Multi-target tracking using tractable approximations of optimal multi-Bayes filters

Michele Pace
Equipe ALEA

26th November 2009
ALEA Working group
Is the problem of estimating the hidden state of a system as a set of observations becomes available in time.

Applications

- Ballistics
- Robotics (localization)
- Visual Tracking hands/cars/people
- Econometrics
- Navigation
- Many more...
Filtering

Is the problem of estimating the hidden state of a system as a set of observations becomes available in time.

Applications

- Ballistics
- Robotics (localization)
  - Visual Tracking hands/cars/people
  - Econometrics
  - Navigation
  - Many more...
Filtering

Is the problem of estimating the hidden state of a system as a set of observations becomes available in time.

Applications

- Ballistics
- Robotics (localization)
- Visual Tracking hands/cars/people
- Econometrics
- Navigation
- Many more...
Filtering

Is the problem of estimating the hidden state of a system as a set of observations becomes available in time.

Applications

- Ballistics
- Robotics (localization)
- Visual Tracking hands/cars/people
- Econometrics
- Navigation
- Many more...
Filtering

Is the problem of estimating the hidden state of a system as a set of observations becomes available in time.

Applications

- Ballistics
- Robotics (localization)
- Visual Tracking hands/cars/people
- Econometrics
- Navigation
- Many more...
Filtering

Is the problem of estimating the hidden state of a system as a set of observations becomes available in time.

Applications

- Ballistics
- Robotics (localization)
- Visual Tracking hands/cars/people
- Econometrics
- Navigation
- Many more...
## Difficulties

- Noise on measurements and model uncertainty
- Full/partial occlusions
- Targets entering/leaving the scene
- False positives/false negatives
- Efficiency
- Multiple models and switching dynamics, multiple targets
- Targets overlapping
- **Many more..**
Single Target tracking

- The state vector contains all available information to describe the investigated system.
- Observations are generally of lower dimension than the state vector.

System evolution:

| System evolution | \( x_t = f_{t|t-1}(x_{t-1}, v_t) \) |
|------------------|-------------------------------------|
| Observation function | \( y_t = h_t(x_t, w_t) \) |
| Objective:       | \( p_t(x_t | y_{1:t}) \) |

Fig: [Vo 2009]
The estimation of $p_t(x_t|y_{1:t})$ is can be recursevely obtained using the Chapman-Kolgomorov equation and the Bayes rule. The Bayes filter is a recursion that consists of two steps: prediction and update.

### Prediction

$$p_{t|t-1}(x_t|y_{1:t-1}) = \int f_{t|t-1}(x_t|x)p_{t-1}(x|y_{1:t-1})dx$$  \hspace{1cm} (1)$$

### Update

$$p_t(x_t|y_{1:t}) = \frac{p_{t|t-1}(x_t|y_{1:t-1})p(y_t|x_t)}{p(y_t|y_{1:t-1})}$$  \hspace{1cm} (2)$$
Multi target filter problem

Objective

Estimate an unknown, time varying number of targets and their states from noisy observations available at discrete intervals of time.

- The number of target changes over time
- Detection uncertainty
- Clutter
- Association uncertainty
The first approaches treated multi-target tracking as two separate problems:

1. Associate the new measurement to a list of current tracks, (the state vector is augmented with association variables)
2. Estimate the state of the targets based on this association. (Generally using Kalman filter, EKF or UKF)

Mahler proposed a different approach based on Point Processes. In this framework the association problem generally doesn’t generate a combinatorial explosion.

**The Model**

The set of targets at time $t$, the set of measurements and the clutter process are modeled as point processes.
Notations

Given a state space $\mathcal{X}$, $\mathcal{F}(\mathcal{X})$ the set of all finite subsets of the state space $\mathcal{X}$. The set of targets at time $t$: $X_t = \{X_{t,1}, \ldots, X_{t,K_t}\} \in \mathcal{F}(\mathcal{X})$ is defined as:

$$X_t = \left(\bigcup_{x \in X_{t-1}} S_{t|t-1}(x)\right) \cup \left(\bigcup_{x \in X_{t-1}} D_{t|t-1}(x)\right) \cup N_t$$

Where:
- $S_{t|t-1}(x)$ is the random set of targets which survived at time $t$ from the set $x \in X_{t-1}$
- $D_{t|t-1}(x)$ is the random set of targets which have been generated from $x \in X_{t-1}$
- $N_t$ is the random set of targets appeared at time $t$.

