

# Multiscan association as a multi-commodity flow optimization problem

G. Battistelli<sup>1</sup>, L. Chisci<sup>1</sup>, F. Papi<sup>1</sup>, A. Benavoli<sup>2</sup>, A. Farina<sup>2</sup>

<sup>1</sup>*DSI, Università di Firenze, Florence, Italy*  
e-mail: {battistelli,chisci}@dsi.unifi.it

<sup>2</sup>*Engineering Division, SELEX Sistemi Integrati, Rome, Italy*  
e-mail: {abenavoli,afarina}@selex-si.com

**Abstract**—Multiscan data association can significantly enhance tracking performance in critical radar surveillance scenarios involving multiple targets, low detection probability, high false alarm probability, evasive target maneuvers and finite radar resolution. Unfortunately, however, this approach is affected by the curse of dimensionality which hinders its real-time application for tracking problems with short scan periods and/or long association windows and/or many measurements. In this paper it is shown how the formulation of the multiscan association as a multi-commodity flow optimization problem allows a relaxation of the association problem which, on one hand, guarantees close-to-optimal association performance and, on the other hand, implies a significant reduction of the computational load.

## I. INTRODUCTION

It is well known that in critical radar scenarios, like e.g. GMTI (Ground Moving Target Indicator) radar tracking, involving many targets with low detection probability moving in a highly cluttered environment, a considerable improvement in tracking performance can be achieved by using *multiscan*, in place of *single scan*, data association [1]. In fact, the memory of multiscan association allows the partial recovery of association errors which represent the main error source in multitarget tracking. On the other hand, the complexity of multiscan association grows exponentially with the size (number of scans) of the association window. More precisely let  $M$  be the number of measurements per scan,  $T$  the number of tracks and  $S$  the number of scans in the association window, then multiscan association amounts to a linear binary programming problem with “in the order of”  $TM^S$  variables, i.e.  $O(TM^S)$ , and  $O(MS)$  constraints. To avoid the “curse of dimensionality” [2], Lagrangian relaxation techniques have been proposed [3]. In this paper, a novel relaxation technique is proposed whereby it is possible to represent the association problem on a suitable graph involving  $O(MS)$  nodes and  $O(M^2S)$  arcs: each track corresponds to a commodity and the minimum cost path needs to be determined for each commodity. This solution, though approximate and sub-optimal, allows to get rid of the computational burden exponentially growing with the number of scans and, hence, allows fast real-time implementation of multiscan association. On the other hand, simulation results obtained from realistic case studies will demonstrate the effectiveness of the proposed approximation. The rest of the paper is organized as follows. Section II provides the exact formulation of the multiscan data association

problem. Section III presents its approximate reformulation as a multi-commodity flow optimization problem. In Section IV a performance evaluation of the proposed approach is given and, finally, section V ends the paper with some concluding remarks.

## II. FORMULATION OF THE MULTISCAN DATA ASSOCIATION PROBLEM

In this section, the multiscan association problem over a window of  $S$  scans (in short *S-D Assignment*) is described in detail. To this end, let  $T$  be the number of tracks<sup>1</sup>. It is assumed that for each track  $n = 1, 2, \dots, T$  an estimate  $\bar{\mathbf{x}}_n$  of the track state at the beginning of the window is available. Further, for each  $k = 1, 2, \dots, S$ , let us denote by  $\mathcal{Z}_k = \{\mathbf{z}_{1,k}, \mathbf{z}_{2,k}, \dots, \mathbf{z}_{M_k,k}\}$  the set of measurements obtained at scan  $k$  ( $M_k$  being the cardinality of  $\mathcal{Z}_k$ ). Given the  $T$  tracks and the  $S$  sets of measurements  $\mathcal{Z}_k$  for  $k = 1, 2, \dots, S$ , the objective of *S-D Assignment* is to assign a sequence of  $S$  measurements to each track, where the  $k$ -th element of such a sequence is either taken from  $\mathcal{Z}_k$  or represents a missed detection.

