.1109/TMC.2019.2953474

TABLE 1
Major Notations and Common Operators

Symbol	Description
z_l, $\mathbf{z}_k$ and $\mathbf{Z}$	A scalar RSS measurement of link l, RSS measurement vector at time k and RSS matrix
$\mathbf{x}_k$, $\hat{\mathbf{x}}_k$, $\mathbf{P}_k$	State of the person at time k, estimate of the state and covariance of the estimate
h_l, $\mathbf{H}_\mathbf{x}$ and $\mathbf{H}$	Measurement model of link l, Jacobian of $h_l(\mathbf{x}_k)$ and linear measurement model
$\Delta_l(k)$ and $\boldsymbol{\theta}_l = [\phi_l \ \lambda_l \ \sigma_l^2]$	Excess path length of link l at time k and model parameters of the link
$(\cdot)^T$, $(\cdot)^{-1}$ and $\|\cdot\|$	matrix transpose, inverse and euclidean norm

The work is motivated by deriving the posterior Cramér-Rao bound (PCRB) for the RSS-based DFLT problem and evaluating two estimators with respect to the bound. The used RSS-based DFLT estimators are: an EKF-based method [16] and a modified radio tomographic imaging (RTI) method [12]. The analysis clearly shows that RTI is lower bounded by the pixel size of the discretized image and this bound is significantly higher than the PCRB. On the other hand, the EKF nearly achieves the bound which encourages its use in RSS-based DFLT. However, the divergence issues of the EKF must be solved and we propose to use the FF for this purpose. The filter is experimentally and numerically evaluated. The results imply that the presented filter nearly achieves the performance of the EKF in ideal scenarios, it outperforms the EKF and PF in more challenging environments and it has the robustness of an imaging method. The presented filter is demonstrated to achieve a localization error as low as 11 centimeters in a $75\ \mathrm{m}^2$ open indoor deployment and an error of 29 centimeters in a $82\ \mathrm{m}^2$ apartment experiment, decreasing the localization error by 30-48 percent with respect to a state-of-the-art imaging method.

In RF sensing, sensor fusion is the inherent way of combining the information from multiple sources (distributed sensor nodes [9], [18], antennas [2], [3] and/or frequency channel [12], [19]). Several works have also used different types of sensors including ultra-wideband radios [20], cameras [21], [22], [23] and acoustic sensors [24] to enhance the performance of a system that solely uses radio channel measurements. The works perform sensor fusion by combining the sensory data from two kinds of sensors to reduce the uncertainty and improve the accuracy. Our solution differs from traditional sensor fusion methods since we do not combine different types of sensory data but instead, we merge the information from two categories of DFLT approaches into one filtering algorithm.

This paper makes the following contributions:

- A closed form solution for the PCRB is derived and two well known estimators are evaluated with respect to the bound.
- A sequential imaging method is proposed allowing recursive image updates whenever new RSS measurements are received.
- We propose a method on how the uncertainties related to RTI position estimates can be taken into account.
- A novel filtering framework is proposed that augments the measurement update recursion of a nonlinear Bayesian filter with position estimates from an imaging solution.

The remainder of the paper is organized as follows. In the next section, related work is discussed. In Section 3, the problem of tracking the kinematic state of the person is formulated and two estimators are presented. The PCRB is derived in Section 4 and the bounds of RSS-based DFLT are analyzed. Motivated by the bound analysis, the FF is developed and the filter is presented in Section 5. The experiments and results are presented in Sections 6 and 7 in respective order and thereafter, the conclusions are drawn. In Table 1, major notations of the paper are summarized.

2 Related Work

In RSS-based DFLT, there are two widely used approaches for locating people: fingerprinting [25], [26], [27], and model-based approaches [9], [12], [14]. Fingerprinting methods use a database of training data labelled with a person's known locations. During runtime, the current set of RSS measurements are compared to those in the database to estimate the current location. Model-based approaches use an a priori model for the changes in RSS with respect to the locations of the sensors and person, and localization is performed for example using an imaging approach [9], [10]. Fingerprinting methods are able to achieve high accuracy also in demanding environments, but the training process is laborious and the performance degrades exponentially as the environment is altered [27]. Model-based approaches can be deployed quickly [28], but the mismatch between the RSS model and measurements can significantly affect the system performance [12]. This paper focuses on model-based DFLT and a novel tracking filter is presented, and it is shown that the filter is robust to modeling errors.

In model-based approaches, the person is typically located and tracked either using an imaging approach [9], [10], [12] or Sequential Monte Carlo (SMC) methods [13], [14], [15]. The imaging methods compute a propagation field image of the monitored area [9], [10], the person is localized from the estimated image, and then a KF is used for estimating the kinematic state of the target [11], [12]. In literature, the changes in the propagation field have been quantified using various RSS link metrics including shadowing [9], RSS variance [11] and kernel distance between two RSS histograms [28]. The benefit of the imaging approaches is that they are computationally efficient, they are robust and an improper prior does not cause the tracking filter to diverge. As a drawback, information can be lost in the two-step process to first estimate the image and then the location. In addition, discretization of the image inevitably degrades the localization accuracy.

The SMC methods typically utilize a PF to solve the problem, and in the tracking algorithms, the RSS measurements are directly related to the person's location using either an empirical model [13], [15] or a theoretical propagation model

[14], [29]. Different variants of the PF have been proposed in literature including the sequential importance resampling (SIR) filter [13], [14], [15], auxiliary particle filter [30] and Markov chain Monte Carlo filter [31]. However, the filters have been reported to yield comparative accuracy at least in the single target tracking case [31]. Furthermore, our recent work demonstrates that an EKF yields similar performance as a PF as long as the time evolution of the measurements is taken correctly into account in the filter recursion [16]. The benefit of the Bayesian filters is that they are more accurate than the imaging approaches if the used RSS-model describes the data well [13], [16], [29]. As a disadvantage, the PF is computationally very demanding, and the filters can diverge if the prior distribution is inaccurate [16]. In this paper, we introduce a Bayesian filter in which the measurement update recursion is augmented with position estimates from an imaging approach in order to improve robustness of the filter. The developed filter is computationally less demanding than the PF, it has the robustness of an imaging method and the tracking accuracy of a Bayesian filter.

The PCRB states that the mean squared error (MSE) of an unbiased estimator is always larger than the bound [32], [33]. The PCRB can be used to evaluate the developed estimator and in addition, as a pre-deployment predictor of system performance, providing an analytical method for system design and pre-deployment performance evaluation. Despite the importance of the PCRB, there exists only a few works in RSS-based DFLT that have used it for evaluation purposes [10], [14], [29], [34], [35], [36]. In [10], the CRB is derived for RTI and in the work, it is studied how the node locations affect the accuracy of image estimation. It is shown that the best node geometry is where the nodes are deployed uniformly around the monitored area. The CRB for five different RTI models is derived in [34] providing an analytical tool on how the system parameters affect the CRB, enabling analysis of the tradeoffs between the parameters in system design. Neither analysis provides a bound on position estimates—they can only bound the covariance of the values of the pixels in the image. In this paper, we derive the PCRB on localization error similar to the works in [14], [29], [35], [36]. Contrary to [14], we provide a closed-form solution to the PCRB. In addition, we incorporate the apriori knowledge of the target dynamics and position into the PCRB as opposed to the works in [29], [35], [36]. In fact, the PF used in [29] outperforms the derived bound and the authors point out that the conventional CRB does not incorporate the apriori information of the motion. Furthermore, the aforementioned works use a diffraction-based RSS model which has been validated only in ideal line-of-sight (LOS) scenarios limiting the usefulness of the derived bounds. In this paper, we use an exponential model [13] that is widely used in literature. In addition, the model has been used in challenging through-wall scenarios [12] and therefore, the derived PCRB can be applied to a wide range of environments.

Radio source localization systems have been investigated and deployed in a variety of forms over the past several decades [37]. Source localization from signal-strength is most relevant to this paper, and bounds and algorithms have been presented [38]. Similarly, localization bounds using temporal characteristics has been widely investigated via geometric dilution of precision [39] or via CRB [40]. Bounds for systems that combine multiple signal characteristics (time, power, and angle) are given in [41]. In general, the RSS source localization variance bound is inversely proportional to the average squared distance between neighboring sensors. It also decreases with increasing path loss exponent and increases with fading variance [42]. Respectively, the RSS-based DFLT variance bound decreases with increasing number of wireless links that cover the location. It also decreases with increasing measurement gain and increases with measurement noise variance.

3 RSS-Based DFLT

This section presents the background information needed to derive the PCRB for RSS-based DFLT in Section 4 and developing the novel Fusion filter in Section 5. This section begins by defining the problem of localizing and tracking a person using RSS measurements of wireless links. Thereafter, two solutions from the literature are summarized [12], [16]. The first solution is based on an EKF that directly relates the RSS measurements to the person's kinematic state [16]. The latter is a two-step method, where a discretized propagation field image is first computed and then, the person's position is estimated from the image [10], [12].

To simplify the notation, we assume in this section that the wireless network consists of S nodes forming $L = S \cdot (S-1)$ unique links and that the link measurements are taken at the same time instance k. We want to emphasize that full connectivity is not a requirement of RSS-based DFLT and that the link measurements can be sampled at different time instances. In Section 5, we present the used communication protocol, how the RSS measurements are sampled and how the time evolution between transmissions is taken into account.

3.1 Problem Formulation

This work aims to localize and track a person using RSS measurements of the wireless nodes. The considered problem can be formulated using a state space model of the form

$$\mathbf{x}_k = \mathbf{f}(\mathbf{x}_{k-1}) + \mathbf{q}_{k-1}, \tag{1a}$$

$$\mathbf{z}_k = \mathbf{h}(\mathbf{x}_k) + \mathbf{r}_k, \tag{1b}$$

where k denotes the time and

$\mathbf{x}_k \in \mathbb{R}^{4\times 1}$ is the person's kinematic state,
$\mathbf{z}_k \in \mathbb{R}^{L\times 1}$ RSS measurement vector,
$\mathbf{q}_{k-1} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q})$ Gaussian process noise,
$\mathbf{r}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$ Gaussian measurement noise,
$\mathbf{f}(\cdot)$ dynamic model of the person,
$\mathbf{h}(\cdot)$ RSS measurement model.

The measurement noise covariance is assumed diagonal and it is defined as $\mathbf{R} = \mathrm{diag}\left(\sigma_1^2, \sigma_2^2, \ldots, \sigma_L^2\right)$. In literature, the DC component is typically removed from the RSS since it does not contain information about location of the person [10], [13], [14], [15]. Thus, what we refer as the RSS of link l is actually the mean removed RSS, that is, $z_l(k) = \tilde{z}_l(k) - \mu_l$,

where $\tilde{z}_l$ is the RSS provided by the radio module and μ_l is the mean RSS computed during an initialization procedure.

The state of the person is defined as

$$\mathbf{x}_k = \begin{bmatrix} p_x(k) & v_x(k) & p_y(k) & v_y(k) \end{bmatrix}^T, \tag{2}$$

where $p_x(k)$ and $p_y(k)$ are the x- and y-coordinates, and the velocity components are denoted as $v_x(k)$ and $v_y(k)$. A common choice for the dynamic model in DFLT is the second-order kinematic model [11], [14], [15] given by [43, Ch 6]

$$\mathbf{F} = \begin{bmatrix} 1 & \tau & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \tau \\ 0 & 0 & 0 & 1 \end{bmatrix}, \mathbf{Q} = q \begin{bmatrix} \frac{1}{3}\tau^3 & \frac{1}{2}\tau^2 & 0 & 0 \\ \frac{1}{2}\tau^2 & \tau & 0 & 0 \\ 0 & 0 & \frac{1}{3}\tau^3 & \frac{1}{2}\tau^2 \\ 0 & 0 & \frac{1}{2}\tau^2 & \tau \end{bmatrix}, \tag{3}$$

where q is the power spectral density of the process noise and τ is the sampling period. In this case, the state evolution in Eq. (1a) can be expressed using $\mathbf{x}_k = \mathbf{F}\mathbf{x}_{k-1} + \mathbf{q}_{k-1}$ because the dynamic model is linear.

