Random Finite Set Approach to Signal Strength Based Passive Localization and Tracking

Ossi Kaltiokallio*, Hüseyin Yiğitler†, Jukka Talvitie*, and Mikko Valkama*
*Unit of Electrical Engineering, Tampere University, Tampere, Finland,
†Department of Communications and Networking, Aalto University, Espoo, Finland,
{ossi.kaltiokallio, jukka.talvitie, mikko.valkama}@tuni.fi, {huseyin.yigitler}@aalto.fi

***Abstract*—Radio frequency sensor networks can be utilized for locating and tracking people within coverage area of the network. The technology is based on the fact that humans alter properties of the wireless propagation channel which is observed in the channel estimates, enabling tracking without requiring people to carry any sensor, tag or device. Considerable efforts have been made to model the human induced perturbations to the channel and develop flexible models that adapt to the unique propagation environment to which the network is deployed in. This paper proposes a noteworthy conceptual shift in the design of passive localization and tracking systems as the focus is shifted from channel modeling to filter design. We approach the problem using random finite set theory enabling us to model detections, missed detections, false alarms and unknown data association in a rigorous manner. The Bayesian filtering recursion applied with random finite sets is presented and a computationally tractable Gaussian sum filter is developed. The development efforts of the paper are validated using experimental data and the results imply that the proposed approach can decrease the tracking error up to 48% with respect to a benchmark solution.**

***Index Terms*—Received signal strength, RF sensor network, localization and tracking, random finite set, Gaussian sum filter**

I. Introduction

Ubiquitous radio frequency (RF) sensing technologies have experienced a surge of interest over the past decade and are considered as a potential candidate to be used in smart homes. We envision that future smart homes would not only regulate the heating, lighting and ventilation, but also monitor its inhabitants. In this regard, RF sensing is particularly suitable for realizing such system capabilities and over the past years, various radio technologies have been demonstrated for vital sign monitoring [1], activity and gesture recognition [2], and localization and tracking [3]. The abundant information that can be extracted with RF sensing can be exploited in various ways including: 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.

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 channel estimates. Research has demonstrated the use of various channel estimates for inference, including time delay [1], phase [2], and signal strength [3]; 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) channel estimates that are ubiquitously available in nearly all receivers. In this paper, we consider narrowband wireless devices that provide the RSS and we utilize the RSS for passive localization and tracking (PLT) in indoor environments.

The performance of PLT systems is significantly influenced by the model that describes the RSS as a function of a person's location. Typically, elaborate modeling [4]–[6], supervised training [7]–[9] or unsupervised parameter estimation [10]–[12] are used to tune the model since the model has an immediate impact on localization and tracking accuracy. As an example, PLT can be implemented using a two-step approach in which radio tomographic imaging (RTI) [13] is first used to reconstruct an image description of the human-induced propagation changes within the monitored area and thereafter, a position detector is used to localize the person from the image and a Kalman filter (KF) to track the person as time evolves [14]. With an ideal model, the person can always be detected from the image and the images do not contain any anomalies that would generate false detections. In practice however, the RSS model will always have inaccuracies and localizing the person accurately from the reconstructed images at all times is difficult. Since the detections are used as input to the KF, inaccurate localization also has a direct impact on tracking performance.

This paper proposes a noteworthy conceptual shift in the design of PLT systems as the focus is shifted from RSS modeling to filter design. First of all, we utilize a simplistic RSS model together with RTI to reconstruct the images. To cope with modeling inaccuracies and noisy RTI images, a new position detector is developed that detects multiple candidate positions corresponding to peaks of the RTI image. Similar approaches have been proposed for multi-target localization but instead of processing the data during the detection phase to come up with a correct number of detections [15], [16], we handle all candidate positions in the tracking filter. The candidate positions are simply referred to as measurements from now on.

In this paper, the measurements are modeled using a random finite set (RFS) based approach since the number of measure-

This work was partially supported by the Academy of Finland under grants #328214, #323244, #319994, and #338224.

ments and their locations are unknown. An element of the RFS has three alternatives: (i) a detection if the measurement is generated by the person; (ii) clutter if the measurement is a false detection; (iii) a misdetection if the person does not generate a measurement and the element is empty. In other words, a detection is a peak in the RTI image that is nearby the actual location of the person, clutter measurements correspond to peaks of the RTI image that are far away from the person, and a misdetection occurs if the RTI image does not have a peak near the person.

To solve the single-object tracking (SOT) in clutter problem we resort to RFS theory [17] which provides a solid mathematical foundation to handle the challenges introduced by the proposed position detector. A measurement model for PLT is presented that incorporates detections, misdetections, clutter, and data association (DA) uncertainty into the measurement likelihood function. Thereafter, Bayesian filtering equations applied with RFSs are presented and a computationally tractable Gaussian sum filter (GSF) is developed to track the person using cluttered measurements. The considered problem is closely related to multi-object tracking (MOT) for which the RFS framework is widely used to model the problem in a Bayesian way [17], [18].

The development efforts of the paper are demonstrated using commodity narrowband wireless transceivers that operate on the 2.4 GHz ISM band. The experiments are conducted in a downtown residential apartment and in an open indoor environment. Using the experimental data, it is shown that the proposed approach can decrease the tracking error up to 48% with respect to a benchmark solution. Thus, the experimental results validate the research premise that the presented approach can deal with larger RSS modeling errors due to the filters ability to handle clutter measurements, missing measurements and DA uncertainty. This paper makes the following contributions:

1) A new position detector is developed that detects multiple candidate positions corresponding to peaks of the RTI image.
2) The detected peaks, referred to as measurements, are modeled using a RFS and a measurement likelihood function is presented that account for detections, misdetections, clutter, and DA uncertainty.
3) A computationally tractable GSF is developed and it is demonstrated that the filter outperforms the KF in a variety of experiments.

The rest of the paper is organized as follows. Section II introduces the widely utilized models for PLT and summarizes the typical RTI+KF approach for tracking. The RFS approach to PLT is presented in Section III. Section IV introduces the conducted experiments and the results are presented in Section V. Thereafter, conclusions are drawn in Section VI.

II. Passive Localization and Tracking

Imaging-based PLT is a two-step process. In the first step, RTI is used to reconstruct an image description of the spatial field describing the RSS changes of the monitored area [13]. In the second step, a position detector is used to localize the person from the RTI image and the detections are utilized as input to a tracking filter [19].

A. Received Signal Strength Model

Let us consider a geographical area indexed by the closed and convex set $\mathcal{A} \subset \mathbb{R}^2$ and a wireless link ℓ between a RF transmitter (TX) located at $\mathbf{p}_{\ell,\text{TX}} \in \mathcal{A}$ and an RF receiver (RX) located at $\mathbf{p}_{\ell,\text{RX}} \in \mathcal{A}$. Let us also consider a single person located within the monitored area at $\mathbf{p}(t) \in \mathcal{A}$. Assuming zero-mean Gaussian noise, the human-induced RSS changes of wireless link ℓ at time t follows

$$p(\tilde{y}_\ell(t) \mid \Delta_{\ell,\mathbf{p}(t)}) = \mathcal{N}(\tilde{y}_\ell(t); h(\Delta_{\ell,\mathbf{p}(t)}, \boldsymbol{\phi}_\ell), \sigma^2_\ell(t)) \quad (1)$$

where $\Delta_{\ell,\mathbf{p}(t)}$ is the excess path length, and the mean and variance of the Gaussian model are given by $h(\Delta_{\ell,\mathbf{p}(t)}, \boldsymbol{\phi}_\ell)$ and $\sigma^2_\ell(t)$, respectively. The excess path length is defined as

$$\Delta_{\ell,\mathbf{p}(t)} \triangleq d_{\ell,\text{TX}} + d_{\ell,\text{RX}} - d_\ell \quad (2)$$

in which $d_{\ell,\text{TX}} = \|\mathbf{p}_{\ell,\text{TX}} - \mathbf{p}(t)\|$ and $d_{\ell,\text{RX}} = \|\mathbf{p}_{\ell,\text{RX}} - \mathbf{p}(t)\|$ are the distances from $\mathbf{p}(t)$ to the TX and RX for link ℓ, and $d_\ell = \|\mathbf{p}_{\text{TX}} - \mathbf{p}_{\text{RX}}\|$ is the distance between the transceivers. The exponential model is given by [20]

$$h(\Delta_{\ell,\mathbf{p}(t)}, \boldsymbol{\phi}_\ell) = \mu_\ell + \kappa_\ell \exp\left(-\Delta_{\ell,\mathbf{p}(t)}/\gamma_\ell\right), \quad (3)$$

where $\boldsymbol{\phi}_\ell = [\mu_\ell, \kappa_\ell, \gamma_\ell]$ are the model parameters in which μ_ℓ is the baseline RSS, κ_ℓ the measurement gain and γ_ℓ the decay rate.

