Multiple Target Tracking with Gaussian Mixture PHD Filter using Passive Acoustic Doppler-Only Measurements

Mehmet B. Guldogan, David Lindgren†, Fredrik Gustafsson, Hans Habberstad†, Umut Orguner
Automatic Control, Department of Electrical Engineering
Linköping University, Sweden
†Swedish Defence Research Agency (FOI), Sweden
Email: {bguldogan, umut, fredrik}@isy.liu.se
†{david.lindgren, hans.habberstad}@foi.se

Abstract—In this paper, we present the performance of the Gaussian mixture probability hypothesis density (GM-PHD) filter in tracking multiple ground targets using a passive acoustic-sensor network. For this purpose, an experimental setup consisting of a network of microphones and a loudspeaker was prepared. Non-cooperative transmissions from a loudspeaker (i.e. illuminator of opportunity) are exploited by non-directional separately located microphones (i.e. Doppler measuring sensors). Experimental proof-of-concept study results show that it is possible to track multiple ground targets using only Doppler shift measurements in a passive multi-static scenario.

I. INTRODUCTION

Recently, there has been a great deal of interest in the field of bi/multi-static radar/sonar systems [1]. Although this type of radar systems have a long history, there have not been enough amount of operational systems. Modern multi-static systems, consisting of cooperative/noncooperative transmitters and multiple receivers, collect several independent target measurements such as time-of-arrival (TOA), direction-of-arrival (DOA) and Doppler shift. These measurements are then fused at the fusion center to obtain an estimate of the target position. For target surveillance, multi-static passive radar systems exploit illuminators of opportunity like FM radio transmitters, digital audio/video broadcasters (DAB/DVB), WiMAX systems and global system for mobile-communication (GSM) base stations [2], [3], [4], [5]. There exist very important advantages of passive radar systems over active systems: no frequency allocation problem, receivers are hidden for a possible jamming, energy efficient and much lower costs. Particularly, there is an increasing research interest in GSM-based passive radar systems, due to several unique advantages provided by these systems [5]. The key advantage is that the cellular phone systems have an almost hundred percent coverage of most parts of the world. Secondly, multiple base stations can be utilized in a multi-static passive radar network to improve the overall performance and robustness. Although the GSM waveform has poor range resolution, it can achieve very good Doppler resolution, which makes the GSM-based passive radar suitable for Doppler detection and tracking [5].

Position estimation of a moving target using only Doppler shift measurements is actually an old problem studied in different contexts [6], [7], [8]. However, although Doppler measuring sensors are very simple sensors and no hardware array is required unlike the DOA measuring arrays, analysis of multi-static passive systems that use Doppler only measurement has not been widely implemented yet. Some very recent studies on multi-static systems using Doppler only measurements can be found in the references [9], [10], [11], [12], [13]. These papers mainly concentrate on the static estimation solutions, observability analysis of the target using Doppler-only measurements and optimal positioning of the passive system. In [14], using recent filtering theories, a realistic multi-target tracking scenario taking into account clutter, missed detections, multi-static Doppler variances and process variances is demonstrated.

In this paper, we make use of the probability hypothesis density (PHD) filter framework, which is based on the random finite sets (RFS) approach [15], to sequentially estimate target position and velocity from Doppler shift measurements. The RFS approach is considered to be a very promising alternative to handle the multi-target multi-detection association problem faced in multi-target tracking applications [15], [16]. The RFS approach treats the collection of individual measurements and the individual targets as a set-valued measurement and set-valued RFS approach is considered to be a very promising alternative to handle the multi-target multi-detection association problem faced in multi-target tracking applications [15], [16]. The RFS approach treats the collection of individual measurements and the individual targets as a set-valued measurement and set-valued state, respectively. It is shown that the sequential estimation of multi-targets buried in clutter with association uncertainties can be formulated in a Bayesian filtering framework by modelling set-valued measurements and set-valued states as RFSs [15]. The PHD filter, an approximation of this theoretically optimal approach to multi-target tracking, propagates the first-order statistical moment of the RFS of states in time and avoids the combinatorial data association problem. The PHD filter recursions are evaluated on the single target state. Despite its advantages, the recursions of the PHD filter involves multiple integrals having no closed form solutions. There exist several different types of implementations of the PHD filter in the literature. Particle based implementations have gained interest from researchers in different fields due...
signals (data) with a known carrier frequency, an opportunity (TX), say a GSM base station, constantly transmits pure sinusoids at a known certain frequency. Re-