Among all feasible assignments, an optimal one is found by minimizing a suitably defined cost. In this connection, let  $c(m_1, m_2, \dots, m_S, n)$  denote the cost of associating a certain sequence  $(m_1, m_2, \dots, m_S)$  to the track  $n = 1, 2, \dots, T$ . Here, each variable  $m_k$ , for  $k = 1, 2, \dots, S$ , takes its value in the set  $\{0, 1, \dots, M_k\}$  and refers either to the  $m_k$ -th measurement of the set  $\mathcal{Z}_k$  (when  $m_k > 0$ ) or to a missed detection (when  $m_k = 0$ ). The cost takes the form

$$c(m_1, m_2, \dots, m_S, n) = \sum_{k=1}^S c_k(m_1, m_2, \dots, m_k, n)$$

where  $c_k(m_1, m_2, \dots, m_k, n)$  is the cost of adding  $m_k$  to the partial sequence  $(m_1, m_2, \dots, m_{k-1})$ .

In order to derive an expression for each  $c_k(m_1, m_2, \dots, m_k, n)$ , some preliminary definitions are needed. Given a measurement  $\mathbf{z}$  and a state  $\mathbf{x}$ , let  $\Lambda(\mathbf{x}|\mathbf{z})$  be the likelihood of  $\mathbf{x}$  given  $\mathbf{z}$ , i.e., the probability that

<sup>1</sup>In this paper, for the sake of brevity, the track deletion and track initialization problems are not considered. However, many solutions to these problems are available in the literature; see, e.g., [3] and the references therein.

the measurement  $\mathbf{z}$  originates from a target with state  $\mathbf{x}$ . Further, consider some filtering mechanism (e.g. the extended Kalman filter or a Sequential Monte Carlo filter) that provides an estimate  $\mathbf{x}$  as a function of a measurement  $\mathbf{z}$  and an one-step-behind estimate  $\mathbf{x}^-$ . The filter is supposed to consist of two parts: a prediction step

$$\mathbf{x}^+ = \text{pred}(\mathbf{x}^-)$$

and an innovation update

$$\mathbf{x} = \text{update}(\mathbf{x}^+, \mathbf{z}).$$

The propagation of the other statistics (e.g., the covariance matrices for the extended Kalman filter) as well as their involvement in the computation of the estimate  $\mathbf{x}$  is omitted for the sake of compactness.

Then, one can write

$$c_k(m_1, m_2, \dots, m_k, n) = \begin{cases} -\log(P_d \Lambda(\bar{\mathbf{x}}_{n,k} | \mathbf{z}_{m_k, k})), & m_k > 0 \\ -\log(1 - P_d), & m_k = 0 \end{cases} \quad (\text{II.1})$$

where  $P_d$  is the detection probability and  $\bar{\mathbf{x}}_{n,k}$  is the prediction of the state of track  $n$  at scan  $k$  on the basis of the partial sequence  $(m_1, \dots, m_{k-1})$ . Given a sequence  $(m_1, m_2, \dots, m_S)$ , the predictions  $\bar{\mathbf{x}}_{n,k}$  can be computed recursively as

$$\bar{\mathbf{x}}_{n,k} = \begin{cases} \text{pred}(\text{update}(\bar{\mathbf{x}}_{n,k-1}, \mathbf{z}_{m_{k-1}, k-1})), & m_{k-1} > 0 \\ \text{pred}(\bar{\mathbf{x}}_{n,k-1}), & m_{k-1} = 0 \end{cases}$$

for  $k = 2, 3, \dots, S$ . The recursion is initialized as

$$\bar{\mathbf{x}}_{n,1} = \bar{\mathbf{x}}_n.$$

By exploiting the foregoing definitions, it is possible to give a mathematical formulation of *S-D Assignment*. With this respect, for any possible sequence of measurements  $(m_1, m_2, \dots, m_S)$  and for any track  $n$ , let us define a binary association variable  $a(m_1, m_2, \dots, m_S, n)$  that takes value 1 if  $(m_1, m_2, \dots, m_S)$  is associated to track  $n$  and value 0 otherwise. Then, the optimal assignments can be obtained by minimizing the loss functional

$$\sum_{n=1}^T \sum_{m_1=0}^{M_1} \sum_{m_2=0}^{M_2} \dots \sum_{m_S=0}^{M_S} c(m_1, m_2, \dots, m_S, n) \times a(m_1, m_2, \dots, m_S, n)$$