B. Image Reconstruction

If N is the total number of transceivers in the RF sensor network, then the total number of wireless links is $L = N \times (N-1)$. Now, considering all L links simultaneously, RTI reconstructs a discretized image description of the RSS changes within the considered geographical area. The RSS for every link is acquired over time window, $\mathcal{T} = \{t \mid (k-1)\,\tau < t \leq k\,\tau\}$, in which k denotes the sample number and τ duration of the time window. Now, the considered problem can be described by the following linear model [13]

$$\mathbf{y}_k = \boldsymbol{\Omega}\mathbf{b}_k + \mathbf{r}_k, \quad (4)$$

where $\mathbf{y}_k \in \mathbb{R}^L$ is the mean-removed RSS at time k, $\mathbf{b}_k \in \mathbb{R}^M$ is the discretized image to be estimated, $\boldsymbol{\Omega} \in \mathbb{R}^{L\times M}$ is an $L \times M$ time-invariant weighting matrix and $\mathbf{r}_k \in \mathbb{R}^L$ is a zero-mean Gaussian noise vector. The linear model in (4) is related to (1) via $\{\mathbf{y}_k\}_\ell = \tilde{y}_\ell(t) - \mu_\ell$ and $\mathbf{r}_k \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_\mathbf{r})$ in which $\boldsymbol{\Sigma}_\mathbf{r} = \text{diag}(\sigma^2_1, \ldots, \sigma^2_L)$. Furthermore, for link ℓ and pixel m located at $\mathbf{p}_m \in \mathcal{A}$, elements of the weighting matrix are given by [12]

$$\{\boldsymbol{\Omega}\}_{\ell,m} = \bar{\kappa}_\ell \exp\left(-\Delta_{\ell,\mathbf{p}_m}/\gamma_\ell\right), \quad (5)$$

where $\bar{\kappa}_\ell = \kappa_\ell/|\kappa_\ell|$ normalizes the measurement gain to one so that the links are weighted equally.

With zero-mean Gaussian image prior $\mathbf{b} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_\mathbf{b})$, the weighted least squares estimate for the model in (4) is [19]

$$\hat{\mathbf{b}}_k = \boldsymbol{\Pi}\mathbf{y}_k, \quad (6)$$

where

$$\mathbf{\Pi} = \left(\mathbf{\Omega}^\top \mathbf{\Sigma}_\mathbf{r}^{-1} \mathbf{\Omega} + \mathbf{\Sigma}_\mathbf{b}^{-1}\right)^{-1} \mathbf{\Omega}^\top \mathbf{\Sigma}_\mathbf{r}^{-1}. \tag{7}$$

Covariance of the image prior between pixels i and j is defined as [21]

$$\{\mathbf{\Sigma}\}_{i,j} = \sigma_b^2 \exp\left(-\|\mathbf{p}_i - \mathbf{p}_j\|/\delta_d\right), \tag{8}$$

where σ_b^2 is the pixel variance and δ_d the correlation distance.

C. Position Detector and Tracking Models

In the second step of imaging-based PLT, a position detector is used to localize the person from the image and the obtained position estimates are used as input to a tracking filter. The location of the person at time k can be estimated by finding the voxel of the RTI image that has the maximum value [14], [19]

$$\mathbf{z}_k = \mathbf{p}_j, \quad \text{where} \quad j = \arg\max_M \hat{\mathbf{b}}_k \tag{9}$$

and $\mathbf{p}_j$ is the position of voxel j. From now on, we refer to the position estimate simply as measurement or position measurement, since it is used as a measurement input to a tracking filter.

In this paper, the RF network is the sensor and it is used for determining position and velocity of a person over time. The dynamics of the target and measurements are most naturally combined using an optimal filter and in the following, the underlying models utilized by the filter are introduced. The state of the system in the 2D Euclidean space can be defined as

$$\mathbf{x}_k = \begin{bmatrix} x_k & \dot{x}_k & y_k & \dot{y}_k \end{bmatrix}^\top, \tag{10}$$

where x_k and y_k denote the coordinates and $\dot{x}_k$ and $\dot{y}_k$ denote the velocity components. The evolution of the state can be represented using a Gaussian transition density

$$p(\mathbf{x}_k \mid \mathbf{x}_{k-1}) = \mathcal{N}(\mathbf{x}_k \mid \mathbf{F}_{k-1}\mathbf{x}_{k-1}, \mathbf{Q}_{k-1}), \tag{11}$$

where $\mathbf{F}_{k-1}$ is the state transition matrix and $\mathbf{Q}_{k-1}$ covariance of the process noise. A common choice for the transition model in PLT is the second-order kinematic model [5], [19], given by [22]

$$\mathbf{F}_k = \mathbf{I} \otimes \begin{bmatrix} 1 & \tau_k \\ 0 & 1 \end{bmatrix}, \quad \mathbf{Q}_k = \mathbf{I} \otimes q \begin{bmatrix} \frac{1}{3}\tau_k^3 & \frac{1}{2}\tau_k^2 \\ \frac{1}{2}\tau_k^2 & \tau_k \end{bmatrix}, \tag{12}$$

where $\mathbf{I}$ is a 2×2 identity matrix, $\otimes$ denotes the Kronecker product, q is the power spectral density of the process noise and τ_k is the sampling interval. In the considered problem, $\mathbf{F}_k$, $\mathbf{Q}_k$ and τ_k are time-variant since the sampling interval contains jitter.

The connection between the measurements and the state is represented using a Gaussian measurement likelihood

$$p(\mathbf{z}_k \mid \mathbf{x}_k) = \mathcal{N}(\mathbf{z}_k \mid \mathbf{H}\mathbf{x}_k, \mathbf{R}), \tag{13}$$

where $\mathbf{H}$ is a linear measurement matrix and $\mathbf{R}$ is the measurement noise covariance, given by

$$\mathbf{H} = \mathbf{I} \otimes \begin{bmatrix} 1 & 0 \end{bmatrix} \quad \text{and} \quad \mathbf{R} = \sigma_r^2 \mathbf{I} \tag{14}$$

in which σ_r^2 is variance of the measurement noise.

D. Bayesian Filtering

Recursive Bayesian filtering computes the marginal posterior distribution of the state $\mathbf{x}_k$ at time step k, given the sequence of measurements $\mathbf{z}_{1:k} = \{\mathbf{z}_1, \ldots, \mathbf{z}_k\}$ up to time instant k

$$p(\mathbf{x}_k \mid \mathbf{z}_{1:k}). \tag{15}$$

The recursive equations to compute the marginal posterior distribution is given by the following Bayesian filtering equations. The recursion starts from the prior distribution $p(\mathbf{x}_{k-1} \mid \mathbf{z}_{1:k-1}) = p(\mathbf{x}_0)$ at $k = 0$. The prediction step of the Bayesian filtering recursion can be computed using the Chapman-Kolmogorov equation [23]

$$p(\mathbf{x}_k \mid \mathbf{z}_{1:k-1}) = \int p(\mathbf{x}_k \mid \mathbf{x}_{k-1}) p(\mathbf{x}_{k-1} \mid \mathbf{z}_{1:k-1}) d\mathbf{x}_{k-1}, \tag{16}$$

in which $p(\mathbf{x}_k \mid \mathbf{x}_{k-1})$ is the transition density. Once measurement $\mathbf{z}_k$ is available at time k, the predictive distribution can be updated using the Bayes' rule [23]

$$p(\mathbf{x}_k \mid \mathbf{z}_{1:k}) = \frac{p(\mathbf{z}_k \mid \mathbf{x}_k) p(\mathbf{x}_k \mid \mathbf{z}_{1:k-1})}{p(\mathbf{z}_k \mid \mathbf{z}_{1:k-1})}, \tag{17}$$

where $p(\mathbf{z}_k \mid \mathbf{x}_k)$ is the likelihood function and the term in the denominator is the normalization constant $p(\mathbf{z}_k \mid \mathbf{z}_{1:k-1}) = \int p(\mathbf{z}_k \mid \mathbf{x}_k) p(\mathbf{x}_k \mid \mathbf{z}_{1:k-1}) d\mathbf{x}_k$.