loudspeaker is directed towards the road and continuously generates Doppler shift measuring sensors (black triangles). Each Doppler shift measuring sensor transmits information to a microphone network which provides Doppler shift measurements. The organization of the paper is as follows. Mathematical formulation of the problem and the measurement model are presented in Section II. Section III presents RFS formulation of multi-target tracking and the PHD filter. Section IV provides GM-PHD formulations. Experimental details are given in Section V. Results are presented in Section VI. Lastly, conclusion and possible future directions are given in Section VII.

II. SENSOR AND TARGET MODEL

The scenario in this work is as follows: An illuminator of opportunity (TX), say a GSM base station, constantly transmits signals (data) with a known carrier frequency, \( f_c \), to mobiles in its corresponding cell as illustrated in Fig 1. Transmitted signals are reflected from the detected moving targets and received by each Doppler shift measuring sensor separately distributed in the area. It is assumed that the location of transmitter and the sensors are known to the fusion center and each sensor sends its measurement to the fusion center. The state vector of a target at time \( k \) is

\[
x_k = [x_k, y_k, \dot{x}_k, \dot{y}_k]^T,
\]

where \([x_k, y_k]\) is the position, \([\dot{x}_k, \dot{y}_k]\) is the velocity of the target and \( T \) denotes transpose operation. The target dynamic is modeled by linear Gaussian constant velocity model [22]:

\[
x_k = F_{k-1} x_{k-1} + v_{k-1},
\]

where

\[
F_{k-1} = \begin{bmatrix}
I_2 & \Delta I_2 \\
0_2 & I_2
\end{bmatrix},
\]

\[
Q_{k-1} = \sigma_v^2 \begin{bmatrix}
\frac{\Delta^4 I_2}{2} & \frac{\Delta^3 I_2}{3} \\
\frac{\Delta^3 I_2}{3} & \frac{\Delta^2 I_2}{2}
\end{bmatrix},
\]

\( F_{k-1} \) is the state transition matrix, \( v_{k-1} \sim \mathcal{N}(v; 0, Q_{k-1}) \) is the white Gaussian process noise, \( Q_{k-1} \) is the process noise covariance matrix, \( \Delta \) is the sampling interval, \( \sigma_v \) is the standard deviation of the process noise, \( I_2 \) and \( 0_n \) denote \( n \times n \) identity and zeros matrices respectively.

Doppler shift measurements are collected by separately located \( N_s \) sensors. Measured Doppler shift by the \( i^{th} \) sensor located at \([x_i, y_i]\), \( i = 1, \ldots, N_s \) can be given as

\[
z_{k,i} = h_i(x_k) + \varepsilon_{k,i},
\]

where

\[
h_i(x_k) = -\frac{\dot{x}_k(x_k - x_i) + \dot{y}_k(y_k - y_i)}{\lambda d_{k,i}^2},
\]

\[
\varepsilon_{k,i} \sim \mathcal{N}(\varepsilon_i; 0, R_{k,i})
\]

is the Doppler shift, \( \lambda \) is the wavelength of the transmitted signal, \((x_i', y_i')\) is the transmitter location and \( \varepsilon_{k,i} \) is measurement noise in sensor \( i \), \( \varepsilon_{k,i} \sim \mathcal{N}(\varepsilon_i; 0, R_{k,i}) \) with \( R_{k,i} = \sigma^2 \varepsilon_i \), \( d_{k,i}^2 \) is the distance between the transmitter and the target at time \( k \):

\[
d^2_{k,i} = (x_k - x_i')^2 + (y_k - y_i')^2,
\]

and \( d_{k,i} \) is the distance between the target and the \( i^{th} \) sensor at time \( k \):

\[
d^2_{k,i} = (x_k - x_i)^2 + (y_k - y_i)^2.
\]

