Computational Optimization and Applications, 3, (1994), 27–57

Multidimensional Assignment Formulation of Data Association Problems Arising from Multitarget and Multisensor Tracking*

AUBREY B. POORE
Department of Mathematics Colorado State University, Fort Collins, Colorado 80523

Received February 17, 1993; Revised September 22, 1993

Abstract. The ever-increasing demand in surveillance is to produce highly accurate target and track identification and estimation in real-time, even for dense target scenarios and in regions of high track contention. The use of multiple sensors, through more varied information, has the potential to greatly enhance target identification and state estimation. For multitarget tracking, the processing of multiple scans all at once yields high track identification. However, to achieve this accurate state estimation and track identification, one must solve an NP-hard data association problem of partitioning observations into tracks and false alarms in real-time. The primary objective in this work is to formulate a general class of these data association problems as multidimensional assignment problems to which new, fast, near-optimal, Lagrangian relaxation based algorithms are applicable. The dimension of the formulated assignment problem corresponds to the number of data sets being partitioned with the constraints defining such a partition. The linear objective function is developed from Bayesian estimation and is the negative log posterior or likelihood function, so that the optimal solution yields the maximum a posteriori estimate. After formulating this general class of problems, the equivalence between solving data association problems by these multidimensional assignment problems and by the currently most popular method of multiple hypothesis tracking is established. Track initiation and track maintenance using an N-scan sliding window are then used as illustrations. Since multiple hypothesis tracking also permeates multisensor data fusion, two example classes of problems are formulated as multidimensional assignment problems.

Keywords: data association, multidimensional assignment problems, multiple hypothesis tracking, multitarget tracking, multisensor data fusion

1. Introduction

The central problem in any multitarget/multisensor surveillance system is the data association problem of partitioning the observations into tracks and false alarms [4, 7, 25, 27]. Current methods for multitarget tracking generally fall into two categories: sequential and deferred logic. Methods for the former include nearest neighbor, one-to-one or few-to-one assignments, and all-to-one assignments as

*This work was partially supported by the Air Force Office of Scientific Research through AFOSR Grant Numbers AFOSR-91-0138 and F49620-93-1-0133 and by the Federal Systems Company of the IBM Corporation in Boulder, CO and Owego, NY.

in the joint probabilistic data association (JPDA) [4]. For track maintenance, the nearest neighbor method is valid in the absence of clutter when there is little or no track contention, i.e., when there is little chance of misassociation. Problems involving one-to-one or few-to-one assignments are generally formulated as (two dimensional) assignment or multi-assignment problems for which there are some excellent algorithms [5, 6, 15]. This methodology is real-time but can result in a large number of partial and incorrect assignments in dense or high contention scenarios, and thus incorrect track identification. The difficulty is that decisions, once made, are irrevocable, so that there is no mechanism to correct misassociations. The use of all observations in a scan or within a neighborhood of a predicted track position (e.g., JPDA) to update a track moderates the misassociation problem and has been highly successful for tracking a few targets in dense clutter [4].

Deferred logic techniques consider several data sets or scans of data all at once in making data association decisions. At one extreme is batch processing in which all observations (from all time) are processed together, but this is computationally too intensive for real-time applications. The other extreme is sequential processing. Deferred logic methods between these two extremes are of primary interest in this work. The principal deferred logic method used to track large numbers of targets in low to moderate clutter is called multiple hypothesis tracking (MHT) in which one builds a tree of possibilities, assigns a likelihood score to each track, develops an intricate pruning logic, and then solves the data association problem by an explicit enumeration scheme. The use of these enumeration schemes to solve this NP-hard combinatorial optimization problem in real-time is inevitably faulty in dense scenarios since the time required to solve the problem optimally can grow exponentially in the size of the problem.

Another important aspect in surveillance systems is the growing use of multisenor data fusion [2, 8, 11, 27] in which one associates reports from multiple sensors together. Once matched, this more varied information has the potential to greatly enhance target identification and state estimation [8]. The central problem is again that of data association and the principal method employed is multiple hypothesis tracking. Although the current setting is that of multitarget/multisensor tracking, these data association problems occur in other areas such as the recent work of Cox, Rehg, and Hingorani on edge grouping and contour segmentation [10].

