A Probabilistic Hypothesis Density Filter for Traffic Flow Estimation in the Presence of Clutter

Matthieu Canaud*, Lyudmila Mihaylova†, Nour-Eddin El Faouzi*‡, Romain Billot*‡ and Jacques Sau*
*Université de Lyon, F-69000, Lyon, France
IFSTTAR, LICIT, F-69500, Bron
ENTPE, LICIT, F-69518, Vaulx en Velin
Email: matthieu.canaud@ifsttar.fr
†School of Computing and Communications
InfoLab21, South Drive, Lancaster University
Lancaster LA1 4WA, UK
Email: mila.mihaylova@lancaster.ac.uk
‡Smart Transport Research Centre
Faculty of Build Environment and Engineering,
Queensland University of Technology
2 George St GPO Box 2434, Brisbane QLD 4001, Australia

Abstract—Prediction of traffic flow variables such as traffic volume, travel speed or travel time for a short time horizon is of paramount importance in traffic control. Hence, the data assimilation process in traffic modeling for estimation and prediction plays a key role. However, the increasing complexity, non-linearity and presence of various uncertainties (both in the measured data and models) are important factors affecting the traffic state prediction. To overcome this problem, new methodologies have been proposed. With this aim, in this paper we propose the use of the Probability Hypothesis Density (PHD) filter for traffic estimation. This methodology is intensively studied, developed and improved for the purposes of multiple object tracking and consists in the recursive state estimation of several targets by using the information coming from an observation process. However, some issues need to be studied, especially the impact of the clutter (false alarm) intensity. The goal of this paper is to expose the potential of the PHD filters for real-time traffic state estimation and to investigate the choice of an appropriate clutter density which is usually chosen following a Poisson distribution. The dynamic evolution is based on a Cell Transmission Model (CTM) coupled with the PHD filter. It brings a novel tool to the state estimation problem and allows one to estimate the densities in traffic networks. In this work, we compare this PHD filter with the particle filter (PF) which has been successfully applied in traffic control and conclude that the PHD filter can be seen as a relevant alternative that opens new research avenues.

I. INTRODUCTION

Advanced estimation methods and sensor data fusion can improve the traffic conditions in Intelligent Transportation Systems (ITS) and provide information about special events. The fact that the full benefits of these systems cannot be realized without an ability to predict short-term traffic conditions has been emphasized by Sussman [20]. Hence, short-term prediction of traffic states appears as a key point in various ITS applications such as Advanced Traffic Management Systems (ATMS). In this context, prediction of traffic flow variables such as traffic volume, travel speed or travel time for a short time horizon (typically 5 up to 30 minutes) is of paramount importance. However, the increasing complexity, non-linearity

and presence of various uncertainties (both in the measured data and models) are important factors affecting the traffic state prediction. These make prediction methods based on deterministic assumptions unable to meet the accuracy needed in ITS applications.

To overcome this limitation, many non-deterministic predictive methods have been investigated [1]. In this scope, our research is focused on developing a stochastic traffic modelling framework that enables estimation and prediction of some information such as traffic flow with high accuracy. The purpose of this work is to show the potential of the Probability Hypothesis Density (PHD) filters for real-time traffic state estimation and to investigate the choice of an appropriate clutter density which is usually chosen following a Poisson distribution. The dynamic evolution is based on a Cell Transmission Model (CTM) coupled with the PHD filter as an estimation engine. The CTM is well known in traffic modelling, however the use of the PHD filter in traffic engineering is novel and it offers an alternative to estimate traffic network densities.

The remaining part of this paper is organised as follows. Section II presents the objectives of this work. Section III describes the main results with real traffic data. Finally, conclusions are summarized in Section V.

II. OBJECTIVES

For the problem of multiple-object tracking, the PHD filter has been intensively studied and used in the literature [11]. This problem consists in the recursive state estimation of several targets based on the information coming from an observation process. However, the PHD filter has not been applied yet to traffic estimation problems (such as to the travel time) except for one recent work [2].

