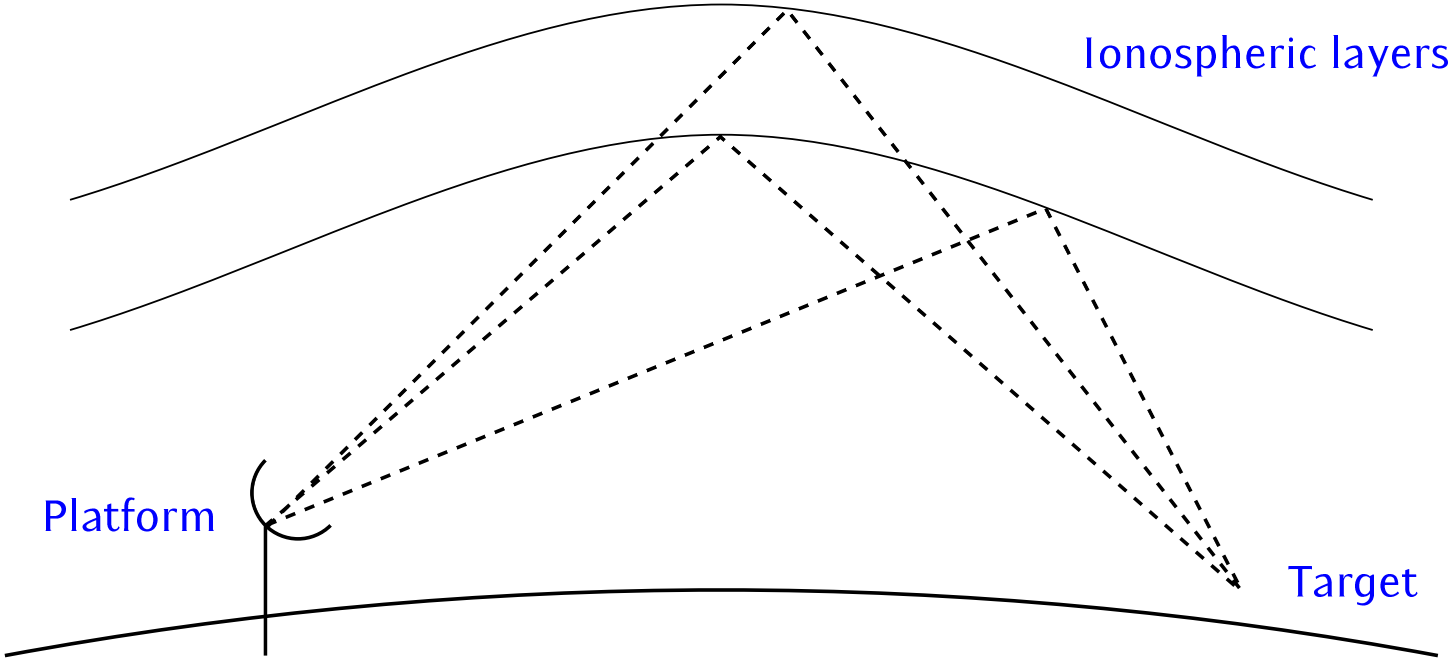
Markov Chain Monte Carlo Data Association for Target Tracking

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
Cambride, UK
ad2@eng.cam.ac.uk

Problem Description

The multiple measurement data association problem ariases 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

$$x_{t+1} = F_t x_t + w_t, \quad t = 0, 1, \ldots \tag{1}$$

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

$$y_i(t) = \begin{cases} H_1 x_t + e_1(t) & \text{for model 1} \\ \vdots & \vdots \\ H_n x_t + e_n(t) & \text{for model } n \\ \text{clutter} & \text{otherwise} \end{cases} \tag{2}$$

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

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

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

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

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 μ_t and is the main challenge in this application.

Markov Chain Algorithm

Given an observed data batch $\mathbb{Y}_N = \{Y(t)\}_{t=0}^N$, 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 $\mathbb{X}_N = \{x_t\}_{t=0}^N$.

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

$$\psi_t^{(p)} \sim p(\psi_t^{(p)} | \mathbb{Y}_N, \Psi_{\neg t}^{(p)}) \tag{3}$$

3. Set $p := p + 1$ and repeat at step 2.

The chosen association event at time t in pass p is denoted $\psi_t^{(p)}$ and
$\Psi_{\neg t}^{(p)} = \{\Psi_{t-1}^{(p)}, \Psi_{t+1:N}^{(p-1)}\} = \{\psi_0^{(p)}, \ldots, \psi_{t-1}^{(p)}, \psi_{t+1}^{(p-1)}, \ldots, \psi_N^{(p-1)}\}$

The full conditional distribution (3) is discrete with π_t possible association events at each time instant, and given by

$$p(\psi_t^{(p)}|\mathbb{Y}_N, \Psi_{\neg t}^{(p)}) = \frac{p(\mathbb{Y}_N|\Psi_{\neg t}^{(p)}, \psi_t^{(p)}) \Pr(\psi_t^{(p)})}{p(\mathbb{Y}_N|\Psi_{\neg t}^{(p)})} \propto p(\mathbb{Y}_N|\Psi_{\neg t}^{(p)}, \psi_t^{(p)}) \Pr(\psi_t^{(p)})$$
$$\propto p(Y_t|\mathbb{Y}_{t-1}, \Psi_t^{(p)}) \Pr(\psi_t^{(p)}) \int p(\mathbb{Y}_{t+1:N}|\Psi_{t+1:N}^{(p-1)}, x_t)\, p(x_t|\Psi_t^{(p)}, \mathbb{Y}_t)\, dx_t$$