These data association problems of partitioning multiple data sets are fundamentally important NP-hard combinatorial optimization problems. Thus the objective in this work is to formulate a large class of these problems as multidimensional assignment problems and to give two examples from multitarget tracking and multisensor data fusion. The development in Section 2 extracts the salient features and assumptions that occur in a large class of these problems; the concrete examples are given in Sections 3 and 5. Our reason for this specific formulation is the recent development of a class of real-time and near-optimal algorithms [12, 20, 23, 24] for solving these multidimensional assignment problems to the noise level in the problem. Also, the current work differs from our earlier work [20, 21, 22] in the generality of the formulation, derivation of the cost coefficients, and the allowable dynamics.

The use of combinatorial optimization in multitarget tracking is not new and dates back to the mid-sixties and the pioneering work of Sittler [26], who used maximum likelihood estimation to evaluate all possible track updates and employed track splitting (several hypotheses were maintained for each track) and pruning (when their probabilities fell below certain threshold). Maximum likelihood estimation was further investigated by Stein and Blackman [7], who developed a comprehensive probability for track initiation, track length expectancy, missed detections and false alarms. Morefield [17] pioneered the use of the integer programming to solve set packing and covering problems arising from a data association problem. Multiple hypothesis tracking has been popularized by the fundamental work of Reid [25]. These works are further discussed in the books of Blackman [7], Bar-Shalom and Fortmann [4], and Waltz and Llinas [27], which also serve as excellent introductions to the field of multitarget tracking and multisensor data fusion. Deb, Pattipati, and Bar-Shalom [11] have also formulated sensor fusion problems in terms of these multidimensional assignment problems. A discussion of these problems is presented in Section 5.

There are a host of topics not considered in this work. We do not discuss algorithms for solving the multidimensional assignment problems such as multiple hypothesis tracking, greedy, branch and bound techniques, and the Lagrangian relaxation methods currently under development [20, 24]. We do not discuss the design of a tracking system and the host of associated issues [2, 3, 7], the use of parallel computation, the various nonlinear estimation techniques for nonlinear dynamics, maneuvering targets characterized by abrupt and unknown changes in the system dynamics, the integration of multisensor data into a multitarget tracking system. Many of these topics are discussed in the aforementioned references and the many references therein.

The outline of the paper is as follows. General conditions under which linear multidimensional assignment problems are applicable to data association are established in Section 2. Utilizing the results of this section, data association problems in multitarget tracking formulated for multiple hypothesis tracking methods are posed as multidimensional assignment problems in Section 3. This development includes track initiation and track maintenance using an N-scan sliding window as developed in earlier work [22]. Multiple hypothesis tracking is also used extensively in multisensor data fusion [8, 9, 11, 27] at each level of the fusion process. Although we do not examine these problem areas in detail, we do present two example problems in Section 5.

2. Assignment formulation of some general data association problems

The goal of this section is to formulate the data association problem for a large class of multiple target tracking and sensor fusion problems as a multidimensional assignment problem. This development extracts the salient features and assumptions that occur in a large class of these problems; concrete examples will be given in subsequent sections. The path to these assignment problems begins with a brief review of the work of Morefield [17], who formulated data association problems as set covering and packing problems. (To explain some of the critical independence assumptions in Morefield's work (e.g., the assumptions in equation (2.7) below), we develop a probabilistic framework for this formulation in Appendix B.) After this review and a refinement of the definition of a partition of the data, the assignment problems are derived in three different ways. This development generalizes the various special and example cases in our earlier work [20, 21, 22]. A reason for this assignment formulation is that it exhibits a rich structure that can be exploited algorithmically [12, 20, 23, 24]. The plan for this section is to begin with a simple description of a model multitarget tracking. The combinatorial optimization problem for data association is then stated in equation (2.2) followed by an explanation of its meaning. The three derivations of the assignment problem are then given in the latter half of the section.