We also provide Jacobian of \( h_i(x_k) \), \( \mathbf{H}_{k,i} \), to be used in the filtering as

\[
\mathbf{H}_{k,i} = \begin{bmatrix}
\frac{\partial h_i(x_k)}{\partial x_k} & \frac{\partial h_i(x_k)}{\partial y_k} & \frac{\partial h_i(x_k)}{\partial \dot{x}_k} & \frac{\partial h_i(x_k)}{\partial \dot{y}_k}
\end{bmatrix}
\]

and each of its elements are

\[
\frac{\partial h_i(x_k)}{\partial x_k} = -\frac{\dot{x}_k d_{k,i}^2 - (x_k - x_i')d_{k,i}^2}{\lambda d_{k,i}^4} \quad \frac{\partial h_i(x_k)}{\partial y_k} = -\frac{\dot{y}_k d_{k,i}^2 - (y_k - y_i')d_{k,i}^2}{\lambda d_{k,i}^4}
\]

Fig. 1. Transmitted signal from an illuminator of opportunity TX (black circle) is reflected from a detected moving target (red square) and received by each Doppler shift measuring sensors (black triangles). \( s_i : i = 1, \ldots, N_s \).
the set of all possible finite subsets of state space $X$ measurement space $Z$

At time $k$, $X_k$ can be written as

$$Z_k = K_k \cup \bigcup_{x \in X_k} \Theta_k(x)$$  \hspace{1cm} (18)

where $K_k$ is the RFS of clutter or false measurements, $\Theta_k(x)$ is the RFS of multi-target state originated measurements, which can take values either $z_k$ if target is detected, or 0 if target is not detected [20].

B. Multi-target Filtering

Having very briefly summarized some key points of the RFS framework, we can define the RFS based multi-target Bayes filter. The optimal multi-target Bayes filter propagates the multi-target posterior density $p_k(:|Z_{1:k})$ conditioned on the sets of measurements up to time $k$, $Z_{1:k}$, in time with the following recursion

$$p_k(X_k|Z_{1:k-1}) = \int f_{k|k-1}(X_k|X)p_{k-1}(X|Z_{1:k-1})\mu_\lambda(dX) \hspace{1cm} (19)$$

$$p_k(X_k|Z_{1:k}) = \frac{g_k(Z_k|X_k)p_{k|k-1}(X_k|Z_{1:k-1})}{\int g_k(Z_k|X_k)p_{k|k-1}(X_k|Z_{1:k-1})\mu_\lambda(dX)} \hspace{1cm} (20)$$

where $f_{k|k-1}$ is the transition density, $g_k(Z_k|X_k)$ is the multi-target likelihood and $\mu_\lambda$ denotes the Lebesgue measure on the set $\mathcal{F}(X)$ [17]. The multi-target Bayes recursion involves multiple integrals and the complexity of computing it grows exponentially with the number of targets. Therefore, it is not practical for scenarios where there exist more than a few targets.

C. The Probability Hypothesis Density Filter

To alleviate the computational burden in calculating the optimal filter given above, the PHD filter was proposed as a practical suboptimal alternative [16]. The PHD filter propagates the first-order statistical moment of the posterior multi-target state, instead of propagating the multi-target posterior density. Consider that, intensities associated with the multi-target posterior density $p_k$ and the multitarget predicted density $p_{k|k-1}$ in the optimal multi-target Bayes recursion are represented with $v_k$ and $v_{k|k-1}$ respectively. The PHD recursion is defined as

$$v_{k|k-1}(x) = \int \mu_\lambda(d\xi) p_{S,k}(\xi)f_{k|k-1}(x|\xi)v_{k-1}(\xi)d\xi + \gamma_k(x)$$  \hspace{1cm} (21)

$$v_k(x) = [1 - p_{D,k}(x)]v_{k|k-1}(x) + \sum_{z \in Z_k} \kappa_k(z) + \int \mu_\lambda(d\xi) g_k(z|\xi)v_{k|k-1}(\xi)$$  \hspace{1cm} (23)

where $p_{S,k}(\xi)$ is the probability of target survival at time $k$ given that its previous state is $\xi$, $\gamma_k(x)$ is the intensity of spontaneous birth RFS at time $k$, $p_{D,k}(x)$ is the probability of target detection at time $k$ and $\kappa_k(z)$ is the intensity of clutter RFS at time $k$. 