The set of measures: $Y_t = \{Y_1, \ldots, Y_{N_t}\}$ at time $t$ is given by the following observation model:

$$Y_t = \Gamma_t \cup \left(\bigcup_{x \in X_t} G_t(x)\right)$$

where $G_t(x)$ is the random set of measures generated from the target $x \in X_t$ and $\Gamma_t$ the set of spurious measurements at time $t$. 
A random set $X \in \mathcal{F}(\mathcal{X})$ can equivalently be represented by the random measure $N_X$ defined by

$$N_X(S) = \sum_{x \in X} 1_S(x) = |X \cap S|$$

where $1_S(x) = 1$ if $x \in S$, 0 otherwise, and $|A|$ is the number of elements in the set $A$. The first moment, or intensity measure of $X \in \mathcal{F}(\mathcal{X})$ is defined by

$$V(S) = \mathbb{E}[N(S)]$$

The intensity $V(S)$ over a region $S$ gives the number of elements of $X$ which are in $S$. The density $\nu : \mathcal{X} \rightarrow [0, +\infty]$ of the intensity measure $V$ with respect to the Lebesgue measure is called intensity function, “Probability Hypothesis Density" or PHD.

$$\int_S \nu(x)dx = \mathbb{E}[N(S)] = \mathbb{E}[|X \cap S|]$$

The maxima of $\nu$ correspond to the points in $X$ where there is the highest local concentration of the random number of targets. The total mass of $V(\mathcal{X})$ gives the total average number of targets.
Intensity function

\[ \nu_\Sigma(x_0) \]

PHD (intensity function) of a RFS \( \Sigma \)

\[ \int_S \nu_\Sigma(x) \, dx = \text{expected number of objects in } S \]
**PHD Filter**

**Structure**

**Mahler**: "The PHD filter is a multitarget statistical analog of the computationally fastest approximate single-target filtering approach (constant-gain Kalman filter). This filter propagates a first-order statistical moment in the place of the multitarget posterior distribution”.

- Each object evolves and generates observation independently of one another
- The measurements and clutter RFS are Poisson RFS
- The birth RFS is a Poisson RFS
- The PHD (or intensity function) $\nu_t$ is not a probability density and the PHD propagation equation is not a standard Bayesian recursion

The predicted and the posterior multi-object RFS are approximated by Poisson RFS.

The PHD filter propagates the probability hypothesis density function, that is the first moment of the target posterior in two steps, a prediction step and an update step:

$$\nu_{k|k}(x) = (\Psi_t \circ \Phi_{t|t-1}) \nu_{k-1|k-1}(x)$$

(3)
Basic properties of Poisson point processes

**Definition**

- **Superposition**: The sum of independent Poisson processes with intensities \( \lambda_1 \) and \( \lambda_2 \) is a Poisson process with intensity \( \lambda = \lambda_1 + \lambda_2 \).

- **Thinning**: If each point \( x \) survives with probability \( 0 \leq \pi(x) \leq 1 \) then the probability of survival of \( x \) is a Poisson process with intensity \( \lambda_{\text{th}} = \lambda(x)\pi(x) \).

A Poisson RFS is completely characterized by its intensity function \( \nu \).

**Property**

The Poisson RFSs:

- are closed under **superposition** and **independent thinning**
- the distribution of the cardinality or \( X \) is Poisson with mean \( N = \int \nu(x)dx \)
- given a cardinality of \( N \), the elements of \( X \) are i.i.d. with probability \( \nu(.)/N \)
- A Poisson RFS is completely characterized by its intensity
The PHD filter propagates in time the posterior intensity function:

**PHD Prediction**

\[
\nu_{t|t-1}(x) = \int_{X} p_{S,t} f_{t|t-1}(x|u) \nu_{t-1|t-1}(u) du + \gamma_t(x)
\]  

(4)

where \(f_{t|t-1}(x|u)\) is the evolution density of a target, \(p_{S,t}\) the survival probability at time \(t\) and \(\gamma_t(x)\) the intensity of new targets.