under the constraints

$$\sum_{n=1}^T \sum_{m_1=0}^{M_1} \dots \sum_{m_{k-1}=0}^{M_{k-1}} \sum_{m_{k+1}=0}^{M_{k+1}} \dots \sum_{m_S=0}^{M_S} a(m_1, \dots, m_{k-1}, m, m_{k+1}, \dots, m_S, n) \leq 1$$

for  $m = 1, 2, \dots, M_k$ ,  $k = 1, 2, \dots, S$  (II.2)

and

$$\sum_{m_1=0}^{M_1} \sum_{m_2=0}^{M_2} \dots \sum_{m_S=0}^{M_S} a(m_1, m_2, \dots, m_S, n) = 1$$

for  $n = 1, 2, \dots, T$ . (II.3)

Condition (II.2) is needed to ensure that each measurement can be assigned to at most one track. As to (II.2), it imposes that exactly one sequence be assigned to each track.

As should be evident, *S-D Assignment* turns out to be a linear binary programming problem with  $O(TM^S)$  variables and  $O(MS)$  constraints. Therefore, the complexity of the problem increases exponentially with the number of scans  $S$  (see also Fig. 1 where a graphical representation of the hypothesis tree for a single track is provided). Indeed, *S-D Assignment* has been shown to be NP-hard (see [8]). As a consequence, unless very small instances of the problem are considered, the possibility of solving it exactly is ruled out by the so-called curse of dimensionality (i.e., the exponential growth of the computational burden).

In [2], an efficient algorithm based on a successive Lagrangian relaxation technique was proposed for the approximate solution of *S-D Assignment*. In the next section, a different graph-based formulation of the problem is provided. Based on such a formulation and exploiting a Markovian approximation, the original *S-D Assignment* problem will be relaxed into a multi-commodity flow optimization problem.

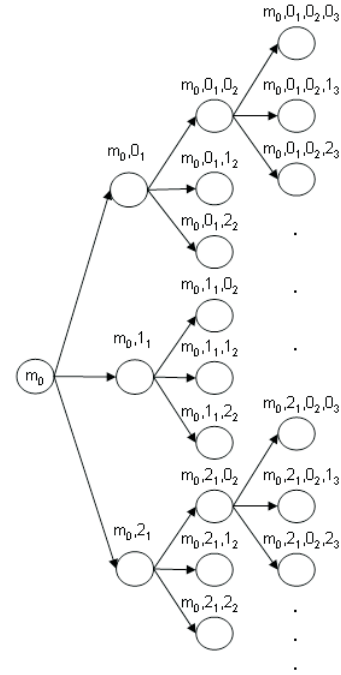


Fig. 1. Hypothesis tree for a single track.

### III. S-D ASSIGNMENT AS MINIMUM COST MULTI-COMMODITY FLOW PROBLEM

In this section, the key idea is to reformulate *S-D Assignment* as the problem of finding the multi-commodity flow



Node  $s$  is the unique source node with a unitary supply for each commodity (see Eq. (III.2)). Node  $d$  is the unique target node with unitary demand for each commodity (see Eq. (III.3)). Condition (III.4) imposes the conservation of the flow in all the internal nodes of the graph and for each commodity. Finally, the capacity constraint (III.5) ensures that at most a unitary flow can go through each node of the graph (this serves the same purpose as condition (II.2) in the original *S-D assignment* problem).

For what concerns the determination of the solution, since the proposed multi-commodity flow optimization problem is a polynomial size linear program, it can be solved in polynomial time and space using either the ellipsoid algorithm or the interior-point method [4]. Further, in spite of an exponential worst-case complexity, the simplex algorithm can solve linear programs quite efficiently in practice with average complexity  $O(\min\{V^2, C^2\})$  for a problem with  $V$  variables and  $C$  constraints [5]. Finally, various decomposition and relaxation techniques are available in the literature to derive approximate solutions with low computational burden to large-scale multi-commodity flow problems (see [6] and the references therein).