3.2 Extended Kalman Filter

The EKF algorithm directly relates the RSS measurements to the person's kinematic state and we model the RSS of link l using the exponential model [13], [15], [31]

$$h_l(\mathbf{x}_k) \triangleq \phi_l \, e^{-\Delta_l(k)/\lambda_l}, \tag{4}$$

where ϕ_l defines the measurement gain when the person is on the LOS and λ_l controls the decay rate. In this paper, we use the exponential model because it has been widely used in literature and its applicability has been demonstrated in challenging indoor environments and through-wall scenarios [12]. Nevertheless, the proposed filter can readily be used used with other RSS models as well. In (4), the excess path length Δ_l relates the person's location $\mathbf{p}_k = \begin{bmatrix} p_x(k) & p_y(k) \end{bmatrix}^T$ to link l with transmitter (TX) i and receiver (RX) j by

$$\Delta_l(k) \triangleq \|\mathbf{p}_i - \mathbf{p}_k\| + \|\mathbf{p}_j - \mathbf{p}_k\| - \|\mathbf{p}_i - \mathbf{p}_j\|, \tag{5}$$

where $\mathbf{p}_i$ and $\mathbf{p}_j$ denote the TX and RX positions in respective order. The EKF requires the Jacobian of $h_l(\mathbf{x}_k)$ for which the elements are given by [16]

$$\begin{bmatrix} \frac{\delta h_l}{\delta p_x} & \frac{\delta h_l}{\delta p_y} \end{bmatrix}^T = \frac{h_l(\mathbf{x}_k)}{\lambda_l} \left(\frac{\mathbf{p}_i - \mathbf{p}_k}{\|\mathbf{p}_i - \mathbf{p}_k\|} + \frac{\mathbf{p}_j - \mathbf{p}_k}{\|\mathbf{p}_j - \mathbf{p}_k\|} \right), \tag{6}$$

so that the Jacobian for link l can be expressed as

$$\{\mathbf{H}_\mathbf{x}\}_l = \begin{bmatrix} \frac{\delta h_l}{\delta p_x} & 0 & \frac{\delta h_l}{\delta p_y} & 0 \end{bmatrix}. \tag{7}$$

Given that the dynamic model in (3) is linear, the prediction step of the first order additive noise EKF can be expressed as [17, Ch 4]

$$\begin{aligned} \hat{\mathbf{x}}_k^- &= \mathbf{F}\hat{\mathbf{x}}_{k-1}, \\ \mathbf{P}_k^- &= \mathbf{F}\mathbf{P}_{k-1}\mathbf{F}^T + \mathbf{Q}. \end{aligned} \tag{8}$$

At time k, measurement $\mathbf{z}_k$ becomes available and the mean $\hat{\mathbf{x}}_k^-$ and covariance $\mathbf{P}_k^-$ can be updated using [17, Ch 5]

$$\begin{aligned} \mathbf{S}_k &= \mathbf{H}_\mathbf{x}(\hat{\mathbf{x}}_k^-)\mathbf{P}_k^-\mathbf{H}_\mathbf{x}^T(\hat{\mathbf{x}}_k^-) + \mathbf{R}, \\ \mathbf{K}_k &= \mathbf{P}_k^-\mathbf{H}_\mathbf{x}^T(\hat{\mathbf{x}}_k^-)\mathbf{S}_k^{-1}, \\ \hat{\mathbf{x}}_k &= \hat{\mathbf{x}}_k^- + \mathbf{K}_k\left(\mathbf{z}_k - \mathbf{h}(\hat{\mathbf{x}}_k^-)\right), \\ \mathbf{P}_k &= \mathbf{P}_k^- - \mathbf{K}_k\mathbf{S}_k\mathbf{K}_k^T. \end{aligned} \tag{9}$$

3.3 Radio Tomographic Imaging

In RTI, the RSS for the L links is assumed to be a linear combination of voxel changes plus noise [9]

$$\mathbf{z}_k = \mathbf{A}\mathbf{b}_k + \mathbf{r}_k, \tag{10}$$

where $\mathbf{r}_k \in \mathbb{R}^{L\times 1}$ is the measurement noise defined in (1b), $\mathbf{A} \in \mathbb{R}^{L\times N}$ is a weight matrix that relates the spatial propagation field $\mathbf{b}_k \in \mathbb{R}^{N\times 1}$ to the RSS $\mathbf{z}_k \in \mathbb{R}^{L\times 1}$ and N is the voxel number. The minimum mean square error estimate (MMSE) for the model in (10), with zero-mean Gaussian priors with image noise covariance $\mathbf{\Sigma}_b$ and measurement noise covariance $\mathbf{R}$ is

$$\hat{\mathbf{b}}_k = \mathbf{\Pi}\mathbf{z}_k, \tag{11}$$

where $\mathbf{\Pi} = (\mathbf{A}^T\mathbf{R}^{-1}\mathbf{A} + \alpha\mathbf{\Sigma}_b^{-1})^{-1}\mathbf{A}^T\mathbf{R}^{-1}$ in which α is a regularization parameter. From the estimated image, $\hat{\mathbf{b}}$, the person can be localized by finding voxel n with highest intensity, given by

$$\hat{\mathbf{p}}_k \triangleq \begin{bmatrix} p_x(n) \\ p_y(n) \end{bmatrix} = \arg\max_n \; \hat{\mathbf{b}}_k(n). \tag{12}$$

The covariance matrix $\mathbf{\Sigma}_b$ for pixels m and n is [9]

$$\{\mathbf{\Sigma}_b\}_{m,n} = \exp\left(-\|\mathbf{p}_m - \mathbf{p}_n\|/\delta_d\right), \tag{13}$$

where δ_d is a user defined space constant. For link l and pixel n, the elements of $\mathbf{A}$ are [12]

$$\{\mathbf{A}\}_{l,n} = \frac{\mathrm{sgn}(\phi_l)}{\sqrt{d_l}} e^{-\Delta_{l,n}/\lambda_l}, \tag{14}$$

where ϕ_l and λ_l are defined by the RSS measurement model in (4), $\mathrm{sgn}(\cdot)$ is the sign function, $\Delta_{l,n}$ the excess path length and $d_l = \|\mathbf{p}_i - \mathbf{p}_j\|$ the distance between TX-RX pair $i-j$. In [12], the direction of RSS change is taken into account by weighting the measurements using $\mathrm{sgn}(\phi_l)\cdot z_l$. Analogously, the $\mathrm{sgn}(\phi_l)$ term can be included into the weight matrix as we have done above. Also other models for $\mathbf{A}$ have been proposed and the reader is referred to [34] and [44] for further details.

4 Posterior Cramér-Rao Bound

This section presents a new lower bound for coordinate tracking in RSS-based DFLT. While the PCRB is a well-established bound for tracking problems, in general, we are not aware of its application to the RSS-based DFLT problem. The earlier works have only bounded the covariance of the values of the pixels in the image [10], [34], provided the CRB on localization error [29], [35], [36] or have approximated the bound numerically [14]. In the following, we provide a closed-form solution to the PCRB for the tracking problem.

The CRB for a time-varying system, referred to as the Van Trees version of the CRB [32], or posterior CRB [33], states that the MSE of an unbiased estimator is always larger than $\mathbf{J}^{-1}$

$$\mathbb{E}\{(\hat{\mathbf{x}}(\mathbf{z})-\mathbf{x})(\hat{\mathbf{x}}(\mathbf{z})-\mathbf{x})^T\} \geq \mathbf{J}^{-1}, \tag{15}$$

where $\mathbf{J}$ is the Fisher information matrix (FIM), $\hat{\mathbf{x}}(\mathbf{z})$ denotes an estimator of $\mathbf{x}$ which is a function of measurements $\mathbf{z}$. The FIM is defined as

$$\mathbf{J} = \mathbb{E}\left[-\Delta_{\mathbf{x}}^{\mathbf{x}} \log p(\mathbf{x},\mathbf{z})\right], \tag{16}$$

where $p(\mathbf{x},\mathbf{z})$ is the joint probability density function (PDF) of the pair and $\Delta_{\mathbf{x}}^{\mathbf{y}}$ is defined as the gradient product $\Delta_{\mathbf{x}}^{\mathbf{y}} = \nabla_{\mathbf{x}}\nabla_{\mathbf{y}}^T$, where $\nabla_{\mathbf{x}}^T = \left[\frac{\partial}{\partial x_1},\ldots,\frac{\partial}{\partial x_N}\right]$. The joint PDF for an arbitrary time instant k is defined as

$$p(\mathbf{x}_k,\mathbf{z}_k) = p(\mathbf{x}_0)\prod_{i=1}^{k} p(\mathbf{z}_i|\mathbf{x}_i)\prod_{j=1}^{k} p(\mathbf{x}_j|\mathbf{x}_{j-1}), \tag{17}$$

and for a time-varying system, the FIM can be calculated recursively using [45]

$$\mathbf{J}_k = \mathbf{D}_{k-1}^{22} - \mathbf{D}_{k-1}^{21}\left(\mathbf{J}_{k-1} + \mathbf{D}_{k-1}^{11}\right)^{-1}\mathbf{D}_{k-1}^{12}, \tag{18}$$

where

$$\begin{aligned}
\mathbf{D}_{k-1}^{11} &= \mathbb{E}\{-\Delta_{\mathbf{x}_{k-1}}^{\mathbf{x}_{k-1}} \log p(\mathbf{x}_k|\mathbf{x}_{k-1})\},\\
\mathbf{D}_{k-1}^{12} &= \mathbb{E}\{-\Delta_{\mathbf{x}_{k-1}}^{\mathbf{x}_k} \log p(\mathbf{x}_k|\mathbf{x}_{k-1})\},\\
\mathbf{D}_{k-1}^{21} &= \mathbb{E}\{-\Delta_{\mathbf{x}_k}^{\mathbf{x}_{k-1}} \log p(\mathbf{x}_k|\mathbf{x}_{k-1})\} = \left[\mathbf{D}_{k-1}^{12}\right]^T,\\
\mathbf{D}_{k-1}^{22} &= \mathbb{E}\{-\Delta_{\mathbf{x}_k}^{\mathbf{x}_k}[\log p(\mathbf{x}_k|\mathbf{x}_{k-1}) + \log p(\mathbf{z}_k|\mathbf{x}_k)]\}.
\end{aligned} \tag{19}$$

4.1 PCRB of RSS-Based DFLT

Considering the nonlinear filtering problem with additive Gaussian noise defined in (1), the conditional PDFs in (19) are

$$\begin{aligned}
-\log p(\mathbf{x}_k|\mathbf{x}_{k-1}) &= c_1 + \frac{1}{2}(\mathbf{x}_k - \mathbf{f}(\mathbf{x}_{k-1}))^T\\
&\quad \cdot \mathbf{Q}^{-1}(\mathbf{x}_k - \mathbf{f}(\mathbf{x}_{k-1})),\\
-\log p(\mathbf{z}_k|\mathbf{x}_k) &= c_2 + \frac{1}{2}(\mathbf{z}_k - \mathbf{h}(\mathbf{x}_k))^T\\
&\quad \cdot \mathbf{R}^{-1}(\mathbf{z}_k - \mathbf{h}(\mathbf{x}_k)),
\end{aligned} \tag{20}$$

where c_1 and c_2 are constant scaling terms of the multivariate distributions. Plugging (20) to (19) yields

$$\begin{aligned}
\mathbf{D}_{k-1}^{11} &= \mathbb{E}\{(\nabla_{\mathbf{x}}\mathbf{f}(\mathbf{x}_{k-1}))\mathbf{Q}^{-1}(\nabla_{\mathbf{x}}\mathbf{f}(\mathbf{x}_{k-1}))^T\},\\
\mathbf{D}_{k-1}^{12} &= -\mathbb{E}\{\nabla_{\mathbf{x}}\mathbf{f}(\mathbf{x}_{k-1})\}\mathbf{Q}^{-1},\\
\mathbf{D}_{k-1}^{22} &= \mathbf{Q}^{-1} + \mathbb{E}\{(\nabla_{\mathbf{x}}\mathbf{h}(\mathbf{x}_k))\mathbf{R}^{-1}(\nabla_{\mathbf{x}}\mathbf{h}(\mathbf{x}_k))^T\},
\end{aligned} \tag{21}$$

and since $\mathbf{F}$ is linear, and $\mathbf{h}(\cdot)$ nonlinear for the problem defined in (1), (21) simplifies to

$$\begin{aligned}
\mathbf{D}_{k-1}^{11} &= \mathbf{F}^T\mathbf{Q}^{-1}\mathbf{F},\\
\mathbf{D}_{k-1}^{12} &= -\mathbf{F}^T\mathbf{Q}^{-1},\\
\mathbf{D}_{k-1}^{22} &= \mathbf{Q}^{-1} + \mathbf{H}_{\mathbf{x}}^T\mathbf{R}^{-1}\mathbf{H}_{\mathbf{x}},
\end{aligned} \tag{22}$$

TABLE 2
Experimental Parameters

Parameter		
Regularization parameter in (11)	α	500
Pixel width	δ_p	0.25 (m)
Correlation distance in (13)	δ_d	2 (m)
Image process noise in (34)	q_i	0.1 (dB/s^2)
Image measurement noise in (34)	σ_i^2	0.03 (dB^2)
Image threshold (37)	γ	0.75
RSS model parameter (4)	ϕ	−2.22
RSS model parameter (4)	λ	0.04
Measurement noise (1)	σ^2	1.42

where $\mathbf{H}_{\mathbf{x}}$ is defined in (7). Now, the recursion to update the FIM can be obtained by substituting (22) to (18) giving

$$\begin{aligned}
\mathbf{J}_k &= \mathbf{H}_{\mathbf{x}}^T\mathbf{R}^{-1}\mathbf{H}_{\mathbf{x}} + \mathbf{Q}^{-1}\\
&\quad - \mathbf{Q}^{-1}\mathbf{F}\left(\mathbf{J}_{k-1} + \mathbf{F}^T\mathbf{Q}^{-1}\mathbf{F}\right)^{-1}\mathbf{F}^T\mathbf{Q}^{-1},
\end{aligned} \tag{23}$$

which can be simplified to

$$\mathbf{J}_k = \mathbf{H}_{\mathbf{x}}^T\mathbf{R}^{-1}\mathbf{H}_{\mathbf{x}} + \left(\mathbf{Q} + \mathbf{F}\mathbf{J}_{k-1}^{-1}\mathbf{F}^T\right)^{-1}, \tag{24}$$

using the matrix inversion lemma. It is to be noted that the bound is not defined when the target location coincides with the position of the TX or RX since this results into division by zero in Eq. (6).