In tracking, a common surveillance problem is to estimate the past, current, or future state of a collection of objects (e.g., airplanes) moving in three dimensional space from a sequence of measurements made of the surveillance region by one or more sensors. The objects will be called targets. The dynamics of these targets are generally modeled from physical laws of motion, but there may noise in the dynamics and certain parameters of the motion may be unknown. (The dynamics are often modeled as stochastic differential equations.) At time $t = 0$ a sensor (or sensors) is turned on to observe the region. In an ideal situation measurements are taken at a finite sequence of times $\{t_k\}_{k=0}^{n}$ where $0 = t_0 < t_1 < \cdots < t_n$. (Due to the finite amount of time required for a sensor to sweep the surveillance region, measurements are generally made asynchronously, i.e., not at the same time, so that a time tag is associated with each measurement.) At each time t_k the sensor produces a sequence of measurements $Z(k) = \{z_{i_k}^k\}_{i_k=1}^{M_k}$ where each $z_{i_k}^k$ is a vector of noise contaminated measurements. The actual type of measurement varies with the sensor. For example, a two dimensional radar measures range and azimuth of each potential target ($z_{i_k}^k = (r_{i_k}^k, \theta_{i_k}^k)^T$), a three dimensional radar that measures range, azimuth, and elevation ($z_{i_k}^k = (r_{i_k}^k, \theta_{i_k}^k, \phi_{i_k}^k)^T$), a three dimensional radar with Doppler measures these and the time derivative of range ($z_{i_k}^k = (r_{i_k}^k, \theta_{i_k}^k, \phi_{i_k}^k, \frac{dr_{i_k}^k}{dt})^T$), and a two dimensional passive sensor measures the azimuth and elevation angle ($z_{i_k}^k = (\theta_{i_k}^k, \phi_{i_k}^k)$). Some of the measurements may be false, and the number of targets and which measurement emanates from which target are not known a priori.

The problems then are to determine the number of targets, which measurements go with which targets and which are false (i.e., the data association problem), and to estimate the state of each target given a sequence of measurements that emanate from that target. As part of posing the data association problem, one must estimate the state of a potential target given a particular sequence of measurements from the data sets $\{Z(k)\}_{k=1}^{N}$. Thus, the two problems of data association and state estimation are inseparable parts of the same problem. Before proceeding to the data association problem, we briefly discuss the form of the governing equations of motion and the relation of the measurements to the state of the target.

Given a sequence of measurements $\{z(k)\}_{k=0}^{n}$ postulated to belong to a particular target, the discrete form of the dynamics of the target and measurement equation are generally assumed to be governed by a state-space system of the form

$$\begin{aligned} x(k+1) &= f_k(x(k)) + g_k(x(k))w(k) \\ z(k) &= h_k(x(k)) + v(k) \end{aligned} \tag{2.1}$$

where $x(k)$ is a vector of n state variables at time t_k, $\{w(k)\}$ and $\{v(k)\}$ are independent white noise sequences of normal random variables with zero mean and covariance matrices $Q(k)$ and $R(k)$, respectively, $z(k)$ represents the measurement at time k associated with this particular target and $h_k(x(k))$ maps the state $x(k)$ to the measurement space $z(k)$. Although we shall not make explicit use of this discrete system, one generally uses Kalman filtering for linear problems and a nonlinear filtering technique such as extended Kalman filtering for nonlinear problems [1].

The combinatorial optimization problem that governs a large number of data association problems in multitarget tracking and multisensor data fusion [2–4, 7–12, 16, 17, 20–27] is generally posed as

$$\text{Maximize}\left\{ \frac{P(\Gamma = \gamma | Z^N)}{P(\Gamma = \gamma^0 | Z^N)} \mid \gamma \in \Gamma^* \right\} \tag{2.2}$$

where Z^N represents N data sets (2.3), γ is a partition of indices of the data ((2.4) and (2.5)), Γ^* is the finite collection of all such partitions (2.5), Γ is a discrete random element defined on Γ^*, γ^0 is a reference partition, and $P(\Gamma = \gamma | Z^N)$ is the posterior probability of a partition γ being true given the data Z^N. Each of these terms must be defined. Having done so, the objective will be to formulate a reasonably general class of these data association problems (2.2) as multidimensional assignment problems (2.16).