Hence, the main objective is to investigate the potential of the PHD filters for real-time traffic state estimation,
integration of the effects of adversary weather. The PHD filter is implemented based on a Sequential Monte Carlo algorithm. Indeed, common traffic estimation issues can be tackled by the attractive properties of PHD filter, especially in dealing with clutter and multiple traffic flows (considering weather conditions). In most of the existing PHD filters, the clutter is modeled as a Poisson random finite set with a given intensity. The clutter intensity is characterized as a product of the average number of clutter (false alarm) points per scan and the probability density of clutter spatial distribution. This work applies the PHD filter jointly with a constant velocity model for estimation of trajectories. In contrast to [2], in which authors estimate road traffic intensity based on mobile vehicles’ coordinates, we develop a PHD filter for traffic networks, and combine the CTM with the PHD filter. Its performance is evaluated over a real-world study case.

III. METHODOLOGY

A. Macroscopic traffic model

In this study, the traffic model used is the well-known first order Lighthill-Whitham-Richards (LWR) model in its sending-receiving cell version of Daganzo [5], which is based on hydrodynamic analogy describing the behavior of the traffic flow. Let us recall it briefly in its discrete version, with application to a simple section (as shown in Figure 1). The motorway section is divided into \( n \) cells of length \( \Delta x \). The cell densities, \( k_i \) for cell \( i \), and flows, \( q_i \) for flow between cell \( i \) and \( i+1 \), are updated every \( \Delta t \) interval where the subscript \( N \) stands for numerical.

The complete model is composed of two equations. The first equation is the conservation equation. The second equation is the flow equation which consists of the demand-supply flows, i.e. the resulting flow in cell \( i \) at time \( t \) will be the minimum between cell \( i + 1 \) supply \( \Sigma_{i+1}(t) \) and the cell \( i \) demand \( \Gamma_i(t) \). Note that the Stochastic Cell Transmission Model (SCTM) developed by Sumalee et al., [18], extends the CTM to consider both supply and demand uncertainties. The numerical time step is taken as a sub-multiple of the CTM to consider both supply and demand uncertainties.

The state vector (of length \( 2n + 1 \)) to be estimated consists of the flows and densities in the \( n \) cells of the section:

\[
x_t = (k_1(t), \ldots, k_n(t), q_0(t), \ldots, q_n(t))^T,
\]

where \( T \) denotes the transpose operator.

The inputs \( u_t \) of the system are the demand upstream and the supply downstream of the considered section and other perturbations (incidents, work zones). The state equation is then written as follows:

\[
x_{t+1} = f(x_t, u_t),
\]

where \( f \) is a complex and highly nonlinear function with no straightforward analytical form.

The state equation (2) is completed by an (linear) observation equation, which maps measurements, \( y_t \), and the state vector of the system:

\[
y_t = Cx_t,
\]

where \( C \) is a real matrix consisting of rows whose elements are all zero except for the element corresponding to the position of the sensor delivering a measurement.

As there are uncertainties in both the measurements and the model, the complete dynamical model is given by

\[
\begin{align*}
x_t &= f(x_{t-1}, u_{t-1}) + w_t, \\
y_t &= Cx_t + v_t,
\end{align*}
\]

where \( w_t \) and \( v_t \) are Gaussian noises of null mean and variance-covariance matrix respectively \( Q_t \) and \( R_t \).

B. The particle PHD filter for traffic state estimation

1) Overview: The PHD filter is a multiple target filter for recursively estimating the number and the state of a set of targets given a set of observations. It works by propagating in time the first moment (called intensity function or PHD function) associated with the multi-target posterior [10]. Multi-target tracking is a common problem in many applications. The literature in this area is spread and the state-of-art and methodology are well summarized in [11]. Recently, one article has been concerned by traffic modeling issues [2], and the potential for real-time traffic state estimation has been investigated in Canaud et al., [4].