III. RANDOM FINITE SETS (RFS) BASED FILTERING

RFS formulation based multiple target tracking framework proposed by R. Mahler, combines the problems of combinatorial data association, detection, classification and target tracking within a unified compact Bayesian paradigm [15], [16]. In the following subsections, some RFS notions, multiple target generalization of a Bayesian filter and its first order approximation probability hypothesis density (PHD) filter are described.

A. RFS Formulation

The RFS approach treats the collection of the individual targets and individual measurements as a set-valued state and set-valued measurement, respectively, as:

$$X_k = \{x_{k,1}, ..., x_{k,M(k)}\} \in \mathcal{F}(\mathcal{X})$$  \hspace{1cm} (15)

$$Z_k = \{z_{k,1}, ..., z_{k,N(k)}\} \in \mathcal{F}(\mathcal{Z})$$  \hspace{1cm} (16)

where $M(k)$ is the number of targets at time $k$, $N(k)$ is the number of measurements at time $k$, $\mathcal{F}(\mathcal{X})$ and $\mathcal{F}(\mathcal{Z})$ are the set of all possible finite subsets of state space $\mathcal{X}$ and measurement space $\mathcal{Z}$, respectively. An RFS model for the time evolution of a multi-target state $X_{k-1}$ at time $k - 1$ to the multi-target state $X_k$ at time $k$ is defined as

$$X_k = \bigcup_{\zeta \in X_{k-1}} S_{k|k-1}(\zeta) \cup \Gamma_k \hspace{1cm} (17)$$

where $S_{k|k-1}(\zeta)$ is the RFS of surviving targets from previous state $\zeta$ at time $k$ and $\Gamma_k$ is the RFS of spontaneous target births at time $k$. The RFS measurement model for a multi-target state

$$\frac{\partial h_i(x_k)}{\partial y_k} = -\left\{ \frac{y_k d_{k,i}^t - (y_k - y_i') d_{k,i}^t}{\lambda d_{k,i}^t} \right\}$$  \hspace{1cm} (10)

$$\frac{\partial h_i(x_k)}{\partial x_k} = -\left\{ \frac{(y_k - x') + (y_k - x_i)}{\lambda d_{k,i}^t} \right\}$$  \hspace{1cm} (11)

$$\frac{\partial h_i(x_k)}{\partial y_k} = -\left\{ \frac{(y_k - y') + (y_k - y_i)}{\lambda d_{k,i}^t} \right\}$$  \hspace{1cm} (12)

where total derivatives of $d_{k,i}$ and $d_{k,i}^t$ are, respectively

$$\dot{d}_{k,i} = \dot{x}_k(x_k - x_i) + \dot{y}_k(y_k - y_i)$$  \hspace{1cm} (13)

$$\ddot{d}_{k,i} = \dot{x}_k(x_k - x_i') + \dot{y}_k(y_k - y_i')$$  \hspace{1cm} (14)
There exist practical implementations of the PHD filter in the literature [17], [19], [23]. In the next section, we describe one of those implementations known as Gaussian mixture PHD [20].