1) Kalman Filter: Typically in PLT [12], [14], [19], the posterior is represented using a Gaussian

$$p(\mathbf{x}_k \mid \mathbf{z}_{1:k}) = \mathcal{N}(\mathbf{x}_k \mid \mathbf{m}_k, \mathbf{P}_k), \tag{18}$$

which is parameterized by the mean $\mathbf{m}_k$ and covariance $\mathbf{P}_k$. Since $p(\mathbf{x}_k \mid \mathbf{z}_{1:k})$ is Gaussian, and the models in (11) and (13) are linear Gaussian, the KF is the closed form solution to the Bayesian filtering equations. The distributions in (16) and (17) are Gaussian and given by [23]

$$p(\mathbf{x}_k \mid \mathbf{z}_{1:k-1}) = \mathcal{N}(\mathbf{x}_k \mid \mathbf{m}_{k|k-1}, \mathbf{P}_{k|k-1}), \tag{19}$$

$$p(\mathbf{x}_k \mid \mathbf{z}_{1:k}) = \mathcal{N}(\mathbf{x}_k \mid \mathbf{m}_{k|k}, \mathbf{P}_{k|k}), \tag{20}$$

$$p(\mathbf{z}_k \mid \mathbf{z}_{1:k-1}) = \mathcal{N}(\mathbf{x}_k \mid \mathbf{H}\mathbf{m}_{k|k-1}, \mathbf{S}_{k|k-1}). \tag{21}$$

The parameters of the predictive distribution in (19) are given by prediction step of the KF

$$\mathbf{m}_{k|k-1} = \mathbf{F}_{k-1}\mathbf{m}_{k-1|k-1}, \tag{22}$$

$$\mathbf{P}_{k|k-1} = \mathbf{F}_{k-1}\mathbf{P}_{k-1|k-1}\mathbf{F}_{k-1}^\top + \mathbf{Q}_{k-1}. \tag{23}$$

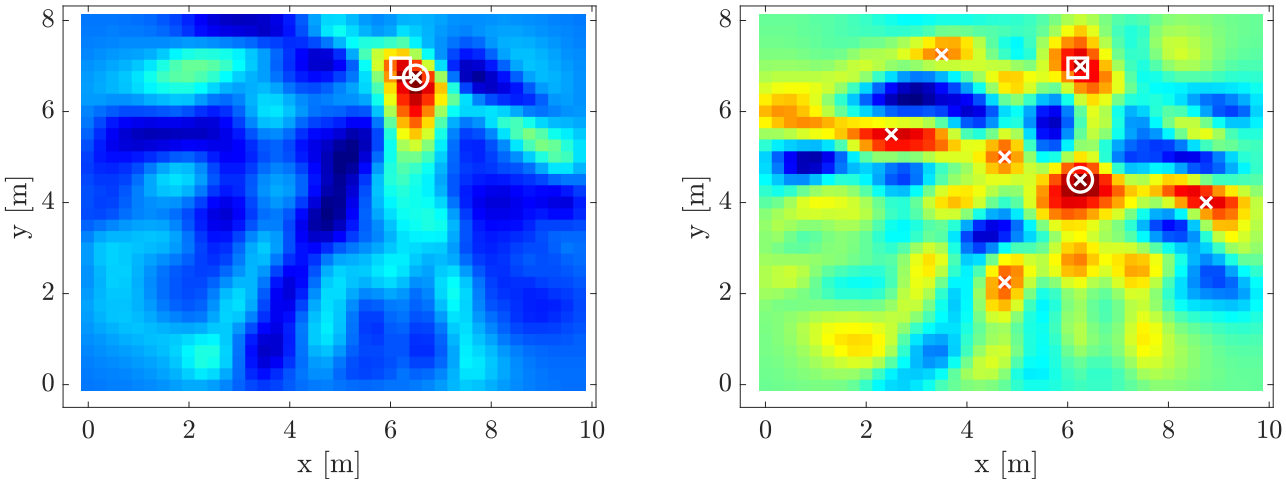
Respectively, parameters of the posterior distribution in (20) are given by update step of the KF

$$\mathbf{S}_{k|k-1} = \mathbf{H}\mathbf{P}_{k|k-1}\mathbf{H}^\top + \mathbf{R}, \tag{24}$$

$$\mathbf{K}_{k|k-1} = \mathbf{P}_{k|k-1}\mathbf{H}^\top \mathbf{S}_{k|k-1}^{-1}, \tag{25}$$

$$\mathbf{m}_{k|k} = \mathbf{m}_{k|k-1} + \mathbf{K}_{k|k-1}\left(\mathbf{z}_k - \mathbf{H}\mathbf{m}_{k|k-1}\right), \tag{26}$$

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_{k|k-1}\mathbf{S}_{k|k-1}\mathbf{K}_{k|k-1}^\top. \tag{27}$$



(a) Low clutter image

(b) High clutter image

Fig. 1. Example RTI images in a low clutter scenario (a) and in a high clutter scenario (b). In the images, the true position of the person is illustrated with the white square, estimate computed using (9) presented with white circle and estimates obtained using (31) shown with white crosses.

III. A Random Finite Set Approach to PLT

An RTI image representing the situation in which one person is located in the monitored area should ideally show a single peak in the image and the pixel with highest intensity should locate near the object as illustrated in Fig. 1a. However, due to noise in the RSS and the complex nature of the indoor radio propagation channel, the RTI images are often noisy, contain spurious peaks that do not correspond to the actual object and the peak with the highest intensity is not always the one that is closest to the person as illustrated in Fig. 1b. For this reason, a peak detector is developed that finds all the local maxima in the image and a Bayesian filter based on RFS theory is implemented to track the person using cluttered measurements.

A. Peak Detector and Measurement Model

Ideally, each peak in the RTI image is a smooth point spread function (PSF) resembling a 2D Gaussian but in reality, the images can contain additive noise such that each PSF peak can have multiple small peaks. To suppress the noise and small peaks, the images are filtered so that there will be only one pixel in each peak that will correspond to the actual maxima of the PSF. As in [15], the images are smoothed using a 2D Gaussian filter

$$\mathbf{B}(x,y) = \frac{\exp(-\frac{1}{2\sigma_B^2}(x^2+y^2))}{\sum_x \sum_y \exp(-\frac{1}{2\sigma_B^2}(x^2+y^2))}, \tag{28}$$

where $\sigma_B^2 = 0.25$ is the variance and elements of the utilized 3×3 filter are

$$\mathbf{B}(x,y) \approx \begin{bmatrix} 0.0000 & 0.0003 & 0.0000 \\ 0.0003 & 0.9987 & 0.0003 \\ 0.0000 & 0.0003 & 0.0000 \end{bmatrix}. \tag{29}$$

The low-pass filtered RTI image is then calculated as

$$\tilde{\mathbf{b}}_k = \hat{\mathbf{b}}_k * \mathbf{B}(x,y), \tag{30}$$

where $*$ represents the 2D convolution operator and it is assumed that the estimated image vector, $\hat{\mathbf{b}}_k \in \mathbb{R}^M$, has been converted to a 2D image above.

The locations of all the peaks at time k are modeled using a RFS with random cardinality and random elements, given by

$$\mathcal{Z}_k = \{\mathbf{z}_k^1, \ldots, \mathbf{z}_k^{m_k}\}, \tag{31}$$

where m_k is the total number of valid peaks and $\mathbf{z}_k^j = \mathbf{p}_i$ is the position of voxel i. Pixel $\mathbf{p}_i$ contains a peak if $\tilde{\mathbf{b}}_{i,k}$ is larger than the eight surrounding pixels. Furthermore, we only consider pixels with intensity higher than $\alpha\mathcal{B}$, where $\mathcal{B} = \max(\tilde{\mathbf{b}}_k)$ denotes the maximum component of $\tilde{\mathbf{b}}_k$ and $\alpha = 0.75$ is a threshold. The threshold is a tuning parameter between two extremes: if $\alpha = 0$ all peaks are taken into account and if $\alpha = 1$ only a single peak is accounted for, and we have empirically found that $\alpha = 0.75$ provides a good overall performance.

Since the peak detector in (31) gives rise to multiple measurements, the problem can be formulated as SOT in clutter [24], [25], which is a special case of MOT [17], [18]. With respect to conventional Bayesian filtering there are new challenges that must be considered and these are: missed detections, clutter detections and unknown DAs. The main difference with respect to the problem introduced in the previous section is the underlying measurement model and in the following, a likelihood function that account for the new challenges is introduced.