In the above surveillance example, the data sets were measurements made of the surveillance region. To allow for more general types of data such as tracks and measurements, we shall call the elements in the data sets *reports* [25]. Thus, let $Z(k)$ denote a data set of M_k reports $\{z_{i_k}^k\}_{i_k=1}^{M_k}$ and let Z^N denote the cumulative data set of N such sets defined by

$$Z(k) = \{z_{i_k}^k\}_{i_k=1}^{M_k} \quad \text{and} \quad Z^N = \{Z(1), \ldots, Z(N)\}, \tag{2.3}$$

respectively. In multisensor data fusion and multitarget tracking the data sets $Z(k)$ may represent different classes of objects. For track initiation in multitarget tracking the objects are measurements that must be partitioned into tracks and false alarms. In our formulation of track maintenance (Section 3.3) one data set will be tracks and remaining data sets will be measurements which are assigned to existing tracks, as false measurements, or to initiating tracks. In sensor level tracking, the objects to be fused are tracks [27]. In centralized fusion [8, 27], the objects may all be measurements that represent targets or false reports, and the problem is to determine which measurements emanate from a common source.

The next task is to define what is meant by a partition of the cumulative data set Z^N in (2.3). Since this definition is to be independent of the actual data in the cumulative data set Z^N, we first define a partition of the indices in Z^N. Let

$$I^N = \{I(1), I(2), \ldots, I(N)\} \text{ where } I(k) = \{i_k\}_{i_k=1}^{M_k} \tag{2.4}$$

denote the indices in the data sets (2.3). A partition γ of I^N and the collection of all such partitions Γ^* is defined

$$\gamma = \{\gamma_1, \ldots, \gamma_{n(\gamma)} \mid \gamma_i \neq \emptyset \text{ for each } i\}, \tag{2.5a}$$

$$\gamma_i \cap \gamma_j = \emptyset \text{ for } i \neq j \tag{2.5b}$$

$$I^N = \cup_{j=1}^{n(\gamma)} \gamma_j, \tag{2.5c}$$

$$\Gamma^* = \{\gamma \mid \gamma \text{ satisfies (2.5a)} - \text{(2.5c)}\}. \tag{2.5d}$$

Here, γ_i in (2.5a) will be called a track, so that $n(\gamma)$ denotes the number of tracks (or elements) in the partition γ. A $\gamma \in \Gamma^*$ is called a set partitioning of the indices I^N if properties (2.5a)–(2.5c) are valid, a set covering of I^N if property (2.5b) is omitted but the other two properties (2.5a) and (2.5c) are retained, and a set packing if (2.5c) is omitted but (2.5a) and (2.5b) are retained [18]. A partition $\gamma \in \Gamma^*$ of the index set I^N induces a partition of the data Z^N via

$$Z_\gamma = \{Z_{\gamma_1}, \ldots, Z_{\gamma_{n(\gamma)}}\} \quad \text{where} \quad Z_{\gamma_i} = \{\{z_{i_k}^k\}_{i_k \in \gamma_i}\}_{k=1}^N. \tag{2.6}$$

Clearly, $Z_{\gamma_i} \cap Z_{\gamma_j} = \emptyset$ for $i \neq j$ and $Z^N = \cup_{j=1}^{n(\gamma)} Z_{\gamma_j}$. Each Z_{γ_i} will be called a track of data.