In [7] and [12] the authors highlighted the fact that the PHD filter outperforms the standard approaches such as the Kalman Filter (KF), Particle Filter (PF) or the Multiple Hypothesis tracking (MHT). One can find many implementations of the PHD filter either via the sequential Monte Carlo (SMC) method [21], [24] or using finite Gaussian mixtures (GM) [22], [23].

The GM method is attractive because it provides a closed form algebraic solution to the PHD filtering equation, with the state estimate (and its covariance) easily accomplished. However, the GM method is based on somewhat restrictive assumptions that single-object transitional densities and likelihood functions are Gaussian, and that the probability of survival and the probability of detection are homogenous ([11] [22]). The SMC method imposes no such restrictions and should therefore provide a more general framework for PHD filtering, even though it is also affected by different issues [11].

Beginning with Sidenbladh [16] and Zajic and Mahler [25], most researchers have implemented the PHD filter using SMC methods. SMC techniques were originally devised to approximate probability densities and the single-target Bayes filter.
Consequently, they must be modified for use with the PHD filter. This is because first: (i) the PHD is not a probability density function, and (ii) the PHD filter equations are more complex than the single-target Bayes filter equations.

Despite these differences, the SMC approximation carries over to the PHD filter relatively easily (with the exception of regularization). Particles represent random samples drawn from a posterior PHD. Particles are supposed to be more densely located where targets are most likely to be present. As with single-target SMC filtering, the basic idea is to propagate particles from time step to time step so that this assumption remains valid. The implementation is introduced in [21]. Then some improvements have been proposed in [13] and [15].

2) General formulation: The purpose of this section is to introduce the general formulation of a probability hypothesis density in the random finite set framework for multiple-target filtering. In this context targets and measurements are considered as random sets (random in values and also in number of values). From time step to time step, some of these targets may disappear. The surviving targets evolve to their new states and new targets may appear. Due to imperfections in the detector, some of the surviving and newborn objects may not be detected, whereas the observation set $Y_t$ may include false alarm detection.

The concept of PHD filter then requires to recursively estimate the set of states of all $n_t$ targets, $X_t = \{x_{t,1}, \ldots, x_{t,n_t}\}$ that are present at time $t$ given random measurements $Y_t = \{y_{k,1}, \ldots, y_{k,n_k}\}$ for all $k$ up to time $t$. In the next parts of the paper, we abusively use $x$ to denote the multi-target vector state, in order to simplify notation and without loss of generality since the context gives enough information about which vector is used. The underlying idea of the PHD filter is to propagate a suitable density function $D(x)$ in the target state space $\chi \subseteq \mathbb{R}^n$ ($n = \text{dim } x$). The PHD is a density function but not a probability density function, such that, for any region $S \subseteq \chi$, the expected number of targets in $S$ is given by

$$n(S) = \int_S D(s)dx,$$

i.e. by integration of $D(\cdot)$ over $S$.

Let’s $D(\cdot)$ be the PHD function and $D_{k|k}$ be the PHD function at time $k$ based on the set $Y_k$ of measurements till time instant $k$. The objective of PHD filter is time propagation

$$D_{k-1|k-1}(x) \rightarrow D_{k|k-1}(x) \rightarrow D_{k|k}(x).$$

Before presenting the predictor and corrector equations, we introduce some assumptions and notations. The PHD filter presumes some multi-target motion model. More precisely, target motions are statistically independent. Targets can disappear from the scene. New targets can be spawned by existing targets; and new targets can appear in the scene independently of existing targets. These possibilities are described as follows.

- **Motion of individual targets**: $f_{k+1|k}(x|x')$ is the single target Markov transition density.
- **Probability of surviving**: $p_{S,k+1|k}(x')$ noted $p_S(x')$ is the probability that a target with state $x'$ at time step $k$ will survive in time step $k+1$.
- **Spawning of new targets by existing targets**: $b_{k+1|k}(x|x')$ but as this phenomenon is negligible in our case, this assumption is not included in this formulation compared to the general PHD recursion [11].
- **Appearance of completely new targets**: $b_{k+1|k}(x)$ is a term characterising that new targets with state set $x$ will enter the scene at time step $k+1$.

