Markov Chain Monte Carlo Data Association for Target Tracking

Niclas Bergman  
Division of Automatic Control  
Dept. of Electrical Engineering  
Linkping University  
SE-58183 Linkping, Sweden  
niclas@isy.liu.se

Arnaud Doucet  
Signal Processing Group  
Department of Engineering  
University of Cambridge  
Trumpington St. CB2 1PZ  
Cambridge, UK  
ad2@eng.cam.ac.uk
The multiple measurement data association problem arises in over the horizon target tracking.

Data association between several measurements and several measurement models of the same target.
Consider a linear dynamical target state model

\[ x_{t+1} = F_t x_t + w_t, \quad t = 0, 1, \ldots \]  

At time \( t \) a set of \( \mu_t \geq 0 \) measurements \( Y_t = \{y_i(t)\}_{i=1}^{\mu_t} \) are detected, where each measurement either originate from one of \( n \) known linear measurement models or is a false detection. Explicitly, we model the received measurement as

\[ y_i(t) = \begin{cases} 
H_1 x_t + e_1(t) & \text{for model 1} \\
\vdots & \vdots \\
H_n x_t + e_n(t) & \text{for model } n \\
\text{clutter} & \text{otherwise}
\end{cases} \]  

where \( y_i(t) \in \mathbb{R}^{n_y} \) and \( e_i(t) \) is a white, zero mean Gaussian sequence with known covariance \( R_i \) of full rank.
Each measurement model in (2) produces a measurement with a probability $P_D$, which for simplicity is assumed identical for all $n$ models. The number of clutter detections in each measurement set is Poisson distributed

$$p_c(q) = \frac{(\lambda V_S)^q}{q!} \exp(-\lambda V_S) \quad q = 0, 1, \ldots$$

and the actual clutter measurements are uniformly distributed in the observation volume, above labeled $V_S$.

The number of possible association hypotheses for a given number of $m_t$ target measurements follows by combinatorics as

$$r_t(m_t) = \frac{n! \mu_t!}{m_t!(n-m_t)!(\mu_t-m_t)!} \quad m_t = 0, 1, \ldots, \nu_t$$

where $\nu_t \triangleq \min\{\mu_t, n\}$. The total number of association hypotheses for the measurement set received at time $t$ is thus $\pi_t = \sum_{m_t=0}^{\nu_t} r_t(m_t)$ which will grow inhibitiley fast with $\mu_t$ and is the main challenge in this application.
Markov Chain Algorithm

Given an observed data batch $Y_N = \{Y(t)\}_{t=0}^{N}$, we consider computing both MMSE estimates and MMAP estimates of the sequence of association events $\Psi_N = \{\psi_t\}_{t=0}^{N}$ and the state sequence $X_N = \{x_t\}_{t=0}^{N}$.

Algorithm structure:
1. Pick the initial association $\Psi_N^{(0)} = \{\psi_t^{(0)}\}_{t=0}^{N}$ randomly or deterministically. Set $p = 1$.
2. For each $t = 0, 1, \ldots, N$ generate a random sample from the full conditional distribution

$$
\psi_t^{(p)} \sim p(\psi_t^{(p)}|Y_N, \Psi_{-t}^{(p)})
$$

(3)


The chosen association event at time $t$ in pass $p$ is denoted $\psi_t^{(p)}$ and $\Psi_{-t}^{(p)} = \{\psi_{t-1}^{(p)}, \psi_{t+1:N}^{(p)}\} = \{\psi_0^{(p)}, \ldots, \psi_{t-1}^{(p)}, \psi_{t+1}^{(p-1)}, \ldots, \psi_{N}^{(p-1)}\}$
The full conditional distribution (3) is discrete with $\pi_t$ possible association events at each time instant, and given by

$$p(\psi_t(p)|Y_{-t}, \psi_{-t}(p)) = \frac{p(Y_t|\psi_t(p), \psi_{-t}(p))\Pr(\psi_t(p))}{p(Y_t|\psi_{-t}(p))} \propto p(Y_t|\psi_t(p), \psi_{-t}(p))\Pr(\psi_t(p))$$

$$\propto p(Y_t|Y_{t-1}, \Psi_t(p))\Pr(\psi_t(p)) \int p(Y_{t+1:N}|\Psi_{t+1:N}, x_t)p(x_t|\Psi_t(p), Y_t)dx_t$$

In the paper we show that (3) is given by

$$p(\psi_t(p)|Y_N, \Psi_{-t}(p)) \propto \mathcal{N}(Y_t; \hat{Y}_{t|t-1}, S_t)\frac{p_c(\mu_t - m_{t,p})P_{D}^{m_{t,p}}(1 - P_{D})^{n-m_{t,p}}V_{S}^{m_{t,p}}}{r_t(m_{t,p})}$$

$$|I + P_{t|t}\Gamma_{t|t+1}|^{-1/2}\exp\left(-\frac{1}{2}\|\hat{x}_{t|t}\|^{2}_{\Gamma_{t|t+1}} + \hat{x}_{t|t}\hat{a}_{t|t+1} + \frac{1}{2}\|\hat{a}_{t|t+1} - \Gamma_{t|t+1}\hat{x}_{t|t}\|^{2}_{(\Gamma_{t|t+1} + P_{t|t}^{-1})^{-1}}\right), \ (4)$$

where $\hat{Y}_{t|t-1}$, $\hat{x}_{t|t}$, and $P_{t|t}$ are based on $\psi_t(p)$ and $\|x\|^{2}_{A} = x^TAx$. 
The full conditional probabilities are computed using a backwards information filter under the association hypothesis \( \psi_N^{(p-1)} \)