The PCRB states that for any unbiased estimator, the root-mean squared (RMS) localization error is lower bounded by

$$\text{RMSE}(k) \geq e_{pk} \triangleq \sqrt{\{\mathbf{J}_k^{-1}\}_{1,1} + \{\mathbf{J}_k^{-1}\}_{3,3}}, \tag{25}$$

where $\{\mathbf{J}_k^{-1}\}_{1,1}$ and $\{\mathbf{J}_k^{-1}\}_{3,3}$ denote elements of the PCRB matrix corresponding to the x- and y-coordinates. The PCRB can be used as a pre-deployment predictor of localization accuracy, providing an analytical method for system design and pre-deployment performance evaluation. As an example, [10] uses the CRB to investigate the effect of node geometry to imaging accuracy, [35] uses it to analytically evaluate the developed model and [36] uses it for pre-deployment performance assessment. In the following section, we compare the RMS errors of RTI and the EKF to the lower bound for localization RMSE, denoted by e_{pk}, to evaluate the system and location estimators.

4.2 Bound Analysis

The bound analysis is conducted using a network of 20 sensors that cover an area of 75 m^2 and the devices are distributed around the monitored area as illustrated in Fig. 1a. The experimental setting is the same as the open environment experiment described in Section 6 but the data is simulated. For now, we consider a time-invariant system so that we can focus on localizing a stationary target. This assumption simplifies the PCRB analysis since time evolution can be neglected and power spectral density of the process noise can be set to zero, that is, $q = 0$. The PCRB and position estimates are calculated assuming full connectivity and that a single RSS measurement is available from each TX-RX pair. The empirical cumulative distribution function (ECDF) for the model parameter estimates in the open environment experiment are shown in

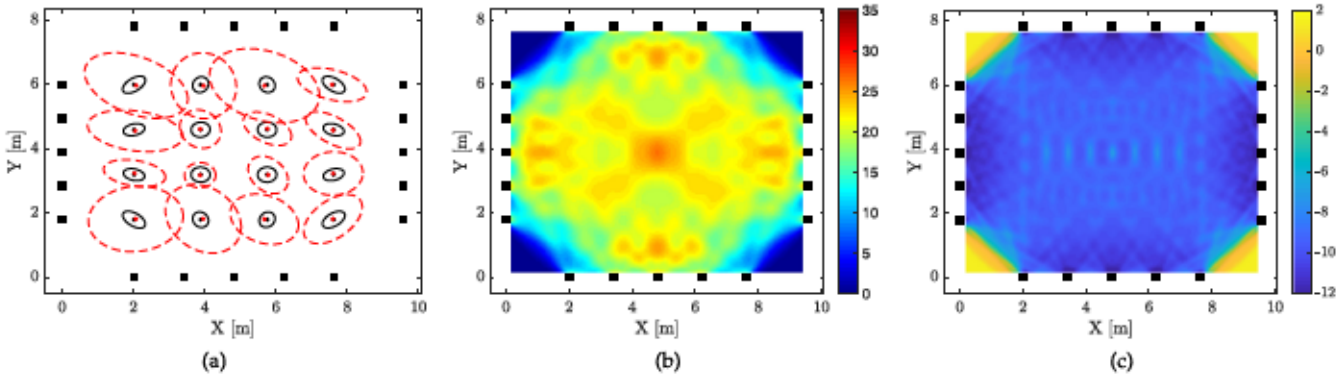

Fig. 1. In (a), the RTI estimator mean (•) and 3σ uncertainty ellipse (– – –) using 1000 trials for 16 example locations with respect to the true location (•) and PCRB on the 3σ uncertainty ellipse (——). In (b), the sum of link contributions $h_{tot}(\mathbf{p}_k) = \sum_{l=1}^{L} e^{-\Delta_l(k)/\lambda_l}$ as a function of target coordinates $\mathbf{p}_k$. In (c), lower bound of localization standard deviation in logarithmic scale, $10\log_{10}(e_{p_k}/1\text{ m})$, as a function of $\mathbf{p}_k$.

Fig. 6 and the median of the ECDFs is used in the following analysis. Thus, the RSS is modeled using: $\phi_l = -2.22$ dB, $\lambda_l = 0.04$ m and $\sigma_l^2 = 1.42$ dB^2 $\forall\, l$ and parameters of the filters are given in Table 2.

In Fig. 1a, 16 example locations and the lower bound on 3σ uncertainty ellipses illustrated together with the mean of the RTI estimates and 3σ uncertainty ellipse. The average lower bound on localization RMSE is $\bar{e}_p = \frac{1}{16}\sum_k^{16} e_{pk} = 0.1191$ m and the RMSE of RTI estimates is 0.3980 m, over three times higher than the bound. The advantage of RTI is that the estimator does not require accurate priori information about state of the target and the method can locate the target as long as a sufficient number of links intersect the location. As a disadvantage, the two-step estimator is inefficient as it never reaches the lower bound as shown by the uncertainty ellipses in Fig. 1a. The other disadvantage is that measurement noise can result in images with multiple peaks leading to inaccurate position estimates. An example of a noisy RTI image is illustrated in Fig. 3b which leads to an inaccurate position estimate and as a result, the uncertainty ellipse of the RTI position estimator is very large in many of the positions in Fig. 1a.

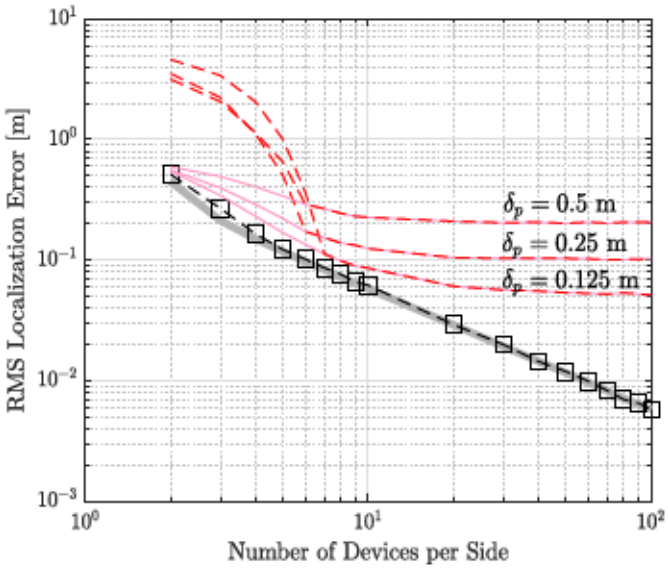

Fig. 2. The PCRB as a function of devices per side (——), and the RMS errors of EKF (–⊟–) and RTI w/ (– – –) and w/o outliers (——) obtained over 1000 trials. An estimate is considered an outlier if it is one meter or more from the true location.

For the node configuration shown in Fig. 1a, the sum of link contributions, $h_{tot}(\mathbf{p}_k) = \sum_{l=1}^{L} e^{-\Delta_l(k)/\lambda_l}$, is calculated and plotted as a function of target location $\mathbf{p}_k$ in Fig. 1b. Respectively, the PCRB on RMSE is computed and illustrated in Fig. 1c. With the given model parameters, the PCRB is entirely defined by the geometrical relationship of the target and nodes, and the PCRB decreases as the number of wireless links that intersect the location increases. The minimum PCRB is 0.0449 m and the lowest values are found beside the wireless devices on the side that faces the monitored area. Respectively, the maximum PCRB is 1.4142 m and the worst localization accuracy is expected in the corners of the monitored area where none of the links intersect the location (see Fig. 1b). On average, the PCRB is 0.1858 m and 80 percent of the monitored area has a PCRB of 0.1308 m or lower. Thus, the localization accuracy is expected to be high and similar performance is anticipated throughout the monitored area as long as the target is not close to the borders.

In Fig. 2, the PCRB and RMS errors of the estimators illustrated as a function of number of devices per side when they are equally spaced on the side of a $X = 9.58$ m by $Y = 7.82$ m deployment. The PCRB is a monotonically decreasing function and the lower bound is cut to half every time the number of devices per side is doubled. The RMS errors of the RTI estimates are illustrated using three pixel size values and as shown, lower RMS errors can be achieved with higher image resolution but with the expense of increased computational complexity. Interestingly, the RMS error of the RTI estimates converge toward $\bar{e}_p \to \delta_p/\sqrt{6}$[1]. This value is the lower bound for RTI and it requires that the target can always be positioned to within the correct pixel. The EKF is not constrained by discretization of the monitored area resulting to lower RMS errors as shown in Fig. 2. Moreover, the EKF is an efficient estimator as it nearly achieves the bound already with a low number of devices and it converges toward the PCRB as the number of sensors increases.

1. If the estimate locates within the correct pixel, the coordinate errors can be considered as i.i.d. uniform random variables $X \sim U(-\frac{1}{2}\delta_p, \frac{1}{2}\delta_p)$ and $Y \sim U(-\frac{1}{2}\delta_p, \frac{1}{2}\delta_p)$. Then, the MSE is $\mathrm{E}[e_p^2] = \mathrm{E}[X^2 + Y^2] = \mathrm{E}[X^2] + \mathrm{E}[Y^2] = \delta_p^2/6$, where the expected value of the uniformly distributed zero-mean random variable is known to be $\mathrm{E}[X^2] = \mathrm{E}[Y^2] = \delta_p^2/12$ [46, Ch.6]. Thus, the RMS error is $\mathrm{E}[e_p] = \delta_p/\sqrt{6}$ and this value is the lower bound for RTI location estimates.

In the analysis of Fig. 2, the EKFs prior on the mean and covariance are set as $\mathbf{x}_0 = \mathbf{x}_{true} + \mathbf{Q}_0\mathbf{w}$ and $\mathbf{P}_0 = 3\mathbf{J}^{-1} + \mathbf{Q}_0$, where $\mathbf{x}_{true}$ denotes the true state and $\mathbf{J}$ is the FIM at $\mathbf{x}_{true}$, $\mathbf{Q}_0 = \text{diag}\big(\begin{bmatrix} 0.25^2\ \text{m}^2 & 0.0\ (\text{m/s})^2 & 0.25^2\ \text{m}^2 & 0.0\ (\text{m/s})^2 \end{bmatrix}\big)$ and $\mathbf{w} \in \mathbb{R}^{4\times 1}$ is a zero-mean white Gaussian noise vector. The EKF requires an accurate prior or otherwise the filter can diverge and result to a biased estimate. This matter will be further discussed in the following.

4.3 Discussion

Ideally, we would prefer to use the EKF since it is able to achieve the PCRB and the filter is computationally efficient. However, the filter has two important limitations. First, the filter is prone to diverge if modeling errors exist. Second, the filter requires the prior distribution to locate near the true state. These reasons restrain using the EKF in real world applications since the model parameters are known to vary for each link [12] and the person's location is not known when initializing the filter. RTI can be considered as complementary to the EKF, since RTI is significantly more robust to modeling errors and it can be initialized without prior information of the person's location. Due to these reasons, we present a novel Bayesian filter in the next section that has the beneficial properties of both approaches.