The PHD filter also presumes the standard multitarget measurements model. More precisely, no target generates more than one measurement and each measurement is generated by no more than a single target, all measurements are conditionally independent of the target state, missed detections, and a multiobject Poisson false alarm process. These assumptions can be summarized as follows:

- **Single-target measurement generation**: $L_k(y|x)$ is the sensor likelihood function for observation $y$ at time step $k$ and state $x$.
- **Probability of detection**: $p_{D,k+1|k}(x')$ noted $p_D(x')$ is the probability that an observation will be collected at time step $k+1$ from a target with state $x$, if the sensor has state $x'$ at that time step.
- **False alarm density**: At time step $k+1$, the sensor collects an average number $\lambda = \lambda_k(x)$ of Poisson-distributed false alarms, the spatial distribution is governed by the probability density $c(y)$.

With these notations, we now describe the basic steps of the PHD filter: initialization, prediction and correction.

1) PHD Filter initialization

Regarding the initialization step of the PHD filter i.e. the choice of $D_{0|0}(x)$, no special recommendations are given in the literature. One could choose this PHD initialization as a sum of Gaussians. Mahler, [11], emphasizes that with no prior information about initial target position, an uniform PHD should be chosen.

2) PHD Filter predictor

At time step $k$, starting from $D_{k|k}$, one can derive a formula for the predictive $D_{k+1|k}$, which can be found in [10], defined as:

$$D_{k+1|k}(x) = b_{k+1|k}(x) + \int F_{k+1|k}(x|x')D_{k|k}(x')dx', \tag{7}$$

where the PHD “pseudo-Markov transition density” is

$$F_{k+1|k}(x|x') = p_{S}(x')f_{k+1|k}(x|x') + b_{k+1|k}(x|x'), \tag{8}$$

with the spawned target equal to zero in our study case. The predicted number of targets is therefore

$$N_{k+1|k} = \int D_{k+1|k}(x|y_k)dx, \tag{9}$$
then we have, under the non restrictive assumption of no spawning,

\[
D_{k+1|k}(x) = b_{k+1|k}(x) + \int p_S(x') f_{k+1|k}(x|x') D_{k|k}(x') dx'.
\]

(10)

3) **PHD Filter corrector**

From the previous step, one has the predicted PHD \( D_{k+1|k}(x) \), given by (10). At time step \( k+1 \), one collects a new observation set \( Y_{k+1} = \{y_1, \ldots, y_m\} \) and requires a formula for the data updated PHD \( D_{k+1|k+1}(x) \). According to Mahler, [10], the PHD corrector step is

\[
D_{k+1|k+1}(x) = L_{k+1}(y|x) D_{k+1|k}(x),
\]

(11)

where the “PHD pseudo-likelihood” function is defined by

\[
L_{k+1}(y|x) = 1 - p_D(x) + \frac{L_k(y|x)}{\sum_{y \in Y} \lambda_c(y) + \int p_D(x) L_k(y|x) D_{k+1|k}(x) dx}.
\]

(12)

Then we have,

\[
D_{k+1|k+1}(x) = [1 - p_D(x)] D_{k+1|k}(x) + \int p_D(x) L_k(y|x) D_{k+1|k}(x) dx.
\]

(13)

**C. Clutter intensity**

In Bayesian multi-target filtering knowledge of parameters such as clutter intensity and sensor field-of-view are of critical importance. Significant mismatches in clutter and sensor field of view model parameters result in biased estimates. In addition to the non-linearity, process and measurement noise, the two main sources of uncertainty which constitutes significant challenges in multi-target filtering are clutter and detection. Clutter is spurious measurements that do not belong to any target.

An unknown and non-homogeneous clutter intensity is accommodated in [9] by dropping the standard Poisson assumption for false alarms, and modelling individual clutter returns based on individual clutter targets or generators. Each clutter generator is analogous to an actual target, in the sense that clutter generators have their own separate models for births and deaths as well as transition, likelihood and detection or missed detection. However, the two types of clutter and actual targets are distinct, and cannot evolve into the other type. The intuition here is that the clutter generators will dynamically distribute themselves around the state space to explain the prevailing false alarm conditions.