It is important to remark that, unfortunately, due to the capacity constraints (III.5), the constraint matrix of the associated linear program is not totally unimodular and the considered multi-commodity minimum cost flow problem does not have the integrality property. As a consequence, at least in principle, the optimal flows  $f^\circ(j_{k-1}, i_k, n)$  might not be integral even if all supplies, demands, and capacities are integral. Of course, this would not be admissible for the purpose of multiscan multitarget association and, in such a case, the optimal flows  $f^\circ(j_{k-1}, i_k, n)$  should be somehow rounded up in order to determine a feasible binary solution. However, as a matter of fact, in all the considered study cases the optimal flows turned out to be integral. This state of affairs can be understood by noting that the flow conservation constraints (III.4) impose that many vertices of the simplex be integer. Then one may conjecture that, while not true in general, the integral property might hold true “with probability one” for the considered problem.

As a final remark, it is pointed out that idea of modelling multiscan multitarget association as a minimum cost flow problem dates back to [7]. The main novelties of the proposed approach concern: i) the construction of the graph based on a relaxation of the *S-D Assignment* problem (as defined in [8]); ii) the possibility of dealing with missed detections and clutter. Attempts in this direction are also in [9], [10] where a different graph-based relaxation of *S-D assignment* is proposed and a Viterbi-like algorithm is adopted for the determination of the multiscan multitarget association.

#### IV. PERFORMANCE EVALUATION

In this section, it is shown that the proposed graph-based approximation has negligible impact on performance by means of simulation experiments in critical multitarget radar tracking scenarios.

For each target, the true state at discrete time  $t$  is  $\mathbf{x}(t) = [x(t), \dot{x}(t), y(t), \dot{y}(t)]'$ , where  $(x(t), y(t))$  provides the position and  $(\dot{x}(t), \dot{y}(t))$  the velocity of the target in Cartesian coordinates. The target’s motion is described by the constant velocity model:

$$\mathbf{x}(t+1) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{w}(t)$$

$$\mathbf{A} = \begin{bmatrix} 1 & T_s & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T_s \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} T_s^2/2 & 0 \\ T_s & 0 \\ 0 & T_s^2/2 \\ 0 & T_s \end{bmatrix}$$

where  $T_s$  is the sampling time and  $\mathbf{w}(t)$  is a zero-mean white process noise with covariance matrix  $\mathbf{Q}$ .

It is assumed that the sensor is a radar providing measurements of range  $r(t)$ , range rate  $\dot{r}(t)$  and azimuth  $\theta(t)$ . Then the measurement equation is given by

$$\mathbf{z}(t) = \mathbf{h}(\mathbf{x}(t)) + \mathbf{v}(t) \quad (\text{IV.1})$$

where  $\mathbf{z}(t) = [r(t), \theta(t), \dot{r}(t)]'$  is the measurement vector at time  $t$ , the nonlinear function  $\mathbf{h}(\cdot)$  is defined as

$$\mathbf{h}(\mathbf{x}) = \begin{bmatrix} \sqrt{x^2 + y^2} \\ \angle(x + iy) \\ \frac{x\dot{x} + y\dot{y}}{\sqrt{x^2 + y^2}} \end{bmatrix},$$

and  $\mathbf{v}(t)$  is a zero-mean white measurement noise with covariance matrix  $\mathbf{R} = \text{diag}\{\sigma_r, \sigma_\theta, \sigma_{\dot{r}}\}$ .

In the multitarget context let  $\mathbf{x}_n(t)$  denote the state of the  $n$ th target (track) at time  $t$ .

Given the solution of the *S-D Assignment Problem* at time  $t$ , the state estimate  $\hat{\mathbf{x}}_n(t)$  is updated via EKF with the last measurement of the multiscan sequence associated to the  $n$ th target. The estimation error for the  $n$ th target is defined as the distance between the true Cartesian position and its estimate, i.e.

$$d_n(t) = \sqrt{(x_n(t) - \hat{x}_n(t))^2 + (y_n(t) - \hat{y}_n(t))^2} \quad (\text{IV.2})$$

Let us also introduce the binary variable  $L_n(t)$  taking value 1 if the measurement associated to track  $n$  is the right one and value 0 otherwise. Monte Carlo simulations have been carried out by randomly varying the noise realizations, the time location of missed detections and clutter. The performance indices  $\bar{d}(t)$  and  $\bar{L}(t)$ , obtained by averaging the previously defined  $d_n(t)$  and  $L_n(t)$  over the various tracks and over 50 Monte Carlo runs, have been evaluated in order to compare the multi-commodity association approach proposed in this paper with the ideal case (exact association).