\[
\hat{a}_{t|t} = \hat{a}_{t|t+1} + H_t^T R_t^{-1} Y_t \\
\Gamma_{t|t} = \Gamma_{t|t+1} + H_t^T R_t^{-1} H_t \\
\hat{a}_{t-1|t} = F_t^T \left( I - \Gamma_{t|t} \left( \Gamma_{t|t} + Q_t^{-1} \right)^{-1} \right) \hat{a}_{t|t} \\
\Gamma_{t-1|t} = F_t^T \left( I - \Gamma_{t|t} \left( \Gamma_{t|t} + Q_t^{-1} \right)^{-1} \right) \Gamma_{t|t} F_t
\]

Where \( Y_t \) in (5a) is the stacked vector of non-clutter measurements under the hypothesis \( \psi_t^{(p-1)} \), and \( H_t \) and \( R_t \) are the corresponding measurement equation matrix and measurement error covariance, respectively.
Algorithm 1 (Monte Carlo Data Association, MCDA)

1. Pick the initial association $\Psi^{(0)}_N$. Set $p = 1$.
2. Compute (5) using $\Psi^{(p-1)}_N$. Store $\Gamma_{t|t+1}$ and $\hat{a}_{t|t+1}$.
3. Initialize the Kalman recursion $\hat{x}_{0|−1}$ and $P_{0|−1}$, set $t = 0$.
   (a) Run measurement updates under each hypothesis for $Y_t$, i.e., for $k = 1, \ldots, \pi_t$. Store every $\hat{x}^k_{t|t}$ and $P^k_{t|t}$.
   (b) Compute (4) under all $\pi_t$ hypotheses, normalize this discrete density, and generate a sample $\psi^{(p)}_t$.
   (c) Set $t := t + 1$, run a time update step with the Kalman filter corresponding to $\psi^{(p)}_t$, store the resulting $\hat{x}_{t+1|t}$ and $P_{t+1|t}$ and return to 3(a). If $t > N$, go to 4.
4. Compute $X^{(p)}_N = \mathbb{E}\left(X_N|Y_N, \Psi^{(p)}_N\right)$.
Instead of sampling from the full conditional in item 3(b) above, MMAP estimates can be formed by deterministically picking the most likely association, according to the Iterated Conditional Mode (ICM). The resulting ICMDA algorithm is iterated until two subsequent passes yields the same association sequence.
Previous Approaches

**Multiple Simultaneous Measurement Filter, MSMF** The MSMF replaces the Gaussian mixture from all hypotheses by a single Gaussian density at each algorithm iteration by matching the first and the second moments.

\[
\begin{align*}
E(x_t|Y_t) & \approx \hat{x}_t|t = \sum_{i=1}^{\pi_t} c_t^{(i)} \hat{x}_{t|i}\t
\end{align*}
\]

\[
Cov(x_t|Y_t) \approx P_{t|t} = \sum_{i=1}^{\pi_t} c_t^{(i)} (P_{t|i} + \hat{x}_{t|i}\hat{x}_{t|i}^T) - \hat{x}_{t|t}\hat{x}_{t|t}^T
\]

where

\[
c_t^{(i)} \propto \frac{pc(\mu_t - m_{t,i})P_D^{m_{t,i}}(1 - P_D)^{n-m_{t,i}}}{r_t(m_{t,i})} \prod_{(j,k) \in \psi_{t,i}} N(y_j(t); H_k\hat{x}_{t|t-1}, H_kP_{t|t-1}H_k^T + R_k)
\]
Expectation Maximization Data Association, EMDA computes MMAP estimates of the association events by treating the target state sequence $X_N$ as missing (unobserved) data and the sequence of association events $\Psi_N$ as the sought parameters using the EM-algorithm

$$\hat{\Psi}^{(p+1)}_N = \arg \max_{\Psi_N} E \left( \ln p(X_N, Y_N, \Psi_N) | Y_N, \hat{\Psi}^{(p)}_N \right)$$

Defining a cost function $J_p(\psi_t)$ which can be maximized independently for each $t$ we have the algorithm:

1. Compute the measurement to target association $\hat{\Psi}^{(p)}_N$ by maximizing the cost function $J_p(\psi_t)$ based on $\hat{X}^{(p-1)}_N$.
2. Compute $\hat{X}^{(p)}_N$ given $\hat{\Psi}^{(p)}_N$ using fixed interval smoothing.
Evaluation and Conclusion

Mimicing the simulation setup from previous work but adding that $\lambda V_S = 4$ and that $P_D = 0.9$.

<table>
<thead>
<tr>
<th></th>
<th>MSMF</th>
<th>EMDA</th>
<th>MCDA</th>
<th>ICMDA</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>96.3682</td>
<td>7.4059</td>
<td>7.5234</td>
<td>7.2387</td>
</tr>
<tr>
<td>Time</td>
<td>1.000</td>
<td>1.8835</td>
<td>6.3295</td>
<td>2.7651</td>
</tr>
</tbody>
</table>

RMSE computed over 50 independent runs.

The EM and both Monte Carlo algorithms are superior to MSMF but have a larger computational complexity. The proposed deterministic ICMDA filter combines a moderate computation time with accurate data association. The execution time of MCDA is highly dependent on the length of the simulation chain, in our simulations it iterates on the average 3 times longer than the MMAP estimators.