However, since in traffic modeling and state estimation problem the use of the PHD filter is a new concept, the first step should be to investigate if some simple assumptions as Uniform or Poisson distribution is far away from the ground truth, or if it possible to choose an adaptive methodology, automatically estimating the clutter intensity ([17]).

**IV. EXPERIMENTAL RESULTS**

**A. Site and collected data**

The test site is the urban freeway located at the Eastern part of Lyon’s ring road (point A to D on figure 2), which consists of three lanes between km point B and km point C (5.6km long), [3], [14]. Traffic data were provided by the urban motorways’ operator CORALY and collected in 2007 from 8 loop sensors.

The upstream demand and the downstream supply and balances from a rainy day are used. The profiles of these external actions on the motorway system come from the real data of highway flow measurements in March 2007, on one weekday (Tuesday). The motorway section under consideration is the most frequently congested part. The upstream flow comprises the flow from North-West Lyon and the on-ramp from the Geneva freeway (point B). Therefore, the upstream demand presents high values on classical peak hours in the morning and at the end of the afternoon.

According to the section configuration, the space discretization of our traffic model has been carried out in such a way that the discretized cells are the segments between two consecutive sensors (Figure 2). Hence, we have 7 cells and 8 sensors. In this urban motorway section, three on and off-ramps are located on cells 3, 5 and 6. Therefore, in the traffic model, source terms have to be considered in these cells. The balance on versus offramp movements are shown in figure 3 as well as upstream demand.

**B. Hypothesis, models parameters and filter considerations**

From previous works, e.g. [14], the calibration of the fundamental diagram for this test site under rainy conditions (which is the case of the selected day), is achieved with the following parameters:
1) Critical density: \( k_c = 100 \text{ [veh/km]} \);
2) Maximum density: \( k_M = 300 \text{ [veh/km]} \);
3) Maximum flow: \( q_M = 138 \text{ [veh/min]} \).

The variance matrices needed in the estimation method also have to be selected. For the flow measurement uncertainty we have chosen a standard deviation of \( \sigma_R = 1.5 \text{ [veh/min]} \) (consistent with empirical analysis conducted on the raw data collected on this network). Hence, the noise variance matrix is then defined by: \( R_t = \text{diag}(\sigma_R^2) \). Regarding the uncertainty of the state equation, we assume the choice that the model is as robust as the measurements, meaning that the standard deviation of state vector flow part is the same as measurements one (\( \sigma_Q = 1.5 \text{ [veh/min]} \)). The density magnitudes are much smaller than those giving the flow; therefore the standard deviation of the density part of the state vector is chosen to be:

\[
\sigma_k = \sigma_Q \times \frac{k_c}{q_M} = 0.0011 \text{ [veh/m]}
\] (14)

Thus, the \( n \) first diagonal elements of the \( Q \) matrix are equal to \( \sigma_k^2 \) whereas the \((n+1)\) last ones are equal to \( \sigma_Q^2 \).

Regarding the filters, the particle PHD filter uses the following parameters:

1) Probability of detection: \( P_d = 0.98 \); 
2) Number of particles: \( N_p = 400 \); 
3) Number of birth particles: \( N_b = 250 \); 
4) Number of clutter (false alarm): \( N_c = 1 \).

We assume for the particle initialization a “Uniform PHD”. Moreover, as the newborn object particles need to cover the entire state-space with reasonable density for the SMC-PHD filter in order to work properly, [13], the birth density is driven by measurements following a uniform distribution. Further, it is supposed that, at each time instant \( t \), on average \( N_c \) clutter measurements are generated with an uniform distribution in the measurement space.

C. Performance evaluation

In order to evaluate performance of PHD filter for real-time traffic state estimation, the proposed methodology has been carried out. The assimilation was performed by two methodologies: (i) Particle filter and (ii) SMC-PHD filter, both under the same conditions (CPU time, filter and model parameters). Finally, we compare the flow estimates obtained from both filters and for each sensor location.