The probability that an object is detected by the sensor is represented by detection probability $P_{\mathrm{D}}(\mathbf{x}_k) \in [0,1]$ and respectively, $1 - P_{\mathrm{D}}(\mathbf{x}_k)$ is the probability of misdetection. The detection process is Bernoulli distributed with probability $P_{\mathrm{D}}(\mathbf{x}_k)$ and if the object is detected, the measurements follow the likelihood function in (13). The clutter measurements form a m_k^c-subset of $\mathcal{Z}_k$, denoted as $\mathcal{C}_k \subseteq \mathcal{Z}_k$, where $\mathcal{C}_k = \{\mathbf{c}_k^1, \ldots \mathbf{c}_k^{m_k^c}\}$. The subset is proper if the object is

detected ($m_k = m_k^c + 1$) and the sets are equal if the object is undetected ($m_k = m_k^c$). The clutter measurements are modeled using a Poisson point process (PPP) [24]

$$p(\mathcal{C}_k) = \frac{\exp(-\bar{\lambda}_c)}{m_k^c!} \prod_{i=1}^{m_k^c} \lambda_c(\mathbf{c}_k^i), \tag{32}$$

where $m_k^c \sim \text{Po}(\bar{\lambda}_c)$ is the number of clutter measurements and it follows a Poisson distribution with rate parameter $\bar{\lambda}_c \in [0, \infty)$. The clutter is assumed i.i.d. $\mathbf{c}_k^i \sim p_c(\mathbf{c})$, where $p_c(\mathbf{c})$ describes the spatial probability distribution function (PDF) and the clutter is parameterized using intensity function $\lambda_c(\mathbf{c}_k^i) = \bar{\lambda}_c p_c(\mathbf{c})$. To describe the DA we use $\theta_k = 0$, if the object is undetected and $\theta_k = i$, if $\mathbf{z}_k^i$ is an object detection. The complete measurement likelihood can now be expressed as [24], [25]:

$$\begin{aligned} p(\mathcal{Z}_k \mid \mathbf{x}_k) &= \sum_{\theta_k=0}^{m_k} p(\mathcal{Z}_k, m_k, \theta_k \mid \mathbf{x}_k) \\ &= \sum_{\theta_k=0}^{m_k} p(\mathcal{Z}_k \mid m_k, \theta_k, \mathbf{x}_k) p(m_k, \theta_k \mid \mathbf{x}_k) \\ &= (1 - P_{\text{D}}(\mathbf{x}_k)) \frac{\exp(-\bar{\lambda}_c)}{m_k!} \prod_{i=1}^{m_k} \lambda_c(\mathbf{z}_k^i) \\ &+ \sum_{\theta_k=1}^{m_k} P_{\text{D}}(\mathbf{x}_k) \frac{p(\mathbf{z}_k^{\theta_k} \mid \mathbf{x}_k)}{\lambda_c(\mathbf{z}_k^{\theta_k})} \frac{\exp(-\bar{\lambda}_c)}{m_k!} \prod_{i=1}^{m_k} \lambda_c(\mathbf{z}_k^i). \end{aligned} \tag{33}$$

In the above equation, the sum over $\theta_k = \{0, \ldots, m_k\}$ accounts for all possible DAs, terms on the third line account for misdetection and clutter, and the terms on the last line account for detection, the distribution of the detection and clutter.

B. Filtering Recursion

Let the sequence of measurements and DA hypotheses up to time k be denoted as $\mathcal{Z}_{1:k} = \{\mathcal{Z}_1, \ldots, \mathcal{Z}_k\}$ and $\theta_{1:k} = \{\theta_1, \ldots, \theta_k\}$, respectively. In SOT, the posterior distribution at time k can be expressed as the summation over all possible data association sequences up to time k [25]

$$p(\mathbf{x}_k \mid \mathcal{Z}_{1:k}) = \sum_{\theta_{1:k}} w_{\theta_{1:k}} p(\mathbf{x}_k \mid \theta_{1:k}, \mathcal{Z}_{1:k}), \tag{34}$$

where $p(\mathbf{x}_k \mid \theta_{1:k}, \mathcal{Z}_{1:k})$ is the density of the object state conditioned on the measurements and a specific DA hypotheses sequence, $w_{\theta_{1:k}} = \Pr(\theta_{1:k} \mid \mathcal{Z}_{1:k})$ denotes the probability of the DA hypotheses sequence and

$$\sum_{\theta_{1:k}} = \sum_{\theta_1=0}^{m_1} \sum_{\theta_2=0}^{m_2} \cdots \sum_{\theta_k=0}^{m_k}. \tag{35}$$

The filtering recursion of SOT in clutter follows the conventional prediction and update steps of Bayesian filtering presented in Section II-D, but now the posterior is a mixture density and the likelihood function is different. Let us denote the posterior at the previous time step as $p(\mathbf{x}_{k-1} \mid \mathcal{Z}_{1:k-1})$, the prediction step of the filter can be computed using [25]

$$\begin{aligned} p(\mathbf{x}_k \mid \mathcal{Z}_{1:k-1}) &= \int p(\mathbf{x}_k \mid \mathbf{x}_{k-1}) p(\mathbf{x}_{k-1} \mid \mathcal{Z}_{1:k-1}) d\mathbf{x}_{k-1}, \\ &= \sum_{\theta_{1:k-1}} w_{\theta_{1:k-1}} p(\mathbf{x}_k \mid \theta_{1:k-1}, \mathcal{Z}_{1:k-1}), \end{aligned} \tag{36}$$

where predicted density of the object state is computed using the Chapman-Kolmogorov equation

$$\begin{aligned} p(\mathbf{x}_k \mid \theta_{1:k-1}, \mathcal{Z}_{1:k-1}) = \int p(\mathbf{x}_k \mid \mathbf{x}_{k-1}) \\ \times p(\mathbf{x}_{k-1} \mid \theta_{1:k-1}, \mathcal{Z}_{1:k-1}) d\mathbf{x}_{k-1}. \end{aligned} \tag{37}$$

During the prediction step, the weights remain the same and the density of every hypotheses is predicted using the standard Bayesian filtering prediction step. The update step of the filter is computed using the Bayes' rule and once measurement set $\mathcal{Z}_k$ is available at time k, the posterior is updated using [25]

$$\begin{aligned} p(\mathbf{x}_k \mid \mathcal{Z}_{1:k}) &\propto \sum_{\theta_{1:k-1}} w_{\theta_{1:k-1}} p_{k|k-1}^{\theta_{1:k-1}}(\mathbf{x}_k)(1 - P_{\text{D}}(\mathbf{x}_k)) \\ &+ \sum_{\theta_{1:k-1}} \sum_{\theta_k=1}^{m_k} \frac{P_{\text{D}}(\mathbf{x}_k)}{\lambda_c(\mathbf{z}_k^{\theta_k})} w_{\theta_{1:k-1}} p_{k|k-1}^{\theta_{1:k-1}}(\mathbf{x}_k) p(\mathbf{z}_k^{\theta_k} \mid \mathbf{x}_k), \end{aligned} \tag{38}$$

where the predicted density is denoted as $p_{k|k-1}^{\theta_{1:k-1}}(\mathbf{x}_k) = p(\mathbf{x}_k \mid \theta_{1:k-1}, \mathcal{Z}_{1:k-1})$ for brevity. It is to be noted that every state density is updated by a misdetection and by every measurement, such that the number of hypotheses grows by a factor of $m_k + 1$ when the posterior is updated.

C. Gaussian Sum Filter

The filtering recursion presented in Section III-B is intractable since the number of hypotheses grows according to $\prod_{i=1}^{k}(m_i+1)$. To obtain a computationally feasible algorithm, the posterior in (34) is approximated using a Gaussian mixture model (GMM) with fewer components similar to the works in [26] and [27]. The GMM is given by

$$p(\mathbf{x}_k \mid \mathcal{Z}_{1:k}) \approx \sum_{h_k=1}^{\mathcal{H}_k} w_k^{h_k} \mathcal{N}(\mathbf{x}_k \mid \mathbf{m}_k^{h_k}, \mathbf{P}_k^{h_k}), \tag{39}$$

where $\mathcal{H}_k$ is the number of hypotheses at time k and h_k the hypotheses index. For linear Gaussian models given in Section II-C and assuming constant detection probability P_{D}, the distributions in (36) and (38) are Gaussian mixtures and given by [25]

$$p(\mathbf{x}_k \mid \mathcal{Z}_{1:k-1}) = \sum_{h_{k-1}=1}^{\mathcal{H}_{k-1}} w_{k|k-1}^{h_{k-1}} \mathcal{N}(\mathbf{x}_k \mid \mathbf{m}_{k|k-1}^{h_{k-1}}, \mathbf{P}_{k|k-1}^{h_{k-1}}), \tag{40}$$

$$p(\mathbf{x}_k \mid \mathcal{Z}_{1:k}) = \sum_{h_k=1}^{\mathcal{H}_k} w_{k|k}^{h_k} \mathcal{N}(\mathbf{x}_k \mid \mathbf{m}_{k|k}^{h_k}, \mathbf{P}_{k|k}^{h_k}), \tag{41}$$

where $\mathcal{H}_k = \mathcal{H}_{k-1} \times (m_k + 1)$ and the above equations form the filtering recursion of the proposed GSF. It is to be noted that the presented GSF recursion is similar to the works in [26]

and [27], with the difference that the weights are calculated differently due to the different likelihood functions.