#### Scenario 1

Let us first consider the multitarget case-study depicted in figure 3. In this scenario, there are  $T = 10$  targets with detection probability  $P_d = 0.8$  and 10 uniformly distributed clutter measurements at every step  $t$ ; the sampling time is  $T_s = 5s$ . The measurement noise standard deviations are  $\sigma_r = 15 m$ ,  $\sigma_{\dot{r}} = 1 m/s$  and  $\sigma_\theta = 0.25$ .

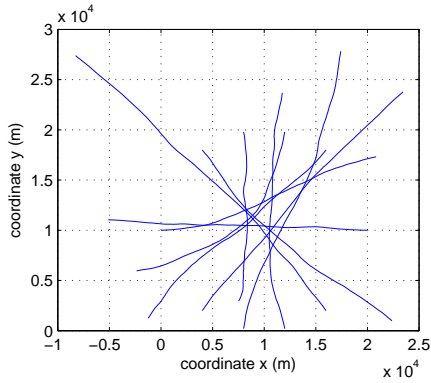


Fig. 3. First Scenario: target trajectories

Figure 4 reports the averaged distance  $\bar{d}(t)$  obtained with the proposed *S-D Assignment* and in the ideal case. The percentage of correct assignments is shown in figure 5. From these figures, it is evident that the proposed multi-commodity relaxation of *S-D Assignment* provides a negligible performance deterioration with respect to the ideal case.

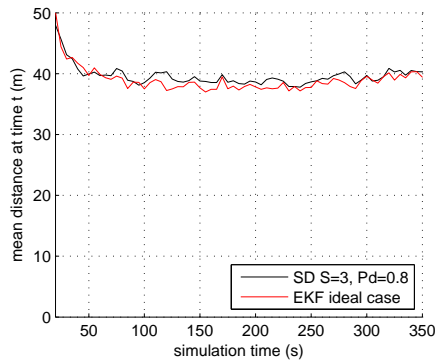


Fig. 4. Average estimation error

### Scenario 2

Let us now consider the multitarget case-study of figure 6 where three maneuvering targets are crossing at time  $t = 160$  s and  $t = 270$  s. The following parameters have been adopted in the simulation: radar scan time  $T_s = 2$  s; range standard deviation  $\sigma_r = 20$  m; range rate standard deviation  $\sigma_{\dot{r}} = 1$  m/s; azimuth standard deviation  $\sigma_\theta = 0.5^\circ$ ; detection probability  $P_d = 0.8$ ; 20 clutter plots at each scan. Fig. 7 plots  $\bar{d}(t)$  for the *S-D* multi-commodity approach as well as in the ideal case, while Fig. 8 reports  $\bar{L}(t)$  for the *S-D* multi-commodity approach. It can be seen that the performance of *S-D* multi-commodity association is comparable with the performance achievable under ideal conditions.

### V. CONCLUSIONS

A novel graph-based relaxation of the well known *S-D Assignment* problem has been proposed. The idea has been to formulate the problem as a multi-commodity flow optimization

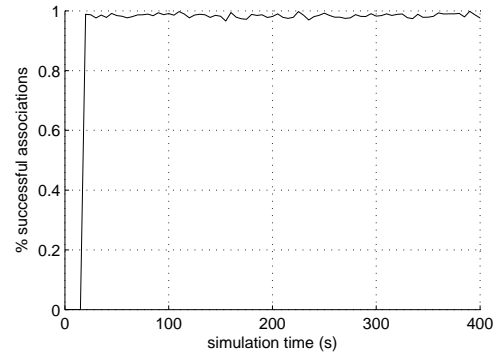


Fig. 5. Percentage of successful associations

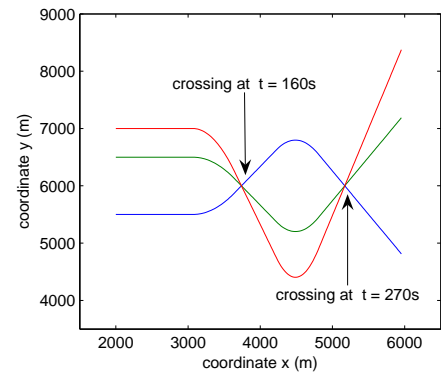


Fig. 6. Second Scenario: target trajectories

on a suitable graph. Simulation experiments in complex multitarget radar tracking scenarios have demonstrated the good association performance of the proposed multi-commodity graph-based relaxation scheme.