We now come to the key independence assumptions that allow one to convert the objective function in the optimization problem (2.2) to an equivalent linear one. Assume $(\hat{Z}^N, \Gamma)$ is a random element defined on the space $(S \times \Gamma^*)$ with probability "density" $p(Z^N, \Gamma = \gamma)$ where Γ^* is defined above, S is the range space of the data Z^N, $\hat{Z}^N$ is a continuous random variable and Γ is discrete. The assumptions are made on the form of the density function $p(Z^N|\Gamma = \gamma)$ and the probability $P_\Gamma(\Gamma = \gamma)$ in Bayes' formula $P(\Gamma = \gamma \mid Z^N) = \frac{1}{p(Z^N)} p(Z^N \mid \Gamma = \gamma)$ $P_\Gamma(\Gamma = \gamma)$ [19] and can be summarized as

$$p(Z^N \mid \Gamma = \gamma) = \prod_{\gamma_i \in \gamma} p(Z_{\gamma_i} \mid \Gamma = \gamma) \tag{2.7a}$$

$$p(Z_{\gamma_i} \mid \Gamma = \gamma) = p(Z_{\gamma_i} \mid \Gamma = \omega) \quad \text{for all} \quad \gamma \text{ and } \omega \in \Gamma^* \tag{2.7b}$$

$$P_\Gamma(\Gamma = \gamma) = C \prod_{i=1}^{n(\gamma)} G(\gamma_i) \tag{2.7c}$$

where C is a constant independent of the partition $\gamma \in \Gamma^*$ and G is a probability distribution on the set of tracks γ_i in (2.5). (If $T = \{\lambda_i \mid \lambda_i$ is a track occurring in some partition $\lambda \in \Gamma^*\}$ and $|T|$ denotes the cardinality of T, then $G(\lambda_i) \geq 0$ and $\sum_{i=1}^{|T|} G(\lambda_i) = 1$.) A probabilistic framework that illustrates each of these assumptions in (2.7) is developed in Appendix B. In particular the existence of a distribution on Γ^* having the desired form (2.7c) is established there. We also note that since for each $\gamma \in \Gamma^*$, Z_γ corresponds to a partition of the data into $n(\gamma)$ feasible tracks of data, assumption (2.7a) says that the $n(\gamma)$ tracks of data, $\{Z_{\gamma_1}, \ldots, Z_{\gamma_{n(\gamma)}}\}$, are independent if γ is the truetrack. Of course, there may be dependence between the reports within a single track. Equation (2.7b) further states that these tracks are independent across all partitions of the data.

By equation (2.7b), the irrelevant conditioning on the partition γ in the expression for $p(Z_{\gamma_i}|\Gamma = \gamma)$ can be dropped and replaced by $p(Z_{\gamma_i})$. This, along with Bayes' formula, and (2.7) yield the following sequence of decompositions

$$P(\Gamma = \gamma \mid Z^N) = \frac{1}{p(Z^N)} p(Z^N \mid \Gamma = \gamma) P_\Gamma(\Gamma = \gamma) \tag{2.8a}$$

$$= \frac{1}{p(Z^N)} \left[\prod_{i=1}^{n(\gamma)} p(Z_{\gamma_i}) \right] P_\Gamma(\Gamma = \gamma) \tag{2.8b}$$

$$= \frac{C}{p(Z^N)} \prod_{\gamma_i \in \gamma} p(Z_{\gamma_i}) G(\gamma_i) \tag{2.8c}$$

Equation (2.8a) is Bayes' formula. If $P_\Gamma(\Gamma = \gamma) = C$, a constant, over all partitions $\gamma \in \Gamma^*$, then (2.8b) is proportional to the likelihood function. Otherwise, the expansion (2.8c) follows from (2.7c). Thus (2.8c) includes both the likelihood function with the identification $C = 1$ and $G(\gamma_i) = 1$ and the posterior function for a more general probability distribution G on the tracks. This completes the review of the required parts Morefield's formulation [17]. The next objective is to refine this formulation in a way that is amenable to the assignment problem.