The simulation has been conducted as follows. First, it is supposed that measurements at some boundary cells are available. Then, it is assumed that the upstream demand and downstream supply are known with a certain probability of detection and false alarms, as well as the ramp balances of in and out flows. Then we estimate by the two proposed methodologies the state vector of densities and flows. For comparison purposes, estimated versus actual traffic flows have been depicted for both estimation engines (see figure 4). Root Mean Square Error (RMSE) is used as a measure of the performance. Namely, for each time step and each sensor, we compute:

\[
RMSE_i = \frac{1}{N_s N_t} \sum_{j=1}^{N_s} \sum_{k=1}^{N_t} \left( \hat{y}_{i,j}(t_k) - y_{j}(t_k) \right)^2,
\] (15)

where the subscript \( i \) = PF or PHD, \( N_s \) is the number of sensors, \( N_t \) the number of time step and \( \hat{y}_{i,j}(t_k) \) the traffic flow estimates for \( i \), sensor \( j \) at time \( t_k \).

Since noises are unknown with real data, first conclusions about the two methodologies could be subject to criticism. For example, if measurements if measurements are far from the real conditions, the less the RMSE, the worst the method. To prevent this, and following the work of Canaud et al., [4], we choose to perform the study on simulated data instead of real data. The aim is that when we form the difference \( \hat{y} - y \), we can easily see the error in \% and then see the performance for all filters with respect to the ground truth data.

For setting up the simulation scenario, the typical days upstream demand and downstream supply and balances of ramp movements introduce before were used. In other word, simulated data were derived from the real one. The model used to simulate data is the same as the one used in the filters in which the level of noises of the simulated data is defined by the diagonal of \( Q \). Results are presented in table I.

As one can see, with only boundaries conditions, both estimates are quite similar, and close to the actual measurements with a noticeable PHD overperformance: \( RMSE_{PF} = 3.44 \) and \( RMSE_{SMC-PHD} = 3.29 \). However, PHD estimates appear smoother that those given by PF, which can be an interesting feature if we look at the confidence interval of estimates.
Thus, PHD filter performed better than PF. Finally, focusing on the transition regime (from free flow to congestion and conversely), PHD detects changes in traffic conditions better.

V. CONCLUSIONS

This paper develops a PHD filter for traffic state estimation with an appropriate clutter intensity. It is implemented based on a SMC algorithm. Its performance is evaluated over a real-world study case. The results are compared with the generic PF, well recognized to solve traffic state estimation problems.

The results show the potential of the PHD filter to be applied to traffic state estimation and lay the cornerstone for the clutter intensity feature. The results demonstrate that the PHD filter outperforms the particle filter on this real study case. Secondly, it opens new avenues for PHD filter research in traffic control, showing useful potential. The PHD filter can be especially powerful in dealing with multiple traffic flows. Specifically, this work can be extended with more complex models to adversary weather conditions (which means more clutter in the data), including switching state space models as developed in [3] or [19] in order to adapt the traffic state estimation according to weather or traffic conditions. Multiple data sources can also be considered since the PHD filter is relevant in multi-target multi-source tracking. One can consider urban traffic and the estimation of other information such as travel time under various noise uncertainties.

In conclusion, this first step in using the PHD filter for traffic estimation issues shows the potential use of such a tool. Some features have still to be investigated now, like the likelihood function, the calibration of some parameters: rate, resampling improvement, detection profile. The question of clutter intensity has to be reconsidered in order to automatically adapt itself. However, this filter has a great interest, with no doubt, and definitely needs more consideration for future traffic applications.

ACKNOWLEDGMENTS

This work is part of COST action TU0702 research activities: Real-time Monitoring, Surveillance and Control of Road Networks under Adverse Weather Conditions [6]. The authors would like to thank CORALY (Lyons urban motorways managers) for providing them with real traffic data.

REFERENCES