IV. THE GAUSSIAN MIXTURE PHD (GM-PHD) FILTER

Vo et al. derived a closed-form solution to the PHD filter, called as the Gaussian mixture PHD (GM-PHD) under linear Gaussian multi-target models in [20]. The GM-PHD filter estimates multi-target states by determining the Gaussian components with highest/thresholded weights. The GM-PHD filter has been successfully used in many different applications [14], [24], [25], [26]. There are several assumptions used in the GM-PHD recursions. The first one is that each target follows a linear Gaussian dynamical and measurement model:

\[
 f_{k|k-1}(x|z) = \mathcal{N}(x; F_{k-1}z, Q_{k-1}) , \quad g_k(z|x) = \mathcal{N}(z; H_kx, R_k) .
\]

(24)

(25)

Secondly, the detection and survival probabilities are state independent: \(p_{D,k}(x) = p_{D,k} \) and \( p_{S,k}(x) = p_{S,k} \). Lastly, the intensity of the birth RFS is Gaussian mixtures of the form

\[
 \gamma_k(x) = \sum_{i=1}^{J_{\gamma,k}} w_{\gamma,k}^{(i)} \mathcal{N}(x; m_{\gamma,k}^{(i)}, P_{\gamma,k}^{(i)}) ,
\]

(26)

where \( J_{\gamma,k}, w_{\gamma,k}^{(i)}, m_{\gamma,k}^{(i)} \) and \( P_{\gamma,k}^{(i)} \) are given model parameters that determine the birth intensity. Posterior intensity at time \( k-1 \) can be written as a sum of Gaussian components with different weights, means and covariances as

\[
 v_{k-1}(x) = \sum_{i=1}^{J_{\gamma,k}} w_{\gamma,k}^{(i)} \mathcal{N}(x; m_{\gamma,k}^{(i)}, P_{\gamma,k}^{(i)})
\]

(27)

and an identifying label \( \ell_{k-1} \) is assigned to each created Gaussian component. A label table, \( \mathcal{L}_{k-1} \), is formed as

\[
 \mathcal{L}_{k-1} = \{ \ell_{1,k-1}^{(1)}, \ldots, \ell_{J_{\gamma,k}}^{(J_{\gamma,k})} \} .
\]

(28)

At time \( k \), the predicted intensity is also a Gaussian mixture:

\[
 v_{k|k-1}(x) = v_{S,k|k-1}(x) + \gamma_k(x) ,
\]

(29)

where

\[
 v_{S,k|k-1}(x) = p_{S,k} \sum_{j=1}^{J_{S,k}} w_{S,k}^{(j)} \mathcal{N}(x; m_{S,k}^{(j)}, P_{S,k}^{(j)})
\]

(30)

\[
 m_{S,k|k-1}^{(j)} = F_{k-1}m_{S,k}^{(j)} ,
\]

(31)

\[
 P_{S,k|k-1}^{(j)} = Q_{k-1} + F_{k-1}P_{k-1}F_{k-1}^T
\]

(32)

Each birth component is assigned a new label and concatenated with the previous time labels,

\[
 \mathcal{L}_{k|k-1} = \mathcal{L}_{k-1} \cup \mathcal{L}_{\gamma,k-1} .
\]

(33)

The posterior intensity at time \( k \) is also a Gaussian mixture and can be written as

\[
 v_k(x) = (1 - p_{D,k}) v_{k|k-1}(x) + \sum_{z \in \mathcal{L}_k} v_{D,k}(x; z) ,
\]

(34)

where

\[
 v_{D,k}(x; z) = \sum_{j=1}^{J_{k|k-1}} w_{k}^{(j)} \mathcal{N}(x; m_{k|k-1}^{(j)}, P_{k|k}^{(j)})
\]

(35)

\[
 w_{k}^{(j)}(z) = \frac{P_{D,k} w_{k|k-1}^{(j)}(z)}{\kappa_k(z) + \sum_{t=1}^{J_{k|k-1}} w_{k}^{(t)} \tilde{q}_{k}^{(t)}(z)}
\]

(36)

\[
 \tilde{q}_{k}^{(j)}(z) = \mathcal{N}(z; H_km_{k|k-1}^{(j)}, H_kP_{k|k-1}^{(j)}H_k^T)
\]

(37)

\[
 m_{k|k}^{(j)}(z) = m_{k|k-1}^{(j)} + K_k^{(j)}(z - H_km_{k|k-1}^{(j)})
\]

(38)

\[
 K_k^{(j)} = P_{k|k-1}^{(j)}H_k^T(H_kP_{k|k-1}^{(j)}H_k^T + R_k)^{-1}
\]