5 Fusion Algorithm

5.1 Sampling and Processing

In RSS-based DFLT, the sensors are programmed to transmit and receive packets from other sensors of the network. Typically, the communication schedule follows a token passing protocol where one sensor transmits at a time while the others are in reception mode [47]. After transmission, the turn is assigned to the next sensor in the schedule following the sensor IDs in sequential order. In the packets, the sensors include the most recent RSS associated with the transmissions of others'.

Let the wireless network consist of S sensors, then, when the last sensor in the schedule transmits at time k the payload of the packet is

$$\mathbf{z}_S(k) = \begin{bmatrix} z_{S,1}(k-S+1) & z_{S,2}(k-S+2) & \cdots & z_{S,S-1}(k-1) & 0 \end{bmatrix}^T,$$

where $z_{i,j}(n)$ denotes the RSS that is transmitted by node j and received by node i at time n. A base station listens to the ongoing transmissions and it stacks the packets to form a measurement matrix

$$\{\mathbf{Z}\}_k^{k-S+1} = \begin{bmatrix} \mathbf{z}_1(k-S+1) & \cdots & \mathbf{z}_S(k) \end{bmatrix}^T, \tag{26}$$

containing the RSS measurements of the last S packets before processing. We denote this time interval as the communication cycle and it contains a transmission from each node and the elements of $\mathbf{Z}$ are

$$\{\mathbf{Z}\}_k^{k-S+1} = \begin{bmatrix} 0 & z_{1,2}(k-2S+2) & \cdots & z_{1,S}(k-S) \\ z_{2,1}(k-S+1) & 0 & \cdots & z_{2,S}(k-S) \\ \vdots & \vdots & \ddots & \vdots \\ z_{S,1}(k-S+1) & z_{S,2}(k-S+2) & \cdots & 0 \end{bmatrix}.$$

The measurements of $\mathbf{Z}$ are clearly taken at different time instances and a sequential processing scheme was introduced in [16] to resolve this issue. The proposed scheme processes the RSS one transmission at a time and the measurements can be associated to the same time instant by delaying the processing by $S-1$ samples so that when TX j transmits at time k, the RSS corresponding to the transmission of TX

$$i = \begin{cases} S-j+1 & \text{if } j \equiv S \\ j+1 & \text{otherwise} \end{cases},$$

becomes available at the base station. Using the time notation $n = k - S + 1$ we can write

$$\mathbf{z}_{n,i} \triangleq \text{col}_i\{\mathbf{Z}\}_k^{k-S+1}. \tag{27}$$

As an example, when the transmitter ID is $j = S$, the measurements of node $i = 1$ can be processed since the first column of the RSS matrix, $\mathbf{z}_{n,1} = \text{col}_1\{\mathbf{Z}\}_k^{k-S+1} = \begin{bmatrix} 0 & z_{2,1}(k-S+1) & \cdots & z_{S,1}(k-S+1) \end{bmatrix}^T$, are related to the same time instant $k-S+1$. At the next time instant, the transmitter ID is $j = 1$ and the measurements of $\{\mathbf{Z}\}_{k+1}^{k-S+2}$ are updated accordingly. Now the measurements of node $i = 2$ can be processed since the second column of the RSS matrix, $\mathbf{z}_{n,2} = \text{col}_2\{\mathbf{Z}\}_{k+1}^{k-S+2} = \begin{bmatrix} z_{1,2}(k-S+2) & 0 & \cdots & z_{S,2}(k-S+2) \end{bmatrix}^T$, are related to the same time instant $k-S+2$.

The sequential processing scheme improves the accuracy and robustness of RSS-based DFLT [16] and taking the time evolution correctly into account as given in (27), the problem defined in (1) can now be written as

$$\mathbf{x}_n = \mathbf{F}\mathbf{x}_{n-1} + \mathbf{q}_{n-1}, \tag{28a}$$

$$\mathbf{z}_{n,i} = \mathbf{h}(\mathbf{x}_n) + \mathbf{r}_n, \tag{28b}$$

where n denotes the time, $\mathbf{z}_{n,i} \in \mathbb{R}^{S\times 1}$ is the new RSS measurement vector and the corresponding measurement noise covariance is $\mathbf{R}_i \in \mathbb{R}^{S\times S}$. Using $\mathbf{z}_{n,i}$, the noise covariance, measurement model vector and Jacobian of the EKF are replaced with

$$\begin{aligned} \{\mathbf{R}_i\}_{j,j} &= \sigma_l^2, \\ \{\mathbf{h}_i(\hat{\mathbf{x}}_n^-)\}_{j,1} &= h_l(\hat{\mathbf{x}}_n^-), \\ \{\mathbf{H}_{\mathbf{x},i}(\hat{\mathbf{x}}_n^-)\}_{j,\cdot} &= \begin{bmatrix} \frac{\delta h_l}{\delta p_x} & 0 & \frac{\delta h_l}{\delta p_y} & 0 \end{bmatrix}, \end{aligned} \tag{29}$$

where i and j are the TX and RX IDs in respective order, $j = 1 \ldots S$ and $l = (i-1)\cdot S + j$.

5.2 A Sequential Imaging Method

5.2.1 Image Filtering

The drawback of using the RTI formulation presented in Section 3.3 is that the estimator requires the complete RSS measurement matrix $\mathbf{Z}$ and it computes a batch estimate of the changes within that communication cycle. In the following, an image filter is presented which allows us to recursively estimate the propagation field image every time new measurements become available. Moreover, the time evolution is more accurate with the proposed method.

The RTI solution in Eq. (11) is equivalent to forming the image by summing together the link contributions

$$\hat{\mathbf{b}}_n = \sum_{i=1}^{S} \mathbf{\Pi}_i \mathbf{z}_{n,i}, \tag{30}$$

where the jth column of $\mathbf{\Pi}_i$ is defined as

$$\{\mathbf{\Pi}_i\}_{\cdot,j} = \{\mathbf{\Pi}\}_{\cdot,l} \tag{31}$$

and indexes i, j and l are defined as in (29). Now, instead of summing the link contributions to form the image, an adaptive filter is implemented to track the changes in the image using $\mathbf{\Pi}_i \mathbf{z}_{n,i}$.

The designed filter is similar to the one presented in [12] with the difference that the presented filter estimates the images from the link contributions $\mathbf{\Pi}_i \mathbf{z}_{n,i}$ whereas the one in [12] uses $\hat{\mathbf{b}}_k$. In the filter, the image state is expressed as

$$\mathbf{m} = \begin{bmatrix} m_1 & m_2 & \cdots & m_N \\ \dot{m}_1 & \dot{m}_2 & \cdots & \dot{m}_N \end{bmatrix}, \tag{32}$$

where m_j denotes the image intensity and $\dot{m}_j$ change rate of voxel j. The state-space model is

$$\mathbf{m}_n = \mathsf{F}\mathbf{m}_{n-1} + \mathbf{q}_{n-1}, \tag{33a}$$

$$\mathbf{b}_n = \mathsf{H}\mathbf{m}_n + \mathbf{r}_n, \tag{33b}$$

where $\mathbf{q}_{n-1} \sim \mathcal{N}(\mathbf{0}, \mathsf{Q})$ is the process noise and $\mathbf{r}_n \sim N(\mathbf{0}, \mathsf{R})$ the measurement noise. The image is modeled using a discrete white noise acceleration model (DWNA) as in [12] for which the transition matrix, measurement model and noises are [43, Ch.6]

$$\mathsf{F} = \begin{bmatrix} 1 & \tau \\ 0 & 1 \end{bmatrix}, \quad \mathsf{Q} = q_i \begin{bmatrix} \frac{1}{3}\tau^3 & \frac{1}{2}\tau^2 \\ \frac{1}{2}\tau^2 & \tau \end{bmatrix}, \quad \mathsf{H} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}^T, \quad \mathsf{R} = \sigma_i^2 \mathbf{I}_N, \tag{34}$$

where q_i is the power spectral density of the image process noise. Given the models, it is straightforward to implement a KF for tracking state of the image. The prediction step of the KF is [17, Ch.4]

$$\begin{aligned} \hat{\mathbf{m}}_n^- &= \mathsf{F}\hat{\mathbf{m}}_{n-1}, \\ \mathbf{C}_n^- &= \mathsf{F}\mathbf{C}_{n-1}\mathsf{F}^T + \mathsf{Q}. \end{aligned} \tag{35}$$

Thereafter, the mean $\hat{\mathbf{m}}_n^-$ and covariance $\mathbf{C}_n^-$ are updated when measurement $\mathbf{z}_{n,i}$ becomes available using [17, Ch.4]

$$\begin{aligned} \mathbf{S}_n &= \mathsf{H}\mathbf{C}_n^-\mathsf{H}^T + \mathsf{R}, \\ \mathbf{K}_n &= \mathbf{C}_n^-\mathsf{H}^T\mathbf{S}_n^{-1}, \\ \hat{\mathbf{m}}_n &= \hat{\mathbf{m}}_n^- + \mathbf{K}_n\left(\mathbf{\Pi}_i\mathbf{z}_{n,i} - \mathsf{H}\hat{\mathbf{m}}_n^-\right), \\ \mathbf{C}_n &= \mathbf{C}_n^- - \mathbf{K}_n\mathbf{S}_n\mathbf{K}_n^T. \end{aligned} \tag{36}$$

5.2.2 Positioning

If a single person locates within the area, it is expected that the pixels with highest intensity locate near the target and therefore, localizing the person can be postulated as finding the mode of $\hat{\mathbf{b}}_n \triangleq \mathsf{H}\hat{\mathbf{m}}_n$ [48]. If $\mathcal{I} = \max(\hat{\mathbf{b}}_n)$ denotes the maximum component of the spatial field, than the mode is in the set of pixels with intensity higher than $\gamma\mathcal{I}$. To simplify the notation, let us define

$$\tilde{\mathbf{b}}_n = \begin{cases} \hat{\mathbf{b}}_n & \text{if } \hat{\mathbf{b}}_n \geq \gamma\mathcal{I} \\ 0 & \text{otherwise} \end{cases}, \tag{37}$$

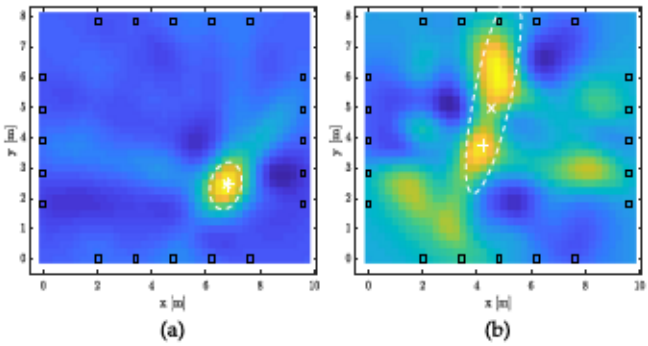

Fig. 3. Two example RTI images and the position and covariance estimates calculated using (38) and (39). In the image, the plus sign indicates the true position, the crosses are the position estimates and the dashed line illustrates the 3σ uncertainty ellipse which is defined as $\hat{\mathbf{p}}_n + \sqrt{\mathbf{N}_n}[\cos(\beta)\ \sin(\beta)]^T$ in which $\beta = [0, \ldots, 2\pi]$ and $\sqrt{\mathbf{N}_n}$ denotes the lower Cholesky factorization of $\mathbf{N}_n$ such that $\mathbf{N}_n = \sqrt{\mathbf{N}_n}\sqrt{\mathbf{N}_n}^T$.

and $\mathbf{w}_n = \tilde{\mathbf{b}}_n / \sum \tilde{\mathbf{b}}_n$. Now, the position can be estimated as the weighted sum of pixels

$$\hat{\mathbf{p}}_n = \mathbf{p}\mathbf{w}_n, \tag{38}$$

where $\mathbf{p} \in \mathbb{R}^{2\times N}$ are the pixel coordinates and $\mathbf{w}_n \in \mathbb{R}^{N\times 1}$ the pixel weights. The covariance of the estimate is defined as

$$\mathbf{N}_n = \mathbf{w}_n \circ \left(\mathbf{p} - \hat{\mathbf{p}}_n\right)\left(\mathbf{p} - \hat{\mathbf{p}}_n\right)^T, \tag{39}$$

where $\circ$ is the Hadamard product. Note that the localization proposed in (38) is only capable of locating one person. Multi-target localization and tracking is a challenging task in DFLT, and it is outside the scope of this paper. Thus, we do not propose coordinate estimators for the multi-target case and for now, readers are referred to [15], [26], [31], [49] for RSS-based multi-target tracking.