Parameters of the predictive distribution in (40) are

$$w_{k|k-1}^{h_{k-1}} = w_{k-1|k-1}^{h_{k-1}}, \tag{42}$$

$$\mathbf{m}_{k|k-1}^{h_{k-1}} = \mathbf{F}_{k-1}\mathbf{m}_{k-1|k-1}^{h_{k-1}}, \tag{43}$$

$$\mathbf{P}_{k|k-1}^{h_{k-1}} = \mathbf{F}_{k-1}\mathbf{P}_{k-1|k-1}^{h_{k-1}}\mathbf{F}_{k-1}^{\top} + \mathbf{Q}_{k-1}. \tag{44}$$

Parameters of the posterior distribution in (41) for misdetection, $\theta_k = 0$, are $w_{k|k}^{h_k} \propto (1 - P_{\mathrm{D}})w_{k|k-1}^{h_{k-1}}$, $\mathbf{m}_{k|k}^{h_k} = \mathbf{m}_{k|k-1}^{h_{k-1}}$ and $\mathbf{P}_{k|k}^{h_k} = \mathbf{P}_{k|k-1}^{h_{k-1}}$. Respectively, parameters of the updated posterior for the detected components, $\theta_k \in \{1, \dots m_k\}$, are

$$w_{k|k}^{h_k} \propto \frac{P_{\mathrm{D}}}{\lambda_c(\mathbf{z}_k^{\theta_k})} w_{k|k-1}^{h_{k-1}} \mathcal{N}(\mathbf{z}_k^{\theta_k}; \mathbf{H}\mathbf{m}_{k|k-1}^{h_{k-1}}, \mathbf{S}_{k|k-1}^{h_{k-1}}), \tag{45}$$

$$\mathbf{m}_{k|k}^{h_k} = \mathbf{m}_{k|k-1}^{h_{k-1}} + \mathbf{K}_{k|k-1}^{h_{k-1}}(\mathbf{z}_k^{\theta_k} - \mathbf{H}\mathbf{m}_{k|k-1}^{h_{k-1}}), \tag{46}$$

$$\mathbf{P}_{k|k}^{h_k} = \mathbf{P}_{k|k-1}^{h_{k-1}} - \mathbf{K}_{k|k-1}^{h_{k-1}}\mathbf{S}_{k|k-1}^{h_{k-1}}(\mathbf{K}_{k|k-1}^{h_{k-1}})^{\top}. \tag{47}$$

where $\mathbf{S}_{k|k-1}^{h_{k-1}}$ and $\mathbf{K}_{k|k-1}^{h_{k-1}}$ are computed using (24) and (25), respectively. During the update step, a new hypotheses h_k is obtained for every old hypotheses h_{k-1} and θ_k, and the new hypotheses is assigned the index $h_k = h_{k-1} + \mathcal{H}_{k-1}\theta_k$ to assure a unique mapping. After every update step, the weights are normalized, $w_{k|k}^{h_k} = w_{k|k}^{h_k} / \sum_{h_k=1}^{\mathcal{H}_k} w_{k|k}^{h_k}$, and the state estimate used for evaluation purposes is computed using the minimum mean square error (MMSE) estimate of the posterior mean, given by

$$\mathbf{m}_{k|k} = \sum_{h_k=1}^{\mathcal{H}_k} w_k^{h_k}\mathbf{m}_{k|k}^{h_k}. \tag{48}$$

To obtain a computationally feasible algorithm and constrain the exponential growth of Gaussian components, the hypotheses reduction algorithm presented in [18] is used to reduce the number of Gaussian components that are propagated to the next time step. The hypotheses reduction algorithm discards components with weights below a pruning threshold. Thereafter, it merges components that are close together and approximates them using a single Gaussian. Lastly, it truncates the Gaussian mixture and only keeps a certain number of components with highest weights. In the experiments, the following parameter values are used in the hypotheses reduction algorithm: pruning threshold is 10^{-6}, merging threshold is 5 and truncating threshold is $\mathcal{H}_{max} = 10$.

IV. Experiments

A. Experiment Description

The development efforts of this paper are demonstrated using narrowband wireless transceivers. The utilized sensors are Texas Instruments CC2531 USB dongle nodes [28] which operate on the 2.4 GHz ISM band according to the IEEE 802.15.4 standard [29]. The wireless nodes communicate using a round-robin schedule on four frequency channels $\mathcal{F} \in \{11, 16, 21, 26\}$. In the transmitted packets, the nodes include the most recent RSS, associated with the transmissions of other nodes. After transmission, the turn is assigned to the next node in the schedule. The time interval between two successive transmission is $\bar{\tau} \approx 2.9$ ms. A gateway that overhears all the traffic extracts the RSS from the packets and relays them to a computer through UART for centralized processing. The readers are referred to [30] for a more elaborate description of the communication protocol.

The experiments are carried out in a downtown residential apartment and in an open indoor environment. In both experiments, 20 nodes are deployed as illustrated in Fig. 2. The time window $\mathcal{T}$ during which each node transmits once is approximately $\tau_k = 20 \times \bar{\tau} \approx 58$ ms defining the sampling interval of the system. In the apartment, the nodes are deployed by the electric sockets so they could be powered from the mains. The size of the apartment is 82 m^2 and distance of the electric sockets from the ground varied from shin to belly height. In the open environment, the nodes are deployed around a 75 m^2 area and they are set on top of podiums at a height of 0.9 m.

In the experiments, reference positions were marked (see Fig. 2) and the test person was given directions to walk along the imaginary lines connecting the different reference positions. During the experiment, the person walked from one reference position to another after which they stopped and remained stationary for a few seconds. Thereafter, the person walked to the next reference positions. The person was carrying a video camera and in post-processing, the RSS and video streams were synchronized and the video was used to define the ground truth trajectory. The experiments were repeated three times in both environments and each trial was approximately three minutes long. In each trial, every reference position was visited at least once and the person followed no specific order when visiting the reference positions.

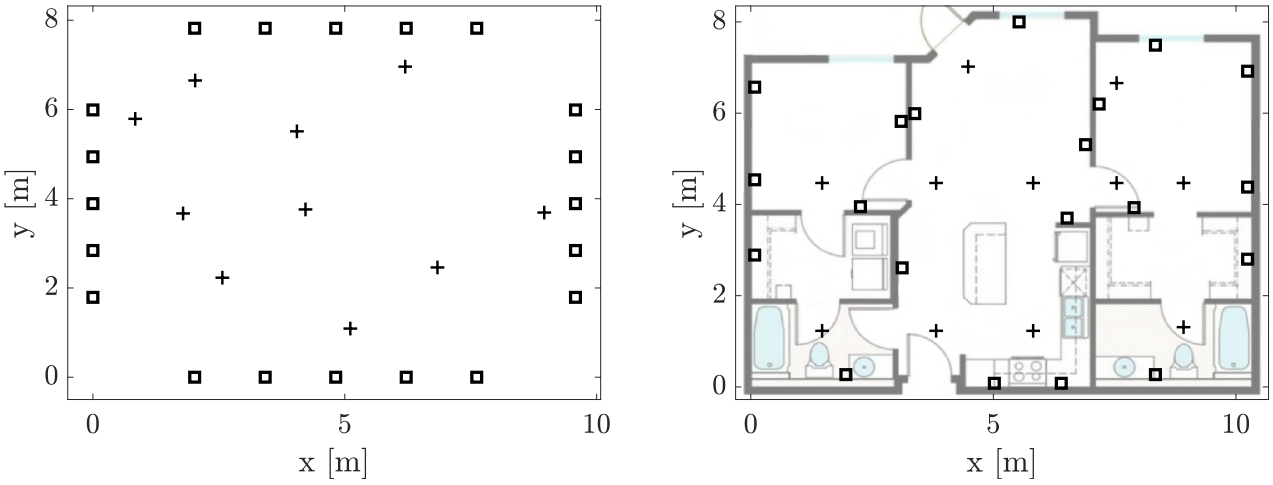
There were several co-existing Wi-Fi networks in both environments but the developed system can tolerate occasional packet drops and furthermore, frequency channel diversity partially mitigates interference issues. It is also to be noted that the developed system could be implement using any device capable of measuring the RSS including Wi-Fi, Bluetooth and RFID.