**PHD Update**

\[
\nu_{t|t}(x) = (1 - p_{D,t}(x)) \nu_{t|t-1}(x) + \sum_{y \in Y_t} \frac{p_{D,t}(x) h_t(y|x) \nu_{t|t-1}(x)}{\kappa_t(y) + \int_{X} p_{D,t}(u) h_t(y|u) \nu_{t|t-1}(u) du}
\]  

(5)

where \(h_t(y|x)\) is the likelihood of an observation, \(p_{D,t}(x)\) the detection probability and \(\kappa_t\) the clutter intensity. This recursion is based on the assumption that the point processes \(X_{t|t}\) and \(X_{t|t-1}\) can be approached by Poisson Point processes.
Description of the model

Denote by $\gamma_n$ and $\eta_n$ the flow of measure defined by the equation:

$$\gamma_n = \gamma_{n-1} Q_n + \mu_n, \text{ and } \eta_n(dx_n) = \gamma_n(dx_n)/\gamma_n$$

In this situation $\gamma_n$ represents the intensity measure associated with a multi-target type branching process associated to the integral operators:

$$Q_n(x_{n-1}, dx_n) = B_n(x_{n-1}, dx_n) + e_n(x_{n-1})P_n(x_{n-1}, dx_n)$$

One traditional way to enter the likelihood of a given observation is to multiply at each time step $\gamma_n$ by a likelihood function which depend on the observation RFS.

$$\hat{G}_{n, \gamma_n}(x) = (1 - d_n(x)) + d_n(x) \sum_{y \in Y_n} \frac{g_{n,y}(x)}{k_n(x) + \gamma_n(d_n g_{n,y})}$$

The resulting evolution equation takes the following form

$$\gamma_n := \gamma_n \tilde{Q}_{n+1, \gamma_n} + \mu_{n+1}$$

With the collection of integral operators defined by:

$$\tilde{Q}_{n+1, \gamma_n}(x_n, dx_{n+1}) = \hat{G}_{n, \gamma_n}(x_n) Q_{n+1}(x_n, dx_{n+1})$$
Possible implementations and variants

**SMC-PHD**
The intensity function is approximated using a set of weighted particles. After the update particles need to be clustered to identify targets position.

**GM-PHD**
Closed-form solution to the PHD recursion exists for linear Gaussian multi-target model. Prior intensity is modeled as Gaussian mixture as well as the posterior intensities.

**Cardinalised PHD Filter [Mahler 06,07]**
Jointly propagate intensity function and probability generating function of cardinality

- More complex PHD update step (higher computational costs)
Algorithm 1 Filtre PHD particulaire

Initialization: At time $t = 0$, initialize the particles

- For $i = 1, \ldots, L_0$, sample $x_0^{(i)} \sim q_0$
- For $i = 1, \ldots, L_0$, set $w_0^{(i)} = N_0/L_0$

where $N_0 = \int \nu_0(x) dx$ and $q_0(x) = \nu_0(x)/N_0$

Iterate: For $t = 1, 2, \ldots$

- Exploration/Mutation
  - For $i = 1, \ldots, L_{t-1} + J_t$, sample

$$
\tilde{x}_t^{(i)} \sim \begin{cases} 
q_t(\cdot|x_{t-1}^{(i)}, Y_t) & i = 1, \ldots, L_{t-1} \\
r_t(\cdot|Y_t) & i = L_{t-1} + 1, \ldots, L_{t-1} + J_t
\end{cases}
$$

(11)
Algorithm 2 Filtre PHD particulaire

- **Selection**
  - For $i = 1, \ldots, L_{t-1} + J_t$:

    $$\tilde{w}_{t|t-1}^{(i)} = \begin{cases} 
      \frac{f_{t|t-1}(\tilde{x}_{t-1}^{(i)} | x_{t-1}^{(i)})}{q_{t}(\tilde{x}_{t}^{(i)} | x_{t-1}^{(i)}, Y_t)} w_{t-1}^{(i)} & i = 1, \ldots, L_{t-1} \\
      \frac{\gamma_t(\tilde{x}_{t}^{(i)})}{J_t r_t(\tilde{x}_{t}^{(i)} | Y_t)} & i = L_{t-1} + 1, \ldots, L_{t-1} + J_t
    \end{cases}$$

  (12)

- Resample the particles according to $\tilde{w}_{t|t-1}^{(i)}$ to obtain $L_t = \rho \tilde{N}_t$ particles $(x_t(i))$ with weights $w_t^{(i)} = \frac{1}{\rho}$ where $\tilde{N}_t = \sum_i \tilde{w}_t^{(i)}$ is the estimated number of targets.
An unknown, varying number of targets move along the line segment $[-100,100]$. The state of the targets consist of position and velocity; only the position is observed. Targets may appear or disappear at any time during and are subjected to random accelerations.