For notational convenience in representing tracks, we add a *zero index* to each of the index sets $I(k)$ $(k = 1, \ldots, N)$ in (2.4), a *dummy report* z_0^k to each of the data sets $Z(k)$ in (2.3), and require that each

$$\begin{aligned} \gamma_i &= (i_1, \ldots, i_N) \\ Z_{\gamma_i} &= Z_{i_1 \cdots i_n} \equiv (z_{i_1}^1, \ldots, z_{i_N}^N) \end{aligned} \tag{2.9}$$

where i_k and $z_{i_k}^k$ can now assume the values of 0 and z_0^k, respectively. The dummy report z_0^k serves several purposes in the representation of missing data, false reports, initiating tracks, and terminating tracks. If Z_{γ_i} is missing an actual report from the data set $Z(k)$, then $\gamma_i = (i_1, \ldots, i_{k-1}, 0, i_{k+1}, \ldots, i_N)$ and

$Z_{\gamma_i} = \{z^1_{i_1}, \ldots, z^{k-1}_{i_{k-1}}, z^k_0, z^{k+1}_{i_{k+1}}, \ldots, z^N_{i_N}\}$. A false report $z^k_{i_k}$ $(i_k > 0)$ is represented by $\gamma_i = (0, \ldots, 0, i_k, 0, \ldots, 0)$ and $Z_{\gamma_i} = \{z^1_0, \ldots, z^{k-1}_0, z^k_{i_k}, z^{k+1}_0, \ldots, z^N_0\}$ in which there is but one actual report. The partition γ^0 of the data in which all reports are declared to be false reports is defined by

$$Z_{\gamma^0} = \{Z_{0\cdots 0 i_k 0 \cdots 0} \equiv \left(z^1_0, \ldots, z^{k-1}_0, z^k_{i_k}, z^{k+1}_0, \ldots, z^N_0\right) \mid i_k = 1, \ldots, M_k; \\ k = 1, \ldots, N\}. \tag{2.10}$$

If each data set $Z(k)$ represents a scan of measurements, a track that initiates on scan $m > 1$ will contain only the dummy report z^k_0 from each of the data sets $Z(k)$ for each $k = 1, \ldots, m-1$. Likewise, a track that terminates on scan m would have only the dummy report from each of the data sets for $k > m$. These representations are discussed further in Section 3 for both track initiation and track maintenance.

The use of the 0-1 variable

$$z_{i_1 \cdots i_N} = \begin{cases} 1 & \text{if } (i_1, \ldots, i_N) \in \gamma, \\ 0 & \text{otherwise}, \end{cases} \tag{2.11}$$

yields an equivalent characterization of a partition (2.5) and (2.9) as a solution of the equations

$$\sum_{(i_1,\ldots,i_{k-1},i_{k+1},\ldots,i_N)=(0,\ldots,0)}^{(M_1,\ldots,M_{k-1},M_{k+1},\ldots,M_N)} z_{i_1 \cdots i_N} = 1 \text{ for } i_k = 1, \cdots, M_k \text{ and } k = 1, \cdots, N. \tag{2.12}$$

With this characterization of a partition of the cumulative data set Z^N as a set of equality constraints (2.12), we now proceed to the formulation of the multidimensional assignment problem.

For the *first derivation* of the assignment problem, observe that for $\gamma_i = \left(i_1 \cdots i_n\right)$ as in (2.9) and the reference partition (2.10), the expansion (2.8c) can be written as

$$\frac{P(\Gamma = \gamma \mid Z^N)}{P(\Gamma = \gamma^0 \mid Z^N)} \equiv L_\gamma \equiv \prod_{(i_1 \cdots i_N) \in \gamma} L_{i_1 \cdots i_N} \tag{2.13a}$$

where

$$L_{i_1 \cdots i_N} = \frac{p(Z_{i_1 \cdots i_N}) G(Z_{i_1 \cdots i_N})}{\prod_{k=1, i_k \neq 0}^{N} p(Z_{0 \cdots 0 i_k 0 \cdots 0}) G(Z_{0 \cdots 0 i_k 0 \cdots 0})}. \tag{2.13b}$$

Here the index i_k in the denominator corresponds to the k^{th} index of $Z_{i_1 \cdots i_N}$ in the numerator. Next define

$$c_{i_1 \cdots i_N} = -\ln L_{i_1 \cdots i_N}, \tag{2.14}$$

so that

$$-\ln\left[\frac{P(\gamma\,|\,Z^N)}{P(\gamma^0\,|\,Z^N)}\right] = \sum_{(i_1,\ldots,i_N)\in\gamma} c_{i_1\cdots i_N}. \tag{2.15}$$