B. System Calibration

In the results section, we consider two different system calibration methods. In the first method, an empty-room calibration period is used to compute the reference RSS for every link [13]. The empty-room calibration period is approximately two minutes long and an independent calibration period precedes each trial. The reference RSS is computed using

$$\mu_\ell = \frac{1}{K}\sum_{k=1}^{K} \tilde{y}_{\ell,k}, \tag{49}$$

where K is the number of samples. The other model parameters of (3) are: $\kappa_l = -5$ dB, $\gamma_l = 0.04$ m and $\sigma_l^2 = 1$ dB^2.



(a) Open environment (b) Apartment

Fig. 2. The experimental layouts in which the nodes (□) and the reference positions (+) are illustrated.

In the second method, the model parameters are estimated using supervised training [7]. The parameter estimates are obtained by minimizing the cost function

$$J(\boldsymbol{\phi}_\ell) = \sum_{k=1}^{K} \left(\tilde{y}_{\ell,k} - h(\Delta_{\ell,\mathbf{p}(t)}, \boldsymbol{\phi}_\ell)\right)^2, \tag{50}$$

where $\mathbf{p}(t)$ is ground truth location of the person, $\boldsymbol{\phi}_\ell = [\mu_\ell, \kappa_\ell, \gamma_\ell]$ are the model parameters and $h(\Delta_{\ell,\mathbf{p}(t)}, \boldsymbol{\phi}_\ell)$ is given in (3). In this paper, a nonlinear least-squares solver based on the interior-reflective Newton method described in [31] is used to find the minimum of $J(\boldsymbol{\phi}_\ell)$ and thereafter, the maximum likelihood estimate of σ_ℓ^2 is computed.

The two calibration methods reflect the two extremes of RTI. When using the empty-room calibration period, the model parameters are fixed apart from μ_ℓ and it is expected that the estimated images are noisy with multiple local maxima making it difficult to locate the person. On the other hand, supervised training can be used to adjust the model parameters to the specific propagation environment so that the estimated images reflect the changes in the environment more accurately. As a result, the images are expected to have less noise and typically they only have one maxima so that the person is easier to localize. Thus, the empty-room calibration method depicts a high clutter scenario whereas the supervised training method depicts a low clutter scenario.

C. Experimental Evaluation

The proposed system is evaluated with respect to a system that utilizes a KF which is the de facto choice for tracking in imaging-based PLT systems [10], [14], [19]. The benchmark system is implemented as described in Section II and the system is simply referred to as KF from now on. The system proposed in this paper is described in Section III and it is referred to as GSF for brevity from now on. Both systems use RTI for estimating the changes in the propagation environment and the difference lies in the position detector and tracking filter.

TABLE I
EXPERIMENTAL PARAMETERS

Parameter		
Pixel Variance	σ_b^2	0.005 (dB^2)
Correlation distance	δ_d	0.5 (m)
Pixel width	δ_p	0.25 (m)
Process noise	q	0.05 (m^2/s^3)
Measurement noise	σ_r^2	0.25 (m^2)
Detection probability	P_D	0.9
Clutter intensity	λ_c	10

The filters are evaluated using the root mean squared error (RMSE)

$$\text{RMSE} = \sqrt{\frac{1}{K}\sum_{k=1}^{K} \|\mathbf{p}_k - \hat{\mathbf{p}}_k\|_2^2}, \tag{51}$$

where K is the total number of estimates, $\mathbf{p}_k$ denotes the ground truth position, $\hat{\mathbf{p}}_k = \mathbf{H}\mathbf{m}_k$ the estimate, and $\|\cdot\|_2^2$ the square of the Euclidean norm.

D. Initialization and Experimental Parameters

Occupancy assessment is an important problem in PLT [11] but for simplicity, we assume we know the time instances when the person has entered the monitored area and is stationary at the first reference position. The tracking filters are initialized using an RTI image as follows.

- *GSF* – Let $\mathcal{Z}_0 = \{\mathbf{z}_0^1, \ldots, \mathbf{z}_0^J\}$ denote the set of position estimates computed using (31). For every measurement $\mathbf{z}_0^j$, the mean is initialized using $\mathbf{m}_0^j = \mathbf{z}_0^j \otimes \begin{bmatrix}1 \text{ m} & 0 \text{ m/s}\end{bmatrix}^\top$ and the covariance as $\mathbf{P}_0^j = \mathbf{I}_4$. Thus, the position of the mixture components are initialized to the peaks of the image, the velocity is set to zero and the initial uncertainty is equal to identity.
- *KF* – The benchmark system is initialized in a similar manner but only using one measurement. Let $\mathbf{z}_0$ denote

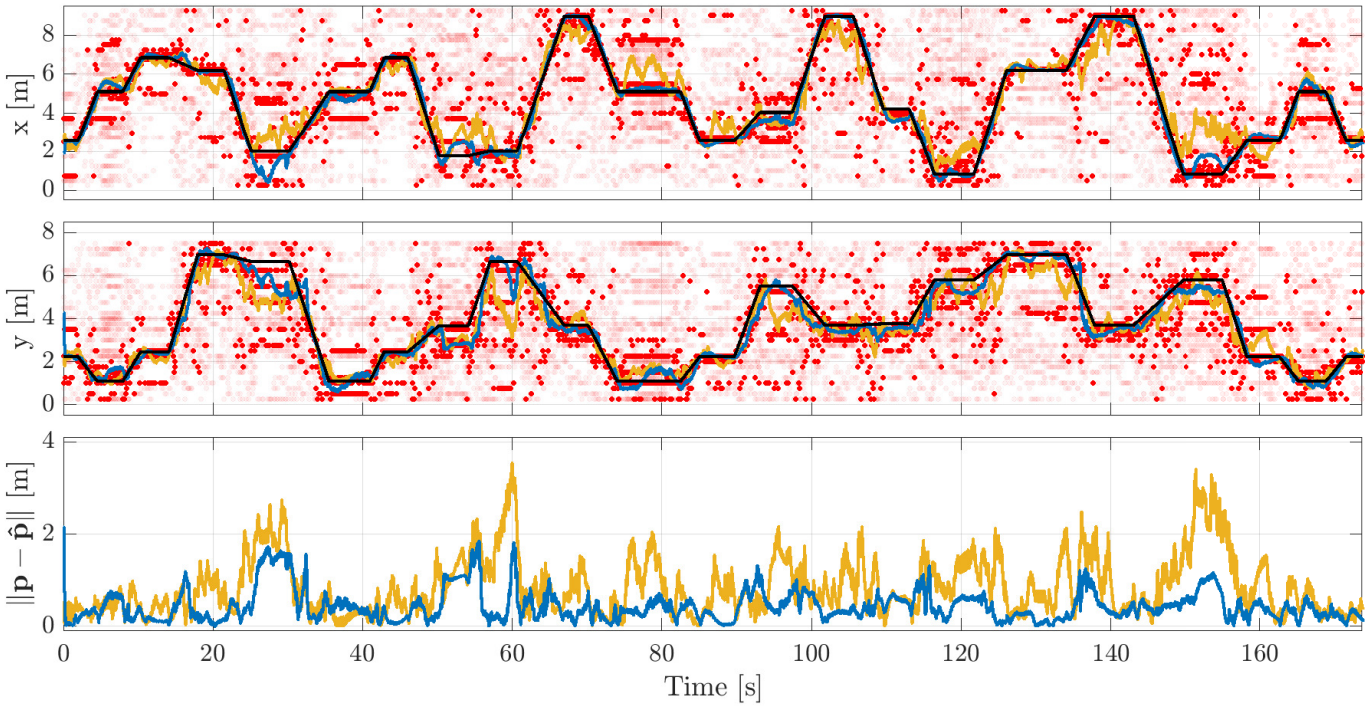


Fig. 3. Example tracking performance in the open environment under the high clutter scenario. On the top and in the middle, the x and y coordinate estimates using KF (——) and GSF (——). The ground truth coordinates are illustrated using (——), $\mathcal{Z}_k$ computed using (31) shown using transparent red dots and the position measurement with the highest pixel intensity computed using (9) illustrated with (•). On the bottom, the positioning error for KF (——) and GSF (——) as a function of time.

the position estimate obtained using (9). The mean is initialized to $\mathbf{m}_0 = \mathbf{z}_0 \otimes \begin{bmatrix} 1 \text{ m} & 0 \text{ m/s} \end{bmatrix}^\top$ and the covariance as $\mathbf{P}_0^j = \mathbf{I}_4$.