$$
\begin{bmatrix}
    x_{t+1} \\
    v_{t+1}
\end{bmatrix} =
\begin{bmatrix}
    1 & \Delta t \\
    0 & 1
\end{bmatrix}
\begin{bmatrix}
    x_t \\
    v_t
\end{bmatrix} + a_t
\begin{bmatrix}
    \frac{1}{2} \Delta t^2 \\
    \Delta t
\end{bmatrix}
$$

(14)

Where $a_t$ is sampled from $\mathcal{N}(0, \sigma_w^2)$

The system evolution is partially observed by the following observation model:

$$
y_t =
\begin{bmatrix}
    1 & 0
\end{bmatrix}
\begin{bmatrix}
    x_t \\
    v_t
\end{bmatrix} + \sigma_v V_t
$$

(15)

Where $V_t$ is sampled from $\mathcal{N}(0, \sigma_v^2)$
Simulation 1

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \rho )</td>
<td>100</td>
</tr>
<tr>
<td>Birth particles</td>
<td>25</td>
</tr>
<tr>
<td>Avg. Targets time step</td>
<td>0.05</td>
</tr>
<tr>
<td>( P_D )</td>
<td>0.9</td>
</tr>
<tr>
<td>( P_E )</td>
<td>1</td>
</tr>
<tr>
<td>( \lambda_k )</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure: Scenario
Simulation 1

Figure: Filtering result
Simulation 1

**Figure:** Intensity function
Simulation 1

\[ \rho = 100 \quad \text{Birth particles} = 25 \]

<table>
<thead>
<tr>
<th>Avg. Targets time step= 0.05</th>
<th>( P_E = 1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P_D = 0.9 )</td>
<td>( \lambda_k = 5 )</td>
</tr>
</tbody>
</table>
Simulation 1

Figure: Filtering result
Simulation 1

Figure: Intensity function
Simulation 1

<table>
<thead>
<tr>
<th>$\rho = 100$</th>
<th>Birth particles = 25</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avg. Targets time step= 0.05</td>
<td>$P_E = 1$</td>
</tr>
<tr>
<td>$P_D = 1$</td>
<td>$\lambda_k = 2$</td>
</tr>
</tbody>
</table>

Figure: Scenario
Simulation 1

Figure: Filtering result
Simulation 1

\[ \rho = 100 \quad \text{Birth particles} = 25 \]

<table>
<thead>
<tr>
<th>Avg. Targets time step = 0.05</th>
<th>( P_E = 1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P_D = 0.7 )</td>
<td>( \lambda_k = 2 )</td>
</tr>
</tbody>
</table>
Simulation 1

![Filtering result](image)

**Figure:** Filtering result
Simulation 1

\[ \rho = 100 \quad \text{Birth particles} = 25 \]

<table>
<thead>
<tr>
<th>Avg. Targets time step= 0.05</th>
<th>( P_E = 1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P_D = 0.5 )</td>
<td>( \lambda_k = 2 )</td>
</tr>
</tbody>
</table>

Figure: Scenario
Simulation 1

Figure: Filtering result
Simulation 1

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \rho )</td>
<td>100</td>
</tr>
<tr>
<td>Avg. Targets time step</td>
<td>0.05</td>
</tr>
<tr>
<td>( P_E )</td>
<td>1</td>
</tr>
<tr>
<td>( P_D )</td>
<td>0.2</td>
</tr>
<tr>
<td>( \lambda_k )</td>
<td>0</td>
</tr>
<tr>
<td>Birth particles</td>
<td>25</td>
</tr>
</tbody>
</table>
Simulation 1

Figure: Filtering result
Simulation 1

- $\rho = 100$
- Birth particles = 25
- Avg. Targets time step= 0.05
- $P_E = 1$
- $P_D = 1$
- $\lambda_k = 2$

![Observation and clutter](image)
Simulation 1

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho$</td>
<td>100</td>
</tr>
<tr>
<td>Birth particles</td>
<td>25</td>
</tr>
<tr>
<td>Avg. Targets time step</td>
<td>0.05</td>
</tr>
<tr>
<td>$P_E$</td>
<td>1</td>
</tr>
<tr>
<td>$P_D$</td>
<td>1</td>
</tr>
<tr>
<td>$\lambda_k$</td>
<td>2</td>
</tr>
</tbody>
</table>
### What’s going on