Thus in view of the of the characterization of a partition (2.5) and (2.6) specialized by (2.9) as a solution of the equation (2.12), the independence assumptions (2.7) and the expansion (2.8), problem (2.2) is equivalently characterized as the following N-dimensional assignment problem:

$$\begin{aligned}
&\text{Minimize} && \sum_{i_1=0}^{M_1}\cdots\sum_{i_N=0}^{M_N} c_{i_1\cdots i_N} z_{i_1\cdots i_N}\\
&\text{Subject To}: && \sum_{i_2=0}^{M_2}\cdots\sum_{i_N=0}^{M_N} z_{i_1\cdots i_N} = 1, \quad i_1 = 1,\ldots,M_1,\\
& && \sum_{i_1=0}^{M_1}\cdots\sum_{i_{k-1}=0}^{M_{k-1}}\sum_{i_{k+1}=0}^{M_{k+1}}\cdots\sum_{i_N=0}^{M_N} z_{i_1\cdots i_N} = 1,\\
& && \qquad \text{for} \quad i_k = 1,\ldots,M_k \text{ and } k = 2,\ldots,N-1,\\
& && \sum_{i_1=0}^{M_1}\cdots\sum_{i_{N-1}=0}^{M_{N-1}} z_{i_1\cdots i_N} = 1, \quad i_N = 1,\ldots,M_N\\
& && z_{i_1\cdots i_N} \in \{0,1\} \text{ for all } i_1,\ldots,i_N,
\end{aligned} \tag{2.16}$$

where $c_{0\cdots 0}$ is arbitrarily defined to be zero. Note that the definition of a partition (2.5) and (2.9) and the 0-1 variable $z_{i_1\cdots i_N}$ in (2.11) imply $z_{0\cdots 0} = 0$. (If $z_{0\cdots 0}$ is not preassigned to zero and $c_{0\cdots 0}$ is defined arbitrarily, then $z_{0\cdots 0}$ is determined directly from the value of $c_{0\cdots 0}$, since it does not enter the constraints other than being a zero-one variable.) Also, each cost coefficient with exactly one nonzero index is zero (i.e., $c_{0\cdots 0 i_k 0\cdots 0} = 0$ for all $i_k = 1,\ldots,M_k$ and $k = 1,\ldots,N$) due to use of the normalizing partition γ^0 in the likelihood ratio in (2.2) and (2.13b).

The *second derivation* begins with the following decompositions:

$$-\ln P(\Gamma = \gamma\,|\,Z^N) = \sum_{(i_1,\ldots,i_N)\in\gamma} \hat{c}_{i_1\cdots i_N} - \ln\left[\frac{C}{p(Z^N)}\right] \tag{2.17a}$$

where

$$\hat{c}_{i_1\cdots i_N} = -\ln\left[p(Z_{i_1\cdots i_N})G(Z_{i_1\cdots i_N})\right] \tag{2.17b}$$

and

$$-\ln P(\Gamma = \gamma^0\,|\,Z^N) = \sum_{k=1}^{N}\sum_{i_k=1}^{M_k} b_{0\cdots 0 i_k 0\cdots 0} - \ln\left[\frac{C}{p(Z^N)}\right] \tag{2.18a}$$

where

$$b_{0\cdots 0i_k0\cdots 0} = -\ln\left[p(Z_{0\cdots 0i_k0\cdots 0})G(Z_{0\cdots 0i_k0\cdots 0}))\right]. \tag{2.18b}$$