Two example RTI images are shown in Fig. 3 together with the position and covariance estimates. The image on the left shows that the pixels with $\hat{\mathbf{b}}_n \geq \gamma\mathcal{I}$ are centered around the true location, the position estimate is accurate, and the estimated covariance is small. The image on the right is noisy and does not clearly indicate the person's location. The estimated position is over a meter away from the true location and the estimated covariance is significantly higher then in the other image. Estimating the covariance allows to take such uncertainties into account and the Kalman filter gives less weight to position estimates that are estimated from noisy images.

5.3 Fusion Filter

The FF composes of two filters running in parallel. The first one is the image filter presented in Section 5.2 and it tracks the changes in the discretized propagation field image. The second filter is the target tracking filter that is implemented using the EKF presented in Section 3.2. However, the update step of the EKF is augmented with position measurements from the imaging solution in order to bound the EKFs measurement residuals by the position errors of the imaging approach. Recursion of the FF at time step n is presented in the following and pseudocode of the filter is given in Algorithm 1. The filter recursion can be divided into three stages:

1) *Predict*—Prediction step of the image and target filter.

(a) open environment (b) apartment environment

Fig. 4. The experimental layouts in which the nodes (□) and the validation positions (+) are illustrated. The apartment was fully furnished with two beds, nightstands, cupboards, a sofa, coffee table, etc., but the furniture is omitted from the figure.

2) *Model concatenation*—Forming the measurement model matrices.
3) *Update*—Update step of the target filter.

Algorithm 1. Fusion Filter at Time Step n

1: **procedure** PREDICT
2: Form $\mathbf{F}$, $\mathbf{Q}$, F and Q using τ
3: Predict $\hat{\mathbf{m}}_n^-$ and $\mathbf{C}_n^-$ using (35) ▷ Image Prediction
4: Predict $\hat{\mathbf{x}}_n^-$ and $\mathbf{P}_n^-$ using (8) ▷ Target Prediction
5: **end procedure**
6: **procedure** MODEL CONCATENATION
7: Update image filter using (36)
8: Estimate $\hat{\mathbf{p}}_n$ and $\mathbf{N}_n$ from image using (38) and (39)
9: Form $\tilde{\mathbf{R}}$, $\tilde{\mathbf{z}}_n$, $\tilde{\mathbf{H}}_{\mathbf{x}}(\hat{\mathbf{x}}_n^-)$ and $\tilde{\mathbf{h}}(\hat{\mathbf{x}}_n^-)$ using (40)
10: **end procedure**
11: **procedure** UPDATE
12: $\mathbf{S}_n = \tilde{\mathbf{H}}_{\mathbf{x}}(\hat{\mathbf{x}}_n^-)\mathbf{P}_n^-\tilde{\mathbf{H}}_{\mathbf{x}}^T(\hat{\mathbf{x}}_n^-) + \tilde{\mathbf{R}}$,
13: $\mathbf{K}_n = \mathbf{P}_n^-\tilde{\mathbf{H}}_{\mathbf{x}}^T(\hat{\mathbf{x}}_n^-)\mathbf{S}_n^{-1}$,
14: $\hat{\mathbf{x}}_n = \hat{\mathbf{x}}_n^- + \mathbf{K}_n(\tilde{\mathbf{z}}_n - \tilde{\mathbf{h}}(\hat{\mathbf{x}}_n^-))$,
15: $\mathbf{P}_n = \mathbf{P}_n^- - \mathbf{K}_n\mathbf{S}_n\mathbf{K}_n^T$
16: **end procedure**

In the prediction step, matrices F and Q in (34), and $\mathbf{F}$ and $\mathbf{Q}$ in (3), are formed using sampling interval τ and thereafter, prediction step of the filters is performed. In the model concatenation step, measurement $\mathbf{z}_{n,i}$ of TX i becomes available and the image filter can be updated. From the filtered image $\hat{\mathbf{m}}_n$, the position and covariance are estimated as described in Section 5.2.2. Thereafter, the measurement model matrices for the FF are formed as follows

$$\begin{aligned}\tilde{\mathbf{R}} &= \text{blkdiag}\,(\mathbf{R}_i, \mathbf{N}_n),\\ \tilde{\mathbf{z}}_n &= \begin{bmatrix}\mathbf{z}_{n,i}^T & \hat{\mathbf{p}}_n^T\end{bmatrix}^T,\\ \tilde{\mathbf{H}}_{\mathbf{x}}(\hat{\mathbf{x}}_n^-) &= \begin{bmatrix}\mathbf{H}_{\mathbf{x},i}^T(\hat{\mathbf{x}}_n^-) & \mathbf{H}^T\end{bmatrix}^T,\\ \tilde{\mathbf{h}}(\hat{\mathbf{x}}_n^-) &= \begin{bmatrix}\mathbf{h}_i^T(\hat{\mathbf{x}}_n^-) & (\mathbf{H}\hat{\mathbf{x}}_n^-)^T\end{bmatrix}^T,\end{aligned} \tag{40}$$

where $\mathbf{R}_i$, $\mathbf{h}_i(\hat{\mathbf{x}}_n^-)$ and $\mathbf{H}_{\mathbf{x},i}(\hat{\mathbf{x}}_n^-)$ are defined in (29) and

$$\mathbf{H} = \begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 0 & 1 & 0\end{bmatrix}.$$

Finally, in the update step, the Fusion Filter can be updated as given by lines 12 through 15 in Algorithm 1.

6 Experiments

This section begins by introducing the experimental setup. Thereafter, the different filtering algorithms are summarized and initialization of the filters is discussed. The section is concluded by presenting the evaluation metrics.

6.1 Experimental Setup

The used wireless sensors are Texas Instruments CC2531 USB dongles operating at the 2.4 GHz ISM band and the sensors communicate in TDMA fashion as described in Section 5.1 and further explained in [47]. The experiments are conducted using all 16 frequency channels for communication in order to increase the system performance [19] and the used channels are 11−26 as defined by the IEEE 802.15.4 standard [50]. The transmission interval between communications is approximately $\tau \approx 2.9$ ms, which defines the sampling period for the filtering algorithms.

The experiments are conducted with 20 nodes that are deployed in two different environments as illustrated in Fig. 4. In the open environment, the nodes are deployed around the monitored area and the network covers a 75 m^2 area. The nodes are set on top of podiums at approximately waist height (≈ 0.9 m). The floor plan of the apartment is 82 m^2 and the nodes are deployed so that the person can be located throughout the entire house. To replicate a realistic deployment scenario, 18 nodes are installed by electric sockets of the apartment so that we could power them using AC adapters. The walk-in closets did not have electric sockets on the exterior walls, so we decided to deploy one battery powered node in each to ensure coverage of the entire apartment. These two nodes are located at $[0.08\ 2.89]^T$ and $[10.24\ 2.80]^T$.

In the experiments, markers are placed inside the monitored area for the test person to follow. During the experiment, the person always walks directly from one validation position (as illustrated in Fig. 4 using +) to another and once the person reaches the destination, they stop for a few seconds before proceeding to the next validation position. In both environments, three different trials are conducted and in each trial, the person is inside the area for approximately three minutes and every validation position is visited at least once. During the experiments, the person is carrying a video camera. In post-processing, the RSS and video streams are synchronized and the video is used to generate the ground truth trajectory. However, we evaluate the tracking accuracy only when the person is stationary to ensure that the ground truth position is correct.

A recent work has presented the means for an adaptive RTI (ARTI) system to use unsupervised training for estimating the unknown model parameters [12]. This improves the RSS models over time as data is gathered and enhances the system performance. We could adapt similar algorithms in this paper but for simplicity, we calibrate the model parameters using the ARTI system [12] and then use these estimates in our filtering algorithms. Thus, the RSS model parameters $\boldsymbol{\theta}_l = \begin{bmatrix} \phi_l & \lambda_l & \sigma_l^2 & \mu_l \end{bmatrix}$ are estimated by minimizing the cost function

$$J(\boldsymbol{\theta}_l) = \sum_{n=1}^{K} \left[z_l(n) - h_l(\hat{\mathbf{x}}_n; \phi_l, \lambda_l) \right]^2, \quad \begin{cases} -30 \leq \phi_l \leq 30 \\ 0.001 \leq \lambda_l \leq 1 \end{cases}, \tag{41}$$

where K is the total number of estimates in one trial (approx. 180 seconds of data), $z_l(n)$ is the measured RSS, $h_l(\cdot)$ is evaluated using ARTI state estimates $\hat{\mathbf{x}}_n$[2], and the dependence of $h_l(\cdot)$ on the model parameters is now explicitly stated. In this paper, constrained nonlinear optimization [51] is used to find the minimum of $J(\boldsymbol{\theta}_l)$ and thereafter, the maximum likelihood (ML) estimates of $\mu_l = \frac{1}{K}\sum_{n=1}^{K}(z_l(n) - h_l(\hat{\mathbf{x}}_n; \phi_l, \lambda_l))$ and $\sigma_l^2 = \frac{1}{K}\sum_{n=1}^{K}(z_l(n) - \mu_l - h_l(\hat{\mathbf{x}}_n; \phi_l, \lambda_l))^2$ are obtained.

6.2 Filtering Algorithms

In the following, the different filters are summarized. The process noise value of the filters has been used as a tuning parameter to maximize performance of each filter. Otherwise, a specific value might favor one filter over another.

- *FF*—The Fusion Filter can be implemented using Algorithm 1 and the imaging parameters of the system are given in Table 2. The selected parameter values work in a wide range of environments and they are close to the ones used in [12]. The process noise of the kinematic state is set to $q = 0.1\ \text{m/s}^2$.
- *EKF*—The first benchmark system is the EKF solution presented in [16]. The algorithm is summarized in Section 3.2, the filter uses sequential processing and the process noise is the same as for the FF.
- *PF*—The second benchmark system is a particle filter, which is the de facto nonlinear Bayesian filter used in RSS-based DFLT [13], [14], [15]. The used PF is a sequential importance resampling filter where the dynamic model is used as the importance distribution. The implemented PF uses 10000 particles, resampling is performed once the number of effective particles is below 1000 and the process noise is increased to $q = 1\ \text{m/s}^2$ and σ_l^2 is increased by one, to mitigate the sample depletion problem [52]. The readers are referred to [17], [53] for further details on PFs.
- *ARTI*—The third benchmark system is ARTI [12]. We modify the original ARTI algorithm so that the measurements can be processed sequentially as explained in Section 5.1. In addition, the online estimator (see Algorithm 2 in [12]) is disabled because the parameters are already estimated. The imaging and filtering parameters are given in Table 2 but the pixel width is decreased to $\delta_p = 0.15$ m so that discretization does not degrade the tracking accuracy. ARTI locates the person using (12) and a KF is implemented to track the target. The process noise of the filter is set to $q = 1\ \text{m/s}^2$ and the measurement noise covariance is $\mathbf{N} = \text{diag}(\begin{bmatrix} 0.25\ \text{m}^2 & 0.25\ \text{m}^2 \end{bmatrix})$.

6.3 Filter Initialization

The image state is initialized with an all zeros matrix and the covariance as $\mathbf{C}_0 = \text{diag}(\begin{bmatrix} 1\ \text{dB}^2 & 1\ (\text{dB/s})^2 \end{bmatrix})$. The target tracking filter is initialized using the true state of the target when the person has reached the first validation position and the covariance matrix is set to $\mathbf{P}_0 = \text{diag}(\begin{bmatrix} 0.1\ \text{m}^2 & 0.1\ (\text{m/s})^2 & 0.1\ \text{m}^2 & 0.1\ (\text{m/s})^2 \end{bmatrix})$. We want to make the following remarks: i) An inaccurate prior does not cause the imaging solutions to diverge, the methods are not particularly sensitive to initialization and only the first few estimates would be affected by inaccurate initialization. ii) The FF could be designed so that the algorithm would only use the position estimates of the imaging solution in the beginning. After convergence, the filter would switch to the augmented measurement model. iii) The EKF requires that the initial estimate is close to the true state. The priori estimate could be obtained for example using an imaging solution or initializing numerous EKFs simultaneously and keeping the one that converges. iv) The PF could be initialized by uniformly distributing the particles inside the monitored area, and with a sufficient number of particles, the posterior distribution is expected to converge to the true state. Initialization of the filters is not within the scope of this work, and for a fair comparison of the different filters, we have chosen to use the true state instead of implementing a different initialization procedure for each filter.