<table>
<thead>
<tr>
<th>Analysis of the conditional distributions of spatial point processes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Techniques developed to complement and simplify more traditional random finite sets analysis involving unnecessary symmetrisation techniques or related to other technicalities associated with moment generating functions derivatives. <em>On the Conditional Distributions of Spatial Point Processes</em></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Particle approximations of branching distribution flows</th>
</tr>
</thead>
<tbody>
<tr>
<td>Design a mean field and interacting particle interpretation of a class of spatial branching intensity models with spontaneous births. In contrast with traditional Feynman-Kac type particle models, the transitions of these interacting particle systems depend on the current particle approximation of the total mass process. Analysis of the stability properties and long time behavior of these distribution flows.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Application and Extension of PHD filter to realistic case scenarios</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Probability Hypothesis Density filter is applied to realistic three-dimensional aerial and naval scenarios. A comparison between the sequential Monte Carlo and the Gaussian Mixture approximation is given using different scenarios and different clutter levels.</td>
</tr>
</tbody>
</table>
Description of the model

$(E_n)_n \geq 0$ sequence of measurable spaces equipped with some $\sigma$-fields $(E_n)_{n \geq 0}$.\n\n$M(E_n)$, $M_+(E_n)$ and $P(E_n)$ be the set of all finite signed measures, the subset of positive measures and the subset of probability measures over $E_n$, with $n \geq 0$.\n\n$\mu_n \in M_+(E_n)$ a collection of measures, $M_n$ a collection of Markov transitions from $E_n$ to $E_{n+1}$.\n\nWe denote by $\gamma_n$ and $\eta_n$ the flow of measure defined by the equation:

$$\gamma_n = \gamma_{n-1} Q_n + \mu_n, \text{ and } \eta_n(dx_n) = \gamma_n(dx_n)/\gamma_n(1)$$

(16)

and some given initial measure $\gamma_0 = \mu_0 \in M_+(E_n)$\n\nThe pair process $(\gamma_n(1), \eta_n) \in (R_+, P(E_n))$ satisfy the following non linear evolution:

$$(\gamma_n(1), \eta_n) = \Gamma_n(\gamma_{n-1}(1), \eta_{n-1})$$

(17)

Where $\Gamma_n^1$ and $\Gamma_n^2$ are the first and the second component mappings from $(R_+ \times P(E_n))$ into $R_+$, and from $(R_+ \times P(E_n))$ into $P(E_n)$. 
Description of the model $\mu_{n+1}(1) = 0$

For null spontaneous branching measures $\mu_n = 0$:

$$
\gamma_n(1) = \eta_{n-1}(G_n) \gamma_{n-1}(1) := \Gamma^1_n(\gamma_{n-1}(1), \eta_{n-1})
$$

$$
\eta_n(dx') = \int \psi_{G_{n-1}}(\eta_{n-1})(dx) M_n(x, dx') := \Gamma^2_n(\gamma_{n-1}(1), \eta_{n-1})(dx')
$$

The second component mapping $\Gamma^2_n(\gamma_{n-1}(1), \eta_{n-1}) := \Phi_n(\eta_{n-1})$ reduces to a mapping $\Phi_n$ that doesn’t depend on the total mass process $\gamma_{n-1}(1)$. 
Markov transport formulation \( \mu_{n+1}(1) \neq 0 \)

Recall

\[
(\gamma_n(1), \eta_n) = \Gamma_n(\gamma_{n-1}(1), \eta_{n-1})
\]

For any \( n \geq 0 \), we have the recursive formula

\[
\begin{align*}
\gamma_{n+1}(1) &= \gamma_n(1) \eta_n(G_n) + \mu_{n+1}(1) \\
\eta_{n+1} &= \Psi_{G_n}(\eta_n)M_{n+1,(\gamma_n(1),\eta_n)}
\end{align*}
\]

(18)

with the collection of Markov transitions \( M_{n+1,(m,\eta)} \) indexed by the parameters \( m \in \mathbb{R}_+ \) and the probability measures \( \eta \in \mathcal{P}(E_n) \) given below

\[
M_{n+1,(m,\eta)}(x, dy) := \alpha_n(m, \eta) M_{n+1}(x, dy) + (1 - \alpha_n(m, \eta)) \tilde{\mu}_{n+1}(dy)
\]

(19)

with the collection of \([0, 1]\)-parameters \( \alpha_n(m, \eta) \) defined below

\[
\alpha_n(m, \eta) = \frac{m\eta(G_n)}{m\eta(G_n) + \mu_{n+1}(1)}
\]