(39)

There will be \( |Z_k| + 1 \) Gaussian components for each predicted term, where \( |\cdot| \) is the cardinality of a set. Then, identifying label at time \( k \) is

\[
 \mathcal{L}_k = \mathcal{L}_{k|k-1} \cup \mathcal{L}_{\gamma,k-1} \cup \ldots \cup \mathcal{L}_{\gamma,k|k} .
\]

(41)

As time progresses, the number of Gaussian components increases and computational problems occur. To alleviate this problem, a simple pruning and merging can be used to decrease the number of Gaussian components propagated [20]. Firstly, weights below a predefined threshold are eliminated. Then, closedly spaced Gaussian components are merged into a single Gaussian component. Starting with the strongest weighted component, \( w_k^{(1)} \), components are merged in a set \( W_k^{(1)} \) by

\[
 W_k^{(1)} := \{ i : (m_k^{(i)} - m_k^{(1)})^T(P_k^{(1)})^{-1}(m_k^{(i)} - m_k^{(1)}) \leq \rho \}
\]

(42)

and the resulting merged component parameters are

\[
 \tilde{w}_k^{(1)} = \sum_{i \in W_k^{(1)}} w_k^{(i)}
\]

(43)

\[
 \tilde{m}_k^{(1)} = \frac{1}{\tilde{w}_k^{(1)}} \sum_{i \in W_k^{(1)}} w_k^{(i)} x_k^{(i)}
\]

(44)

\[
 \tilde{P}_k^{(1)} = \frac{1}{\tilde{w}_k^{(1)}} \sum_{i \in W_k^{(1)}} w_k^{(i)} (m_k^{(i)} - \tilde{m}_k^{(1)})(m_k^{(i)} - \tilde{m}_k^{(1)})^T
\]

(45)

In order to extract multi-target states, means of the Gaussian components, that have weights greater than some predefined threshold, are selected:

\[
 \mathcal{L}_k = \{ \ell_k^{(i)} : w_k^{(i)} > \rho \} ,
\]

(46)

and the estimated target states set is

\[
 \hat{X}_k = \{ m_k^{(i)}, P_k^{(i)} : \mathcal{L}_k^{(i)} \in \mathcal{L}_k \} .
\]

(47)

In the following two sections, we present details of the
due to the high loudspeaker sound pressure in fact needed.
Certainly, the experiment could have been designed in other
ways, but recall that the motivation for the tracking technique
in focus is that it in the future will be used with (narrow-band)
radio transmitters of opportunity, and that we have used the
acoustics as a first step to prove the tracking concept.

At the trial, the acoustic background noise level is rather
high due to nearby construction work and traffic; around
60–70 dB SPL. The Doppler detection is however restricted
to the frequencies just above and below the 10 kHz tone,
frequency regions we define as 10k±[100 800] Hz. The noise
contributions in these frequency regions sum up to around
10 dB SPL. When the LS is switched on, the noise level
rises to 30 dB SPL (with no target present), so 20 dB SPL
noise happens to be self-induced and attributable to non-linear
paths and driver imperfections (speculatively). According to
the free-space propagation model, the sound pressure drops
as 1/d, where d denotes the sound travel distance. If we
consider the reflecting target as a new acoustic source, the
combined drop of will however be closer to 1/d^2, which
is in analogy with the radar equation. At 10 kHz, there is
also significant atmospheric attenuation; typically within the
interval 0.1 through 0.3 dB SPL/m (depending on the air
humidity, temperature etc.) For example, if both the LS/target,
and the target/sensor distance are 50 m, and 0.2 dB SPL/m
attenuation is assumed, the drop will be

\[ 20 \log_{10} \left( \frac{50^{-2}}{d} \cdot 2 \cdot 50 \cdot 0.2 \right) \approx -88 \text{ dB SPL}. \]

Then there are of course a loss as the sound reflects in various
surfaces on the vehicle. These reflection surfaces are not trivial
to identify, and thus it is difficult to predict the reflection
loss. Our observations from the experiment indicate however
that the signal falls below the noise floor at 30 dB SPL
(SNR=0 dB SPL) at around 50 m target distance (with the
source at 129 dB SPL), which leaves a reflection loss of
roughly −11 dB SPL (= 129 − 88 − 30). In conclusion, 50 m
would be close to the range limit for acoustic Doppler radar,
at least in the case with a narrow-band 10 kHz source.