6.4 Evaluation Metrics

The filters are evaluated using the localization RMSE

$$\bar{e}_p = \sqrt{\frac{1}{K}\sum_{n=1}^{K}(e_p(n))^2}, \tag{42}$$

where K is the total number of estimates in one trial (approx. 180 seconds of data) and the distance error at sample n is calculated as

2. In the model calibration phase, the RTI position estimates are substituted with the person's true coordinates whenever the person is stationary and used as input in the filtering recursions.

TABLE 3
RMSE [cm] in the Open Environment Experiment

Scenario / Trial	FF	EKF	PF	ARTI
1 / 1	**75.2**	–	–	99.3
1 / 2	**103.6**	–	–	115.1
1 / 3	**80.8**	–	–	97.2
1 / average	**86.3**	–	–	103.9
2 / 1	10.0	–	**9.6**	22.7
2 / 2	10.9	**10.2**	11.1	21.7
2 / 3	12.3	**9.5**	10.1	19.9
2 / average	11.1	–	**10.3**	21.4

$$e_p(n) = \sqrt{(p_x(n) - \hat{p}_x(n))^2 + (p_y(n) - \hat{p}_y(n))^2},$$

in which p_x and p_y denote the true coordinates and the hat accent indicates the estimate. In addition, we use the ratio of measurements outside a defined threshold for examining robustness of the filters. The metric is defined as

$$e_{\%} = \left(1 - \frac{M}{K}\right) \cdot 100\%, \tag{43}$$

where M is the number of estimates within one meter of the true location. Typically, the filter has diverged from the correct trajectory if the estimate is one meter or further from the true position.

It is to be noted that the sequential processing scheme introduces a delay of $(S-1)\tau \approx 55$ ms to the estimates which corresponds to a distance error of 0.055m if the person moves at 1 m/s. This might have an impact in critical real-world applications if for example the system would be used in collaborative human-robot workspaces [22]. In non-critical applications such a delay can be neglected, especially when duration of the delay is known. In this paper, the delay caused by the sequential processing scheme can be calculated from the transmission times and the lag is removed from the position estimates before calculating the evaluation metrics given above.

7 Results

The development efforts are experimentally and numerically evaluated in this section and the FF is compared with respect to the EKF, PF and ARTI solutions. It is to be noted that the system performance strongly depends on the used model parameters and their accuracy. Thus, we investigate two scenarios:

- *Scenario 1:* Using an educated guess for the model parameters and the same value is used for all links.
- *Scenario 2:* Using the training scheme explained at the end of Section 6.1 and using the unique model parameter estimates for each link.

The ECDFs of the model parameter estimates in the open environment are illustrated in Fig. 6 and these values are used in scenario 2. In scenario 1, the educated guess is the median of the ECDFs and the RSS is modeled using: $\phi_l = -2.22$, $\lambda_l = 0.04$ and $\sigma_l^2 = 1.42\ \forall\ l$.

7.1 Experimental Results

The RMS errors of the filters in the different scenarios and trials are summarized in Table 3 for the open environment experiment. In scenario 1, when an educated guess is used for the model parameters, the EKF and PF always diverge because the filters can not tolerate significant modeling errors. The imaging solutions are more robust to such modeling errors and despite individual position estimates can be inaccurate, the imaging solutions are able to track the person at least to some extent. On average, the FF performs slightly better than ARTI and both methods are able to achieve an RMS error of approximately one meter or below.

Next, the filters are run on the same experimental data and using scenario 2. The EKF solution results to the lowest RMSE in two of the trials when the filter does not diverge. With respect to the EKF solution, the PF is more robust but it is not as accurate despite using 10000 particles. The lower accuracy results from increasing the process and measurement noises, which in our case was mandatory to avoid the sample depletion problem and divergence issues in the open environment. With respect to the nonlinear filters, ARTI has a higher RMSE because the position estimates are always affected by discretization of the image. In addition, the time evolution can not be solved as accurately because the image formation always requires low pass filtering which is performed with the KF in this paper. In this experiment, the FF never diverges and the accuracy of the filter is comparable to the other nonlinear filters and it provides a superior combination of robustness and accuracy.

The coordinate estimates of the filters in scenario 2 are illustrated in Fig. 5a for trial 1 and the position estimates are shown with respect to the validation positions in Fig. 5c for trial 2. On most parts, the coordinate estimates of the filters overlap one another. However, the ARTI estimates are slightly more spread out as shown in Fig. 5c and the trajectory is not as smooth as can be seen in Fig. 5a. In the figure, the time instance ($t = 158$ s) when the EKF diverges is also shown. The trajectory should be a straight line from one validation position to another but all filters result to inaccurate estimates with the difference that the EKF diverges and the other filters are able to recover. It is to be noted that if the modeling assumptions hold, the EKF is the best performing filter in the mean squared sense and no other filter can outperform it. However, the open environment experiment is actually very simplifying and modeling errors are very common due to the complex nature of the indoor propagation channel. Next, we show that the EKF and PF fail in more realistic deployment scenarios even though the model parameters are trained.

Next, the filters are run on experimental data from the apartment experiment and using scenario 2. The coordinate estimates of the filters are illustrated in Fig. 5b and the position estimates are shown with respect to the validation positions in Fig. 5d for trial 1. The apartment experiment is significantly more challenging since most of the nodes have non line-of-sight (NLOS) communication with one another and multipath propagation is common. As an outcome, the person's presence in between the transceivers does not necessarily cause the RSS to change. On the other hand, the person can alter static multipath components causing a significant RSS change even for large Δ values. These can result to

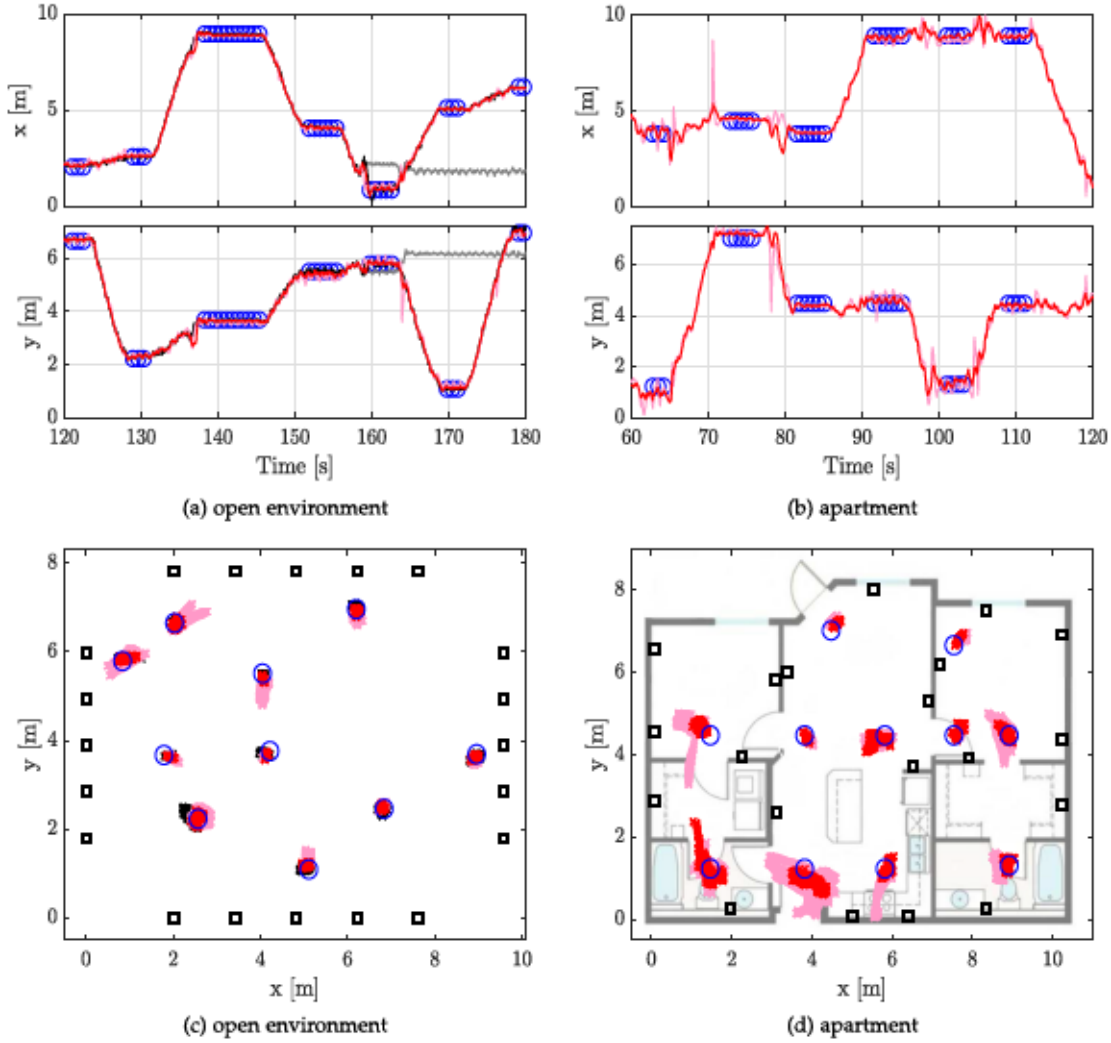

Fig. 5. *Experimental results* - In (a) and (b), coordinate estimates of the filters in the two environments. In the figures, the validation positions are illustrated with (○) and filter estimates are shown using: EKF (——), PF (——), ARTI (——) and FF (——). In (c) and (d), position estimates of the filters in the validation positions. In (b) and (d), the EKF and PF estimates are omitted from the figure because the filters have diverged.

significant modeling errors and the EKF and PF always diverge in the apartment experiment due to this reason. Thus, the coordinate estimates of the filters are omitted from Figs. 5b and 5d. Since the imaging solutions do not rely on an accurate image prior, the methods are significantly more robust to modeling errors and can track the person even in challenging environments as shown in the figures and as given in Table 4. More quantitatively, the average RMS error is 40.9 cm for ARTI and 28.6 cm for FF, decreasing the RMSE by 30 percent with respect to ARTI. In scenario 1, the results are similar to the ones in the open environment and the results are summarized in Table 4.

TABLE 4
RMSE [cm] in the Apartment Experiment

Scenario / Trial	FF	EKF	PF	ARTI
1 / 1	**80.3**	–	–	84.7
1 / 2	**91.0**	–	–	100.9
1 / 3	**82.9**	–	–	94.3
1 / average	**84.7**	–	–	93.3
2 / 1	**28.4**	–	–	43.0
2 / 2	**30.7**	–	–	39.4
2 / 3	**26.8**	–	–	40.3
2 / average	**28.6**	–	–	40.9

The average computation times to initialize the filters and to compute a single recursion are given in Table 5 for the open environment experiment. The results are obtained using a Matlab implementation and a standard laptop equipped with a 2.70 GHz Intel Core i7-4800MQ processor and 16 GB of RAM. Initializing the nonlinear filters is negligible, whereas the imaging methods require computing the projection matrix given in (11) which requires inverting large matrices. Complexity of matrix inversion is at least quadratic[3] and $N \propto (1/\delta_p)^2$ so that the overall complexity of calculating the projection matrix relates to the pixel size via $\mathcal{O}[(1/\delta_p)^4]$. The quartic proportion increases the computational complexity rapidly as the pixel size decreases, and ARTI requires seven times longer to initialize than FF. For ARTI, the initialization time can be reduced by using a larger pixel size, but at the same time, the RMS error increases. As an example, the RMSE of ARTI increases by

3. The complexity is at least $\mathcal{O}(n^2)$ because an $n \times n$ matrix has n^2 values but naive algorithms can have a complexity of $\mathcal{O}(n^3)$.

TABLE 5
CPU Times in Milliseconds

	FF	EKF	PF	ARTI
Initialization	2107.3	0.5	1.5	13920.3
Filter recursion	0.410	0.075	5.945	0.590

27 percent to 27.2 cm in the open environment experiment when $\delta_p = 0.25$ m.