The parameters used in the experiments are given in Table I.

V. Results

A. High Clutter Scenario

In the high clutter scenario, the model parameters are the same for every wireless link and therefore, the model is unable to capture the RSS changes accurately. As a consequence, the images are noisy with multiple peaks and it is common that the pixel that has the highest intensity is far away from the actual location of the person as illustrated in Fig. 1b. Since the KF utilizes the pixel with highest intensity as the measurement in the update step, the performance of the filter is inevitably affected by the inaccurate position measurements. On the other hand, the multiple peaks can be considered as different location hypotheses and either the peak corresponds to clutter or a detection. The clutter measurements are not related to the actual position of the person whereas a detection corresponds to a measurement that is generated by the person. The GSF incorporates the clutter measurements and detections in the tracking filter and it can track multiple hypotheses over time. Furthermore, it is expected that the filter converges to the correct track since typically the image has a local maxima near the person's true location whereas the clutter measurements follow a PPP and are randomly located within the monitored area.

Example tracking performance in the open environment experiment is illustrated in Fig. 3 and as shown, the images in the high clutter scenario are very noisy since majority of the images produce multiple position measurements. Moreover, the position measurements with the highest intensity is typically inaccurate as shown in the figure. The KF does not use the entire set of measurements for tracking, only the peak with the highest intensity. The inaccurate position measurements utilized by the KF result to satisfactory tracking performance as illustrated in the figure. The GSF utilizes all the measurements and its ability to track multiple hypotheses over time is very beneficial in the high clutter scenario. As shown, the filter converges to the correct trajectory and tracks the person with high accuracy even with very cluttered position measurements. In the example scenario, the RMSE of the Kalman filter is 107 cm and the RMSE of the proposed Gaussian sum filter is 56 cm. Thus, the GSF is able to decrease the RMSE by 48 %

B. Low Clutter Scenario

In the low clutter scenario, supervised training is utilized to estimate the model parameters for every wireless link. This allows adapting the model to the particular propagation environment so that it captures the RSS changes more accurately. As a consequence, the quality of the image estimates improve which leads to enhanced position measurements. Moreover,

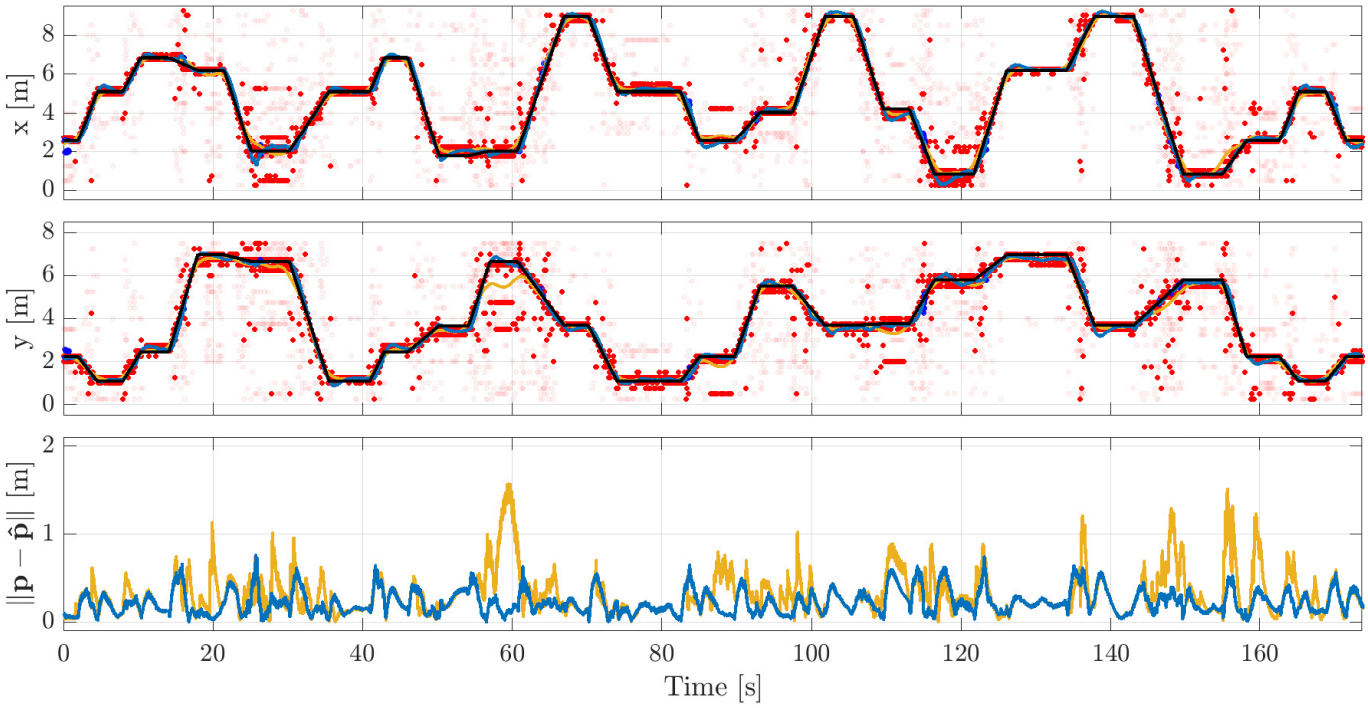


Fig. 4. Example tracking performance in the open environment under the low clutter scenario. On the top and in the middle, the x and y coordinate estimates using KF (——) and GSF (——). The ground truth coordinates are illustrated using (——), $\mathcal{Z}_k$ computed using (31) shown using transparent red dots and the position measurement with the highest pixel intensity computed using (9) illustrated with (•). On the bottom, the positioning error for KF (——) and GSF (——) as a function of time.

the clutter decreases and commonly, the image only contains one global maxima that is close to the actual location of the person as illustrated in Fig. 1a. With better position measurements, the performance of both filters improve.

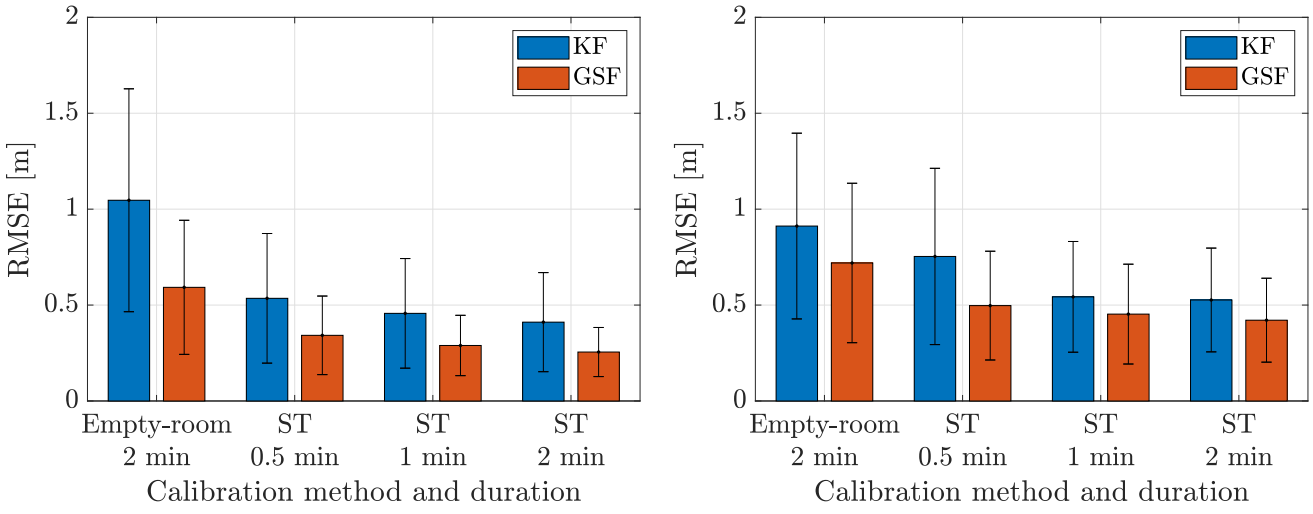
Example tracking performance in the open environment experiment is illustrated in Fig. 4 and as shown, the images in the low clutter scenario have much less noise compared to the high clutter scenario. Moreover, the position measurements with the highest intensity are typically much closer to the actual location of the person as shown in Fig. 4. Now the two filters yield comparative performance most of the time and the reason is obvious; there is either only one position measurement or the position measurement with highest image intensity is the one closest to the person and there is no benefit of considering multiple hypotheses. However, there are also time instances when the GSF is more accurate, for example at time $= 55 - 61$ s. During this time interval, the images are cluttered resulting to position measurements that are inaccurate and due to the reasons outlined in the previous section, the GSF is superior to the KF. Thus, even in low clutter scenarios, the GSF outperforms the KF. More quantitatively, the RMSE of the Kalman filter is 40 cm and the RMSE of the proposed Gaussian sum filter is 26 cm. Thus, even in the low clutter scenario, the GSF is able to decrease the RMSE by $35\,\%$.