VI. EXPERIMENTAL RESULTS

In this section, we present performance results of the
GM-PHDFilter in tracking 2 vehicles using noisy Doppler
shift measurements of separately distributed 3 microphones.
Extended Kalman PHD (EK-PHD) filter is implemented to
accommodate nonlinear models [20]. For the sake of clarity,
we considered the single sensor update of the PHD filter in
formulations. Extending the PHD filters for multi-sensor multi-
target tracking is still an open problem in the literature. In
[27] and [28] Mahler discusses possible ways of extending
PHD algorithms for multiple sensors. Here, we use an iterative
corrector approach, in which the sensors apply the correction
step over the corrected PHD of the previous sensor. Total
duration of the experiment is 13 s. Carrier frequency, f_c, of
the transmitted signal is 10 kHz. Vehicle-1 starts its motion at
time k = 0 s, position [78, 72.8] m and dies at time k = 13 s,
position [9.76, 43.9] m. Vehicle-2 is born at time k = 7.5 s,
position \([17.2, 46.7]m\) and dies at time \(k = 13s\), position \([76.5, 73.3]m\). True vehicle trajectories are plotted in Fig. 4. Blue and red lines denote the vehicle trajectories of the vehicle-1 and vehicle-2, respectively. The symbol \(\Box\) represents locations at which vehicles are born and \(\square\) represents locations at which vehicles die. Blue stars are for microphone positions and black star is for LS location. Black \(\Diamond\) is the look direction of the LS. Short-time Fourier transform (STFT) is used to get the time-frequency diagram of microphone outputs and detection is performed in the time-frequency domain. STFT of the first microphone is seen in Fig. 5. Peaks of the STFT that exceed a threshold level are collected as measurements. However, different kinds of instantaneous frequency estimation techniques can also be adapted to have better detection results [29], [30], [31]. Doppler-shifts corresponding to two vehicles around the carrier frequency can be distinguished in the time-frequency domain. Depending on the look direction of the LS and field-of-view (FOV) of the microphones, signal-to-noise ratio (SNR) of the reflections vary in time. Moreover, strong harmonics appear around the fundamental frequencies, which have negative effect on tracking performance.

Some parameters used in the filtering are: \(\Delta = 85\text{ms}, \sigma_v = 4\text{m/s}^2, \sigma_{D,k} = 0.78, \sigma_{S,k} = 0.98, \sigma_c = 5\text{Hz}, \rho = 2\) and \(J_{\text{max}} = 70\). The clutter RFS, \(K_k\), follows a uniform Poisson model over the surveillance region \([-800, 800]\text{Hz}\), with an average number of clutter returns per unit region, \(\gamma_c = 1 \times 10^{-5}\text{Hz}^{-1}\). Considering the FOV of the microphones, relative position of the road segment and the two-way acoustic path attenuation, \(N_\gamma = 7\) Gaussian distributions means of which are uniformly distributed on the road segment, form the birth process intensity. The spontaneous birth RFS is Poisson with intensity

\[
\gamma_k(x) = \frac{0.2}{N_\gamma} \sum_{i=1}^{N_\gamma} \mathcal{N}(x; m^{(i)}, P) ,
\]

where

\[
m^{(i)} = [85, 75, 0, 0, 0]^T, \ldots, m^{(N_\gamma)} = [15, 45, 0, 0, 0]^T,
\]

\[
P = \begin{bmatrix} P_{x,y} & 0_2 \\ 0_2 & P_{x,y} \end{bmatrix}, \quad P_{x,y} = \begin{bmatrix} 25 & 8 \\ 15 & 20 \end{bmatrix}, \quad P_{x,y} = \begin{bmatrix} 30 & 15 \\ 15 & 20 \end{bmatrix}
\]