In the experiments, the time interval between two transmissions is approximately 2.9 ms which defines the system's sampling rate and also sets an upper bound on how long a single filter recursion can last. As shown, the PF exceeds this value and real-time operation is not possible when 10000 particles are used. The computation time could be decreased by using fewer particles but at the same time, the filter becomes even more vulnerable to divergence issues. The other filters can easily be implemented in real-time and the EKF is superior to the other methods. However, its use is limited to very simplistic environments and when the model parameters are known. Lastly, the FF is more efficient than ARTI because the real-time computation of the image estimate has complexity $\mathcal{O}[NL]$ and $N \propto (1/\delta_p)^2$. In the next section, the PF uses $N = 1000$ particles so that a real-time implementation would be possible.

7.2 Numerical Results

The performance and differences between the filters is further analyzed in this section using simulations which replicate the open environment tests. This experimental setting is chosen because all filters were capable of tracking the person successfully in this environment. First, we compute the ECDFs for the model parameter estimates in the open environment. Thereafter, various distributions are fitted to the data and the one that maximizes the log likelihood is used to describe the model parameter. The ECDFs and the fitted distributions are illustrated in Fig. 6 and the selected distributions are: non-standardized Student's t-distribution $\phi \sim \mathcal{T}(-2.22, 4.09, 8.09)$, log-normal $\lambda \sim \mathcal{L}(-3.14, 1.23)$ and Gamma $\sigma \sim \Gamma(2.66, 0.51)$. Interestingly, the same distributions can be used to describe the apartment experiment model parameters and similar behavior was reported in [12]. This suggests that each parameter might follow a specific distribution where the parameters of the distribution are specified by the environment and layout of the network. Such information would aid pre-deployment performance assessment and could be used as informative priors to the model parameters. However, further investigation is left for future work.

In the simulations, the model parameters for each link are randomly drawn from the fitted distributions, and in an ideal scenario the model parameter is known. In reality, the model parameters are unknown in advance and an educated guess must be used. We use the median of the ECDFs as given in Table 2 and the DC term μ is assumed known for each link. In the simulations, the educated guess is used for link l randomly with probability P and we denote P as the fraction of links with incorrect θ_l. Note that the educated guess is likely to differ from the drawn parameter values. In the analysis, the filter performance between the two extremes is studied, that is, having perfect knowledge of the model parameters $P = 0\%$ and having no prior information of the model parameters $P = 100\%$.

The simulated trajectory and filter estimates averaged over 1000 Monte Carlo simulations are illustrated in Fig. 7a and the PCRB and RMS errors are shown in Fig. 7b when $P = 0\%$. As shown in Figs. 7a and 7b, the state estimates of the nonlinear filters are close to the true coordinates' while the ARTI estimates are not as accurate. As discussed earlier, ARTI performance is degraded by discretization and the estimates are lagging because the image formation always requires low pass filtering. For the y coordinate estimates in Fig. 7a, the impact of discretization can be seen between $t = 88{-}92$ s and the effect of low pass filtering between $t = 84{-}88$ s. The nonlinear filters do not have such drawbacks and as shown in Fig. 7b, the nonlinear filters nearly reach the PCRB. However, RMSE of the PF is slightly higher because the process and measurement noise of the filter have been increased to mitigate the sample depletion problem [52]. The time averaged PCRB is 2.53 cm over the 1000 Monte Carlo simulations and the RMSE of the filters in increasing order are: 3.01, 3.17, 3.62 and 14.42 cm for EKF, FF, PF and ARTI in corresponding order. Thus, the results are inline with the experimental results, that is, the EKF achieves the lowest RMSE in ideal scenarios, the FF and PF are slightly more inaccurate and ARTI has the highest RMSE in ideal scenarios. It is to be noted that acceleration of

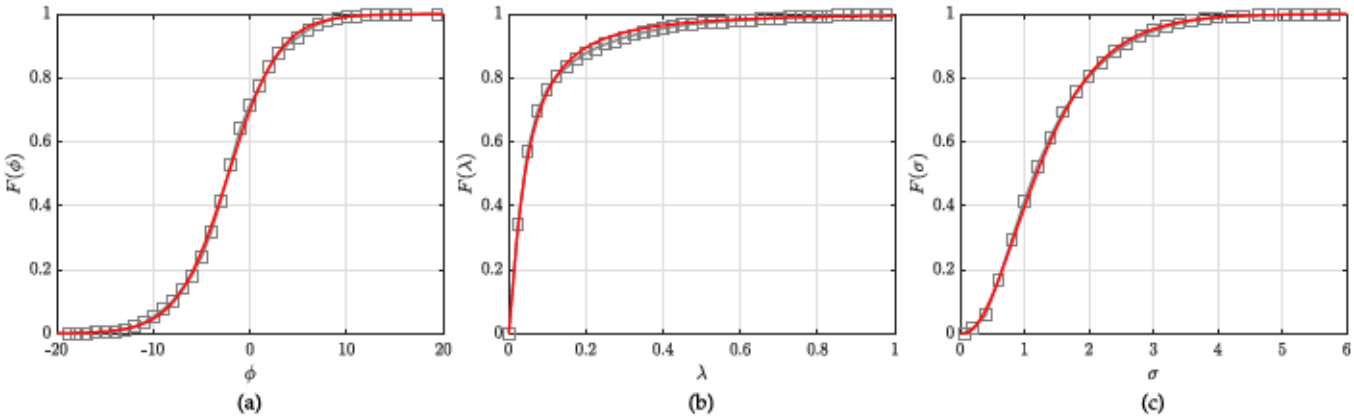

Fig. 6. The empirical CDFs for $\boldsymbol{\theta} = [\phi \ \lambda \ \sigma]$ in the open environment experiment (□) and the distribution fits (——). The fitted distributions are: non-standardized Student's t-distribution $\phi \sim \mathcal{T}(-2.22, 4.09, 8.09)$, log-normal $\lambda \sim \mathcal{L}(-3.14, 1.23)$ and Gamma $\sigma \sim \Gamma(2.66, 0.51)$.

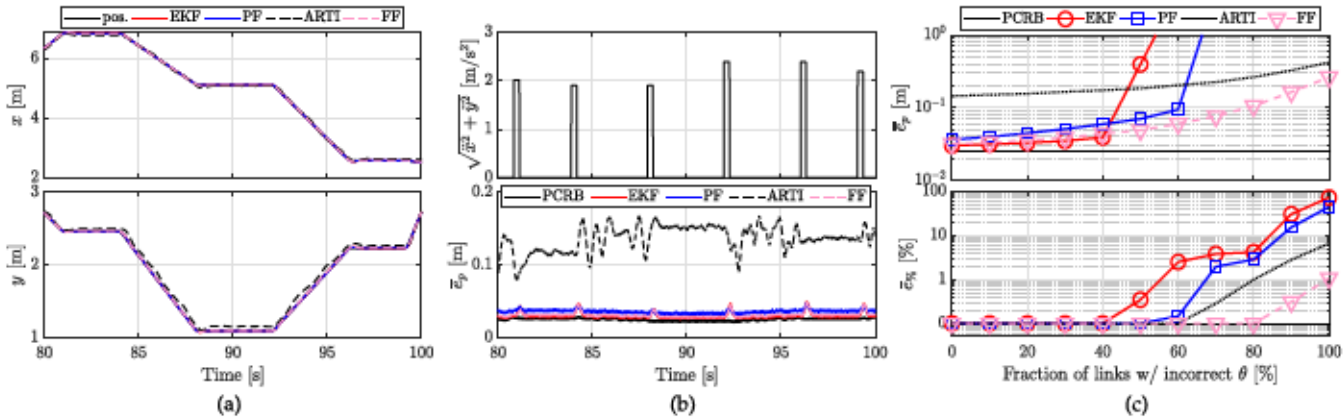

Fig. 7. In (a), the trajectory and coordinate estimates of the different filters averaged over the 1000 Monte Carlo simulations. In (b), the trajectory's acceleration magnitude shown above and the PCRB and RMS errors of the filters shown below. In (c), the evaluation metrics illustrated when the modeling error increases.

the person is non-zero for short time periods when the person stops and starts moving as shown in Fig. 7b. During these time intervals, the RMSE of the nonlinear filters increase as shown in Fig. 7b. A multiple model (MM) approach [43, Ch. 11] could be adopted where multiple filters would be run in parallel, each having a different process noise value. However, the improvement is expected to be marginal because maneuverability of the target is low and the maneuvers last for short time periods.

Next, robustness of the filters is studied by incrementing P so that the number of links that use the educated guess instead of the true model parameter value increases. The results are shown in Fig. 7c, and the advantages and disadvantages of the filters are clearly conveyed by the evaluation metrics. Ideally, the model parameters would be known and a good choice for the filter is an EKF due to its high tracking performance and low computational overhead. However, as the fraction of links with modeling error increases, the PF and EKF start to experience divergence issues and the filters can not be used as stand alone solutions. The PF can cope slightly better with the modeling errors but the improvement is insignificant considering the added computational complexity. The imaging solutions are significantly more robust to modeling errors and the filters are able to track the target even though all links are using an incorrect model parameter value. The main reason for this is that the imaging methods do not rely on an accurate prior when estimating the images. As shown in Fig. 7c, FF always outperforms ARTI and in the most challenging scenario when $P = 100\%$, the RMSE is 0.26 m with FF and 0.42 m with ARTI, an increase of 62 percent in tracking error. Thus, the numerical results support the experimental findings and we can conclude that the FF provides a superior combination of robustness and accuracy. It is to be noted that the numerical and experimental results do not correspond one another precisely because the used model is a simplification of the actual propagation phenomena and error sources.

8 Conclusion

The PCRB of RSS-based DFLT is derived in this paper and two estimators are evaluated with respect to the bound. The first estimator is a two-step imaging approach which first estimates the changes in the propagation field and then the person is localized from the image. The second estimator is a Bayesian filter which is realized using an EKF and the method directly relates the RSS measurements to the person's kinematic state. The bound analysis clearly shows that the EKF is efficient as it nearly reaches the bound and it is superior to the imaging approach in terms of localization accuracy. However, the EKF has practical limitations which restrict its use in real world deployments and in this paper, we address these limitations by introducing a Fusion Filter which merges the EKF and imaging solutions. The benefit of the proposed approach is that the filter's measurement residuals are bounded by the position errors of the imaging approach and as an outcome, the filter has the robustness of an imaging method and the tracking accuracy of a nonlinear Bayesian filter. The results imply that the presented filter nearly achieves the performance of the EKF in ideal scenario, and it is as robust as the imaging solution in non-ideal scenarios.

Developments of this paper open interesting opportunities in smoothing and parameter estimation since the Rauch-Tung-Striebel smoother can be directly applied to improve the state estimates of the presented filter. Moreover, the required expectations and maximization step for an expectation maximization algorithm can be computed in closed form using the smoothing distributions. These topics will be explored in future research.

Acknowledgments

This work was supported in part by the Academy of Finland Project #299099 and #295080. This work was also supported in part by the U.S. ARO under grant #69215CS.

References

[1] R. Yang, M. W. Newman, and J. Forlizzi, "Making sustainability sustainable: Challenges in the design of eco-interaction technologies," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, 2014, pp. 823–832.

[2] F. Adib, H. Mao, Z. Kabelac, D. Katabi, and R. C. Miller, "Smart homes that monitor breathing and heart rate," in *Proc. 33rd Annu. ACM Conf. Human Factors Comput. Syst.*, 2015, pp. 837–846.

[3] Q. Pu, S. Gupta, S. Gollakota, and S. Patel, "Whole-home gesture recognition using wireless signals," in *Proc. 19th Annu. Int. Conf. Mobile Comput. Netw.*, 2013, pp. 27–38.

[4] M. Youssef, M. Mah, and A. Agrawala, "Challenges: Device-free passive localization for wireless environments," in *Proc. 13th Annu. ACM Int. Conf. Mobile Comput. Netw.*, 2007, pp. 222–229.

[5] C. Alippi, M. Bocca, G. Boracchi, N. Patwari, and M. Roveri, "RTI goes wild: Radio tomographic imaging for outdoor people detection and localization," *IEEE Trans. Mobile Comput.*, vol. 15, no. 10, pp. 2585–2598, Oct. 2016.

[6] C. R. Anderson, R. K. Martin, T. O. Walker, and R. W. Thomas, "Radio tomography for roadside surveillance," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 1, pp. 66–79, Feb. 2014.

[7] J. Xiao, Z. Zhou, Y. Yi, and L. M. Ni, "A survey on wireless indoor localization from the device perspective," *ACM Comput. Surveys*, vol. 49, no. 2, pp. 25:1–25:31, Jun. 2016.

[8] S. Palipana, B. Pietropaoli, and D. Pesch, "Recent advances in RF-based passive device-free localisation for indoor applications," *Ad Hoc Netw.*, vol. 64, pp. 80–98, 2017.