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

$$p(\psi_t^{(p)}|\mathbb{Y}_N, \Psi_{\neg t}^{(p)}) \propto \mathsf{N}(Y_t; \widehat{Y}_{t|t-1}, S_t) \frac{p_c(\mu_t - m_{t,p}) P_D^{m_{t,p}} (1 - P_D)^{n - m_{t,p}} V_S^{m_{t,p}}}{r_t(m_{t,p})} \times$$
$$|I + P_{t|t}\Gamma_{t|t+1}|^{-1/2} \exp\Bigg(-\frac{1}{2}\|\widehat{x}_{t|t}\|^2_{\Gamma_{t|t+1}} + \widehat{x}_{t|t}^T \widehat{a}_{t|t+1} +$$
$$\frac{1}{2}\|\widehat{a}_{t|t+1} - \Gamma_{t|t+1}\widehat{x}_{t|t}\|^2_{(\Gamma_{t|t+1} + P_{t|t}^{-1})^{-1}}\Bigg), \quad (4)$$

where $\widehat{Y}_{t|t-1}$, $\widehat{x}_{t|t}$, and $P_{t|t}$ are based on $\psi_t^{(p)}$ and $\|x\|^2_A = x^T A x$.

The full conditional probabilities are computed using a a backwards information filter under the association hypothesis $\Psi_N^{(p-1)}$

$$\hat{a}_{t|t} = \hat{a}_{t|t+1} + \mathbb{H}_t^T \mathbb{R}_t^{-1} Y_t \qquad \hat{a}_{N|N+1} = 0 \tag{5a}$$

$$\Gamma_{t|t} = \Gamma_{t|t+1} + \mathbb{H}_t^T \mathbb{R}_t^{-1} \mathbb{H}_t \qquad \Gamma_{N|N+1} = 0 \tag{5b}$$

$$\hat{a}_{t-1|t} = F_t^T \left(I - \Gamma_{t|t} \left(\Gamma_{t|t} + Q_t^{-1} \right)^{-1} \right) \hat{a}_{t|t} \tag{5c}$$

$$\Gamma_{t-1|t} = F_t^T \left(I - \Gamma_{t|t} \left(\Gamma_{t|t} + Q_t^{-1} \right)^{-1} \right) \Gamma_{t|t} F_t \tag{5d}$$

Where Y_t in (5a) is the stacked vector of non-clutter measurements under the hypothesis $\psi_t^{(p-1)}$, and $\mathbb{H}_t$ and $\mathbb{R}_t$ are the corresponding measurement equation matrix and measurement error covariance, respectively.

Algorithm 1 (Monte Carlo Data Association, MCDA)

1. *Pick the initial association* $\Psi_N^{(0)}$. *Set* $p = 1$.
2. *Compute* (5) *using* $\Psi_N^{(p-1)}$. *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* $\hat{x}_{t|t}^k$ *and* $P_{t|t}^k$.
   (b) *Compute* (4) *under all* π_t *hypotheses, normalize this discrete density, and generate a sample* $\psi_t^{(p)}$.
   (c) *Set* $t := t+1$, *run a time update step with the Kalman filter corresponding to* $\psi_t^{(p)}$, *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* $\mathbb{X}_N^{(p)} = \mathsf{E}\left(\mathbb{X}_N | \mathbb{Y}_N, \Psi_N^{(p)}\right)$.
5. *Output* $\Psi_N^{(p)}$ *and* $\mathbb{X}_N^{(p)}$. *Set* $p := p + 1$ *and return to item 2.*

Instead of sampling from the full conditional in item 3(b) above, MMAP estimates can be formed by deterministically picking the most likeliy association, according to the Iterated Conditional Mode (ICM). The resulting ICMDA algorithm is iterated until two subsequent passes yields the same association sequence.

Previous Apporoaches

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.

$$\mathsf{E}(x_t|\mathbb{Y}_t) \approx \hat{x}_{t|t} = \sum_{i=1}^{\pi_t} c_t^{(i)} \hat{x}_{t|t,i}$$

$$\mathsf{Cov}(x_t|\mathbb{Y}_t) \approx P_{t|t} = \sum_{i=1}^{\pi_t} c_t^{(i)} (P_{t|t,i} + \hat{x}_{t|t,i}\hat{x}_{t|t,i}^T) - \hat{x}_{t|t}\hat{x}_{t|t}^T$$

where

$$c_t^{(i)} \propto \frac{p_c(\mu_t - m_{t,i}) P_D^{m_{t,i}} (1 - P_D)^{n - m_{t,i}}}{r_t(m_{t,i})} \prod_{(j,k) \in \psi_{t,i}} \mathsf{N}(y_j(t); H_k \hat{x}_{t|t-1}, H_k P_{t|t-1} H_k^T + R_k)$$

Expectation Maximization Data Association, EMDA computes MMAP estimates of the association events by treating the target state sequence $\mathbb{X}_N$ as missing (unobserved) data and the sequence of association events Ψ_N as the sought parameters using the EM-algorithm

$$\widehat{\Psi}_N^{(p+1)} = \arg\max_{\Psi_N} \mathsf{E}\left(\ln p(\mathbb{X}_N, \mathbb{Y}_N, \Psi_N) | \mathbb{Y}_N, \widehat{\Psi}_N^{(p)}\right)$$

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

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

Evaluation and Conclusion

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

	MSMF	EMDA	MCDA	ICMDA
RMSE	96.3682	7.4059	7.5234	7.2387
Time	1.000	1.8835	6.3295	2.7651

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.