C. Filter Performance

The performance of the filters in the two environments and with different calibration schemes are summarized in Fig. 5. As expected, the accuracy of both filters improve using supervised training and the performance gain is larger the longer the training period is. However, diminishing returns are already visible at the two minute mark and it is not expected that either filter would greatly benefit from a longer training period. The reason being, the used models cannot capture the human-induced RSS changes nor the dynamics of the person precisely. The used models are sufficient for localization and can be calibrated with a relatively short calibration period. However, the accuracy of the model cannot be improved indefinitely by just using more data and it is expected that the positioning sensor will always have imperfections since the indoor radio propagation channel is very complex. The results imply that the GSF can cope better with the imperfect positioning sensor and that it can be calibrated with less data. These two properties are desired in practical real-world deployments.

VI. Conclusions

This paper presents a RFS approach to RSS-based PLT. By modeling the measurements as a RFS, allows rigorous treatment of detections, misdetections, clutter and unknown DA. A new position detector is developed, a measurement likelihood function based on RFS theory is presented and a computationally tractable GSF for tracking is implemented. Analysis was carried out using two experimental indoor data sets. Results demonstrated the accuracy and robustness of the



(a) Open environment (b) Apartment

Fig. 5. Filter performance in the open environment and apartment experiment using empty-room calibration and different length supervised training (TS) periods. The whiskers illustrate the one standard deviation of the positioning errors.

proposed method, as well as, improved performance with respect to a benchmark system. Furthermore, the newly proposed RFS approach to RSS-based PLT admits numerous possibilities of future research into other RFS filters. For example, the probability hypotheses density (PHD) and Poisson multi-Bernoulli mixture (PMBM) filters could be utilized for MOT which is a difficult problem in RSS-based PLT.

References

[1] F. Adib, H. Mao, Z. Kabelac, D. Katabi, and R. C. Miller, "Smart homes that monitor breathing and heart rate," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 2015, p. 837–846.

[2] Q. Pu, S. Gupta, S. Gollakota, and S. Patel, "Whole-home gesture recognition using wireless signals," in *Proceedings of the 19th Annual International Conference on Mobile Computing and Networking*, 2013, p. 27–38.

[3] M. Youssef, M. Mah, and A. Agrawala, "Challenges: Device-free passive localization for wireless environments," in *Proceedings of the 13th Annual ACM International Conference on Mobile Computing and Networking*, 2007, p. 222–229.

[4] O. Kaltiokallio, M. Bocca, and N. Patwari, "A fade level-based spatial model for radio tomographic imaging," *IEEE Transactions on Mobile Computing*, vol. 13, no. 6, pp. 1159–1172, 2014.

[5] 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 Transactions on Mobile Computing*, vol. 14, no. 3, pp. 484–494, March 2015.

[6] 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 Journal on Selected Areas in Communications*, vol. 33, no. 11, pp. 2391–2403, 2015.

[7] S. Savazzi, M. Nicoli, F. Carminati, and M. Riva, "A Bayesian approach to device-free localization: modeling and experimental assessment," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 1, pp. 16–29, Feb 2014.

[8] 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 Journal*, vol. 16, no. 5, pp. 1253–1264, March 2016.

[9] 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 of Things Journal*, vol. 4, no. 2, pp. 351–362, April 2017.

[10] O. Kaltiokallio, R. Jäntti, and N. Patwari, "ARTI: an adaptive radio tomographic imaging system," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 8, pp. 7302–7316, Aug 2017.

[11] P. Hillyard and N. Patwari, "Never use labels: signal strength-based Bayesian device-free localization in changing environments," *IEEE Transactions on Mobile Computing*, vol. 19, no. 4, pp. 894–906, 2020.

[12] O. Kaltiokallio, R. Hostettler, H. Yiğitler, and M. Valkama, "Unsupervised learning in RSS-based DFLT using an EM algorithm," *Sensors*, vol. 21, no. 16, 2021.

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

[14] ——, "See-through walls: motion tracking using variance-based radio tomography networks," *IEEE Transactions on Mobile Computing*, vol. 10, no. 5, pp. 612–621, May 2011.

[15] M. Bocca, O. Kaltiokallio, N. Patwari, and S. Venkatasubramanian, "Multiple target tracking with RF sensor networks," *IEEE Transactions on Mobile Computing*, vol. 13, no. 8, pp. 1787–1800, Aug 2014.

[16] Q. Wang, H. Yiğitler, R. Jäntti, and X. Huang, "Localizing multiple objects using radio tomographic imaging technology," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 5, pp. 3641–3656, 2016.

[17] R. P. Mahler, *Advances in Statistical Multisource-Multitarget Information Fusion*. Artech House, 2014.

[18] B.-N. Vo and W.-K. Ma, "The Gaussian mixture probability hypothesis density filter," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4091–4104, 2006.

[19] Y. Zhao and N. Patwari, "Robust estimators for variance-based device-free localization and tracking," *IEEE Transactions on Mobile Computing*, vol. 14, no. 10, pp. 2116–2129, Oct 2015.

[20] Y. Li, X. Chen, M. Coates, and B. Yang, "Sequential Monte Carlo radio-frequency tomographic tracking," in *Acoustics, Speech and Signal Processing, 2011 IEEE International Conference on*, 2011, pp. 3976–3979.

[21] N. Patwari and P. Agrawal, "Effects of correlated shadowing: connectivity, localization, and RF tomography," in *Information Processing in Sensor Networks, 2008 International Conference on*, 2008, pp. 82–93.

[22] Y. Bar-Shalom and X.-R. Li, *Estimation with Applications to Tracking and Navigation*. John Wiley & Sons, Inc., 2001.

[23] S. Särkkä, *Bayesian Filtering and Smoothing*. Cambridge University Press, 2013.

[24] S. Challa, M. R. Morelande, D. Mušicki, and R. J. Evans, *Fundamentals of Object Tracking*. Cambridge University Press, 2011.

[25] L. Svensson, Y. Xia, and K. Granström, "Section 2: Single-object tracking in clutter," in *Multi-Object Tracking for Automotive Systems—ChalmersX - ChM013x*, 2023. [Online]. Available: https://www.edx.org/course/multi-object-tracking-for-automotive-systems

[26] H. Sorenson and D. Alspach, "Recursive Bayesian estimation using Gaussian sums," *Automatica*, vol. 7, no. 4, pp. 465–479, 1971.

[27] D. Alspach and H. Sorenson, "Nonlinear Bayesian estimation using Gaussian sum approximations," *IEEE Transactions on Automatic Control*, vol. 17, no. 4, pp. 439–448, 1972.

[28] *A USB-enabled system-on-chip solution for 2.4 GHz IEEE 802.15.4 and ZigBee applications*, Texas Instruments, June 2010, http://www.ti.com/lit/ds/symlink/cc2531.pdf.

[29] *IEEE 802.15.4-2003 standard*, IEEE, October 2003, http://www.ieee802.org/15/pub/TG4Expert.html.

[30] M. Bocca, O. Kaltiokallio, and N. Patwari, "Radio tomographic imaging for ambient assisted living," in *Evaluating AAL Systems Through Competitive Benchmarking*, S. Chessa and S. Knauth, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 108–130.

[31] T. F. Coleman and Y. Li, "An interior trust region approach for nonlinear minimization subject to bounds," *SIAM Journal on Optimization*, vol. 6, no. 2, pp. 418–445, 1996.