[9] N. Patwari and P. Agrawal, "Effects of correlated shadowing: Connectivity, localization, and RF tomography," in *Proc. Int. Conf. Inf. Process. Sensor Netw.*, 2008, pp. 82–93.

[10] J. Wilson and N. Patwari, "Radio tomographic imaging with wireless networks," *IEEE Trans. Mobile Comput.*, vol. 9, no. 5, pp. 621–632, May 2010.

[11] J. Wilson and N. Patwari, "See-through walls: Motion tracking using variance-based radio tomography networks," *IEEE Trans. Mobile Comput.*, vol. 10, no. 5, pp. 612–621, May 2011.

[12] O. Kaltiokallio, R. Jäntti, and N. Patwari, "ARTI: An adaptive radio tomographic imaging system," *IEEE Trans. Veh. Technol.*, vol. 66, no. 8, pp. 7302–7316, Aug. 2017.

[13] Y. Li, X. Chen, M. Coates, and B. Yang, "Sequential Monte Carlo radio-frequency tomographic tracking," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, 2011, pp. 3976–3979.

[14] Z. Wang, H. Liu, S. Xu, X. Bu, and J. An, "A diffraction measurement model and particle filter tracking method for RSS-based DFL," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 11, pp. 2391–2403, Nov. 2015.

[15] Y. Guo, K. Huang, N. Jiang, X. Guo, Y. Li, and G. Wang, "An exponential-Rayleigh model for RSS-based device-free localization and tracking," *IEEE Trans. Mobile Comput.*, vol. 14, no. 3, pp. 484–494, Mar. 2015.

[16] O. Kaltiokallio, R. Hostettler, N. Patwari, and R. Jäntti, "Recursive Bayesian filters for RSS-based device-free localization and tracking," in *Proc. Int. Conf. Indoor Positioning Indoor Navigation*, Sep. 2018, pp. 1–8.

[17] S. Särkkä, *Bayesian Filtering and Smoothing*. Cambridge, U.K.: Cambridge Univ. Press, 2013.

[18] S. Jovanoska, R. Zetik, R. Thoma, F. Govaers, K. Wilds, and W. Koch, "Device-free indoor localization using a distributed network of autonomous UWB sensor nodes," in *Proc. Workshop Sensor Data Fusion: Trends Solutions Appl.*, 2013, pp. 1–6.

[19] O. Kaltiokallio, M. Bocca, and N. Patwari, "Enhancing the accuracy of radio tomographic imaging using channel diversity," in *Proc. IEEE 9th Int. Conf. Mobile Ad-Hoc Sensor Syst.*, 2012, pp. 254–262.

[20] M. McCracken, M. Bocca, and N. Patwari, "Joint ultra-wideband and signal strength-based through-building tracking for tactical operations," in *Proc. IEEE Int. Conf. Sens. Commun. Netw.*, 2013, pp. 309–317.

[21] C. Xu, M. Gao, B. Firner, Y. Zhang, R. Howard, and J. Li, "Towards robust device-free passive localization through automatic camera-assisted recalibration," in *Proc. 10th ACM Conf. Embedded Netw. Sensor Syst.*, 2012, pp. 339–340.

[22] S. Savazzi, V. Rampa, F. Vicentini, and M. Giussani, "Device-free human sensing and localization in collaborative human-robot workspaces: A case study," *IEEE Sensors J.*, vol. 16, no. 5, pp. 1253–1264, Mar. 2016.

[23] S. Kianoush, S. Savazzi, F. Vicentini, V. Rampa, and M. Giussani, "Device-free RF human body fall detection and localization in industrial workplaces," *IEEE Internet Things J.*, vol. 4, no. 2, pp. 351–362, Apr. 2017.

[24] Y. Chen, W. Dong, Y. Gao, X. Liu, and T. Gu, "Rapid: A multimodal and device-free approach using noise estimation for robust person identification," *Proc. ACM Interactive Mobile. Wearable Ubiquitous Technol.*, vol. 1, no. 3, pp. 41:1–41:27, Sep. 2017.

[25] M. Seifeldin, A. Saeed, A. E. Kosba, A. El-Keyi, and M. Youssef, "Nuzzer: A large-scale device-free passive localization system for wireless environments," *IEEE Trans. Mobile Comput.*, vol. 12, no. 7, pp. 1321–1334, Jul. 2013.

[26] C. Xu *et al.*, "SCPL: Indoor device-free multi-subject counting and localization using radio signal strength," in *Proc. ACM/IEEE Int. Conf. Inf. Process. Sensor Netw.*, 2013, pp. 79–90.

[27] B. Mager, P. Lundrigan, and N. Patwari, "Fingerprint-based device-free localization performance in changing environments," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 11, pp. 2429–2438, Nov. 2015.

[28] D. Maas, J. Wilson, and N. Patwari, "Toward a rapidly deployable radio tomographic imaging system for tactical operations," in *Proc. 38th Annu. IEEE Conf. Local Comput. Netw. Workshops*, 2013, pp. 203–210.

[29] S. Savazzi, M. Nicoli, F. Carminati, and M. Riva, "A Bayesian approach to device-free localization: Modeling and experimental assessment," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 1, pp. 16–29, Feb. 2014.

[30] X. Chen, A. Edelstein, Y. Li, M. Coates, M. Rabbat, and A. Men, "Sequential Monte Carlo for simultaneous passive device-free tracking and sensor localization using received signal strength measurements," in *Proc. 10th ACM/IEEE Int. Conf. Inf. Process. Sensor Netw.*, 2011, pp. 342–353.

[31] S. Nannuru, Y. Li, Y. Zeng, M. Coates, and B. Yang, "Radio-frequency tomography for passive indoor multitarget tracking," *IEEE Trans. Mobile Comput.*, vol. 12, no. 12, pp. 2322–2333, Dec. 2013.

[32] H. L. Van Trees, *Detection, Estimation and Modulation Theory*. New York, NY, USA: Wiley, 1968.

[33] P. Tichavsky, "Posterior Cramér-Rao bound for adaptive harmonic retrieval," *IEEE Trans. Signal Process.*, vol. 43, no. 5, pp. 1299–1302, May 1995.

[34] R. K. Martin, A. Folkerts, and T. Heinl, "Accuracy vs. resolution in radio tomography," *IEEE Trans. Signal Process.*, vol. 62, no. 10, pp. 2480–2491, May 2014.

[35] V. Rampa, S. Savazzi, M. Nicoli, and M. D'Amico, "Physical modeling and performance bounds for device-free localization systems," *IEEE Signal Process. Lett.*, vol. 22, no. 11, pp. 1864–1868, Nov. 2015.

[36] S. Kianoush, V. Rampa, S. Savazzi, and M. Nicoli, "Pre-deployment performance assessment of device-free radio localization systems," in *Proc. IEEE Int. Conf. Commun. Workshops*, 2016, pp. 1–6.

[37] W. G. Figel, N. H. Shepherd, and W. F. Trammell, "Vehicle location by a signal attenuation method," in *Proc. 19th IEEE Veh. Technol. Conf.*, 1968, vol. 19, pp. 105–109.

[38] A. Dogandzic and P. P. Amran, "Signal-strength based localization in wireless fading channels," in *Proc. Asilomar Conf. Signals Syst. Comput.*, 2004, vol. 2, pp. 2160–2164.

[39] G. Frenkel, "Geometric dilution of position (GDOP) in position determination through radio signals," *Proc. IEEE*, vol. 61, no. 4, pp. 496–497, Apr. 1973.

[40] E. G. Larsson, "Cramer-Rao bound analysis of distributed positioning in sensor networks," *IEEE Signal Process. Lett.*, vol. 11, no. 3, pp. 334–337, Mar. 2004.

[41] Y. Shen and M. Z. Win, "Fundamental limits of wideband localization–Part I: A general framework," *IEEE Trans. Inf. Theory*, vol. 56, no. 10, pp. 4956–4980, Oct. 2010.

[42] N. Patwari, J. N. Ash, S. Kyperountas, A. O. Hero, R. L. Moses, and N. S. Correal, "Locating the nodes: Cooperative localization in wireless sensor networks," *IEEE Signal Process. Mag.*, vol. 22, no. 4, pp. 54–69, Jul. 2005.

[43] Y. Bar-Shalom and X.-R. Li, *Estimation with Applications to Tracking and Navigation*. Hoboken, NJ, USA: Wiley, 2001.

[44] H. Yiğitler and R. Jäntti, "Experimental accuracy assessment of radio tomographic imaging methods," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. Workshops*, 2016, pp. 1–6.

[45] P. Tichavsky, C. H. Muravchik, and A. Nehorai, "Posterior Cramér-Rao bounds for discrete-time nonlinear filtering," *IEEE Trans. Signal Process.*, vol. 46, no. 5, pp. 1386–1396, May 1998.

[46] R. Walpole, R. Myers, S. Myers, and K. Ye, *Probability and Statistics for Engineers and Scientists (9th Edition)*. Upper Saddle River, NJ, USA: Prentice Hall, 2012.

[47] M. Bocca, O. Kaltiokallio, and N. Patwari, "Radio tomographic imaging for ambient assisted living," in *Proc. Int. Competition Evaluating AAL Syst. Through Competitive Benchmarking*, 2013, vol. 362, pp. 108–130.

[48] H. Yiğitler, R. Jäntti, O. Kaltiokallio, and N. Patwari, "Detector based radio tomographic imaging," *IEEE Trans. Mobile Comput.*, vol. 17, no. 1, pp. 58–71, Jan. 2018.

[49] J. Wilson and N. Patwari, "A fade-level Skew-Laplace signal strength model for device-free localization with wireless networks," *IEEE Trans. Mobile Comput.*, vol. 11, no. 6, pp. 947–958, Jun. 2012.
[50] IEEE 802.15.4 standard technical specs. 2003. [Online]. Available: http://www.ieee802.org/15/pub/TG4Expert.html
[51] R. H. Byrd, J. C. Gilbert, and J. Nocedal, "A trust region method based on interior point techniques for nonlinear programming," *Math. Program.*, vol. 89, no. 1, pp. 149–185, Nov. 2000.
[52] N. J. Gordon, D. J. Salmond, and A. F. M. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," *IEE Proc. F - Radar Signal Process.*, vol. 140, no. 2, pp. 107–113, Apr. 1993.
[53] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 174–188, Feb. 2002.

Ossi Kaltiokallio received the BSc and MSc degrees in electrical engineering from Aalto University, both in 2011, and the DSc (with distinction) degree in communications engineering from the School of Electrical Engineering, Aalto University, Finland, in 2017. He held a postdoctoral fellow position at the University of Utah, in 2018 and currently, he holds a postdoctoral researcher position at Aalto University. He has received best paper awards from SenseApp, in 2012, and IPSN, in 2014. His current research interests include RF propagation, localization and tracking technologies, and statistical signal processing.

Roland Hostettler (S'10-M'14) received the Dipl Ing degree in electrical and communication engineering from the Bern University of Applied Sciences, Switzerland, in 2007, the MSc degree in electrical engineering and the PhD degree in automatic control from the Luleå University of Technology, Sweden, in 2009 and 2014, respectively. He has held postdoctoral researcher positions at the Luleå University of Technology, Sweden and postdoctoral researcher and research fellow positions at Aalto University, Finland. Currently, he is an associate senior lecturer with the Department of Engineering Sciences, Uppsala University, Sweden. His research interests include statistical signal processing with applications to target tracking, biomedical engineering, and sensor networks. He is a member of the IEEE.

Neal Patwari received the BS and MS degree in EE from Virginia Tech, and the PhD degree in EE from the University of Michigan. He is a professor with the McKelvey School of Engineering, Washington University in St. Louis. He is appointed with the Department of Electrical and Systems Engineering and the Department of Computer Science and Engineering. He was with the University of Utah in electrical and computer engineering from 2006 to 2018. He directs the Sensing and Processing Across Networks (SPAN) Lab, which performs research with the intersection of statistical signal processing and wireless networking, for improving wireless sensor networking and for RF sensing, in which the radio interface is the sensor. His research perspective was shaped by the research he conducted during his BS, MS, and PhD degrees. He received the NSF CAREER Award, in 2008, the 2009 IEEE Signal Processing Society Best Magazine Paper Award, and the 2011 University of Utah Early Career Teaching Award. He has coauthored papers with best paper awards at IEEE SenseApp 2012 and ACM/IEEE IPSN 2014 conference. He has served on technical program committees for IPSN, MobiCom, SECON, IPIN, and SenSys.