where \(0_2\) is the two dimensional zero matrix. Tracking results of the GM-PHD filter are presented in Figs. 6, 7. It is seen from the figures that the GM-PHD filter successfully detects and tracks position and velocity of the vehicles from only Doppler shift measurements. At some time instants, short discontinuities occur in the tracks. This is due to the harmonics of the fundamental acoustic reflection and having less number of measurements around carrier frequency. Note that, during the transient regions where the sign of the Doppler shifts change, vehicles behave like extended targets and produce more than one measurement. This effect should be taken care of for better accuracy. As stated, the EK approximation is used in this paper. However, it is expected that the results can be improved by using unscented Kalman (UK) approximation [32]. Moreover, the number of target estimates of the filter is plotted in Fig. 8. The fluctuations in the number of targets estimated can be improved by using the GM implementation of the cardinalized PHD (CPHD) filter which requires additional computational resources [33], [21]. Observability of targets is a problem in target tracking when only Doppler measurements are used [10]. However, this problem can be alleviated by systematically placing sensors with respect to the transmitter [14]. Therefore, we believe that, with a better microphone placement around road, tracking results will significantly improve. Moreover, based on the results in this paper and in [14], it could be safely said that the off-road tracking of multiple vehicles is possible using a carefully located microphone sensor network covering the area in consideration. In that case, much less number of Gaussian distributions with large position variances (to cover the area of interest) can be used to form
the birth process intensity.

VII. CONCLUSION

In this paper, we studied the performance of the GM-PHD filter in tracking multiple targets using passive acoustic Doppler shift measurements from multiple separately located sensors. Transmissions at a known frequency from a loudspeaker are exploited by non-directional microphones. This experimental proof-of-concept study shows that it is possible to track multiple ground targets from cluttered Doppler measurements collected by multiple microphones by using the GM-PHD filter. A future direction would be to increase the coverage of the passive system by adding more microphones systematically to be able to cover a longer segment of the road and get more accurate and robust results.

ACKNOWLEDGMENT

The authors gratefully acknowledge fundings from the EL-LIIT joint research program, Swedish Foundation for Strategic Research (SSF) in the center MOVIII and the Swedish Defence Research Agency (FOI) Center for Advanced Sensors, Multisensors and Sensor Networks (FOCUS) funded by the Swedish Governmental Agency for Innovation Systems (VINNOVA).

REFERENCES

PHD filter for multi-target tracking using passive Doppler-only mea-
surements,” in IET Data Fusion & Tracking Conf., 2012.
[16] R. Mahler, “Multitarget Bayes filtering via first-order multitarget mo-
[17] B. N. Vo, S. Singh, and A. Doucet, “Sequential Monte Carlo implementa-
[18] D. Clark and J. Bell, “Convergence results for the particle-PHD filter,”
[19] B. Ristic, D. Clark, and B. Vo, “Improved SMC implementation of the
4104, Nov. 2006.
cardinalized probability hypothesis density filter,” IEEE Trans. Signal
[22] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, Estimation with Applica-
multiple target tracking with a network of acoustic sensor arrays
using PHD and CPHD filters,” in IEEE Int. Conf. Information Fusion
(FUSION), 2011.
in sonar images,” in SPIE Proc. SPIE Defense and Security Symposium,
2006.
mixture cardinalized probability hypothesis density filter,” IEEE Trans.
[26] D. Laneuville and J. Houssineau, “Passive multi target tracking with
GM-PHD filter,” in IEEE Int. Conf. Information Fusion (FUSION),
2010.
[27] R. Mahler, “The multisensor PHD filter: I. general solution via multi-
[28] ——, “The multisensor PHD filter: II. erroneous solution via Poisson
[29] B. Boashash, “Estimating and interpreting the instantaneous frequency
of a signal. II. algorithms and applications,” Proc. IEEE, vol. 80, no. 4,
based instantaneous frequency estimation in a high noise environment,”
[31] E. Sejdic, I. Djurovic, and J. Jiang, “Timefrequency feature represen-
tation using energy concentration: An overview of recent advances,”