### ACKNOWLEDGMENTS

This work was partially supported by a grant of SELEX-Sistemi Integrati.

### REFERENCES

- [1] Y. Bar-Shalom and X.R. Li, *Multitarget-multisensor tracking: principles and techniques*, YBS Publishing, Storrs, CT, 1995.
- [2] F. Daum and J. Huang, "Curse of dimensionality and particle filters," *Proc. Aerospace Conference*, 2003, vol. 3, pp. 4.1979 - 4.1993, 2003.
- [3] L. Lin, Y. Bar-Shalom, and T. Kirubarajan, "New assignment-based data association for tracking move-stop-move targets," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 40, no. 2, pp. 714 - 725, 2004.
- [4] N. K. Karmarkar, "A new polynomial-time algorithm for linear programming," *Combinatorica*, vol. 26, pp. 373395, 1984.
- [5] I. Adler and N. Megiddo, "A simplex algorithm whose average number of steps is bounded between two quadratic functions of the smaller dimension," *Journal of the Association for Computing Machinery*, vol. 32, pp. 871895, 1985.
- [6] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin, *Network Flows: Theory, Algorithms, and Applications*, Prentice Hall, Englewood Cliffs, N.J., 1993.
- [7] D. A. Castanon, "Efficient algorithms for finding the  $K$  best paths through a trellis," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 26, no. 2, pp. 405 - 410, 1990.

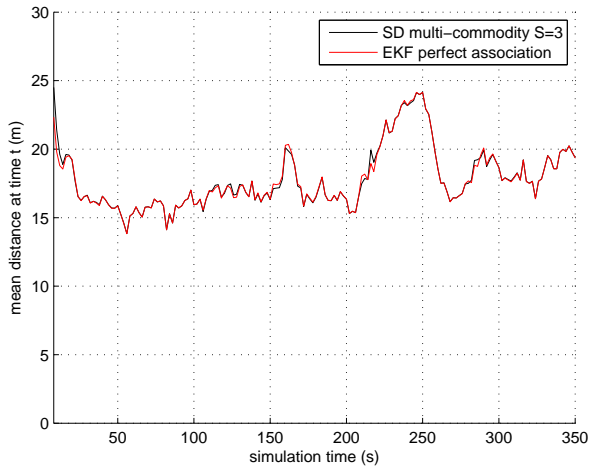


Fig. 7. Average estimation error

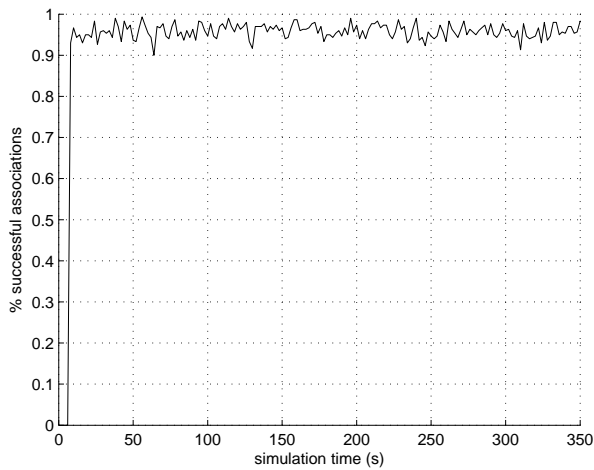


Fig. 8. Percentage of successful associations

- [8] S. Deb, M. Yeddanapudi, K. Pattipati, and Y. Bar-Shalom, "A generalized S-D assignment algorithm for multisensor-multitarget state estimation," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 33, no. 2, pp. 523 - 538, 1997.
- [9] G. W. Pulford, "Multi-target Viterbi data association," *Proc. of the 9th International Conference on Information Fusion, Fusion 2006*, Florence, Italy, July 2006.
- [10] G. W. Pulford and B. F. La Scala, "Multi-hypothesis Viterbi data association tracking: a comparative assessment," submitted to *IEEE Trans. on Aerospace and Electronic Systems*.