Although $Z_{0\cdots 0}$ does not occur in any partition, we arbitrarily set $b_{0\cdots 0} = \hat{c}_{0\cdots 0} = 0$. Thus

$$-\ln\left[\frac{P(\Gamma = \gamma \mid Z^N)}{P(\Gamma = \gamma^0 \mid Z^N)}\right] = \sum_{(i_1,\ldots,i_N)\in\gamma} \hat{c}_{i_1\cdots i_N} - \sum_{k=1}^{N}\sum_{i_k=1}^{M_k} b_{0\cdots 0i_k0\cdots 0}$$

$$= \sum_{(i_1,\ldots,i_N)\in\gamma}\left(\hat{c}_{i_1\cdots i_N} - \sum_{k=1}^{N} b_{0\cdots 0i_k0\cdots 0}\right) \tag{2.19}$$

where the last equality is valid because each nonzero index i_k arises in exactly one $Z_{i_1\cdots i_N} \in Z_\gamma$. Now with the identification

$$c_{i_1\cdots i_N} = \hat{c}_{i_1\cdots i_N} - \sum_{k=1}^{N} b_{0\cdots 0i_k0\cdots 0}, \tag{2.20}$$

the formulation of the multidimensional assignment problem (2.16) follows exactly as in that of the first derivation.

The third derivation is based on the following invariance property of the multidimensional assignment problem (2.16).

Invariance property. *Let $N > 1$ and $M_k > 0$ for $k = 1,\ldots,N$, and assume $\hat{c}_{0\cdots 0} = b_{0\cdots 0} = 0$. Then the minimizing solution of the following multidimensional assignment problem is independent of any choice of $b_{0\cdots 0i_k0\cdots 0}$ for $i_k = 1,\ldots,M_k$ and $k = 1,\ldots,N$.*

$$\text{Minimize } \sum_{i_1=0}^{M_1}\cdots\sum_{i_N=0}^{M_N}\left(\hat{c}_{i_1\cdots i_N} - \sum_{k=1}^{N} b_{0\cdots 0i_k0\cdots 0}\right) z_{i_1\cdots i_N}$$

$$\text{Subject To : } \sum_{i_2=0}^{M_2}\cdots\sum_{i_N=0}^{M_N} z_{i_1\cdots i_N} = 1, \quad i_1 = 1,\ldots,M_1,$$

$$\sum_{i_1=0}^{M_1}\cdots\sum_{i_{k-1}=0}^{M_{k-1}}\sum_{i_{k+1}=0}^{M_{k+1}}\cdots\sum_{i_N=0}^{M_N} z_{i_1\cdots i_N} = 1,$$

$$\text{for } \quad i_k = 1,\ldots,M_k \text{ and } k = 2,\ldots,N-1,$$

$$\sum_{i_1=0}^{M_1}\cdots\sum_{i_{N-1}=0}^{M_{N-1}} z_{i_1\cdots i_N} = 1, \quad i_N = 1,\ldots,M_N$$

$$z_{i_1\cdots i_N} \in \{0,1\} \text{ for all } i_1,\ldots,i_N. \tag{2.21}$$

Proof. Let A(z) and B(z)denote the objective function in (2.21) with b removed and b present, respectively. Let $\hat{z}$ and $\hat{y}$ both be solutions of the constraints in

(2.21). It suffices to show that $A(\hat{z}) \leq A(\hat{y})$ if and only if $B(\hat{z}) \leq B(\hat{y})$. To see this, observe

$$\begin{aligned}
B(\hat{z}) &= \sum_{i_1=0}^{M_1} \cdots \sum_{i_N=0}^{M_N} \left(\hat{c}_{i_1 \cdots i_N} - \sum_{k=1}^{N} b_{0 \cdots 0 i_k 0 \cdots 0} \right) \hat{z}_{i_1 \cdots i_N} \\
&= \sum_{i_1=0}^{M_1} \cdots \sum_{i_N=0}^{M_N} \hat{c}_{i_1 \cdots i_N} \hat{z}_{i_1 \cdots i_N} - \sum_{k=1}^{N} \sum_{i_1=0}^{M_1} \cdots \sum_{i_N=0}^{M_N} b_{0 \cdots 0 i_k 0 \cdots 0} \hat{z}_{i_1 \cdots i_N} \\
&= A(\hat{z}) - \sum_{k=1}^{N} \sum_{i_1=0}^{M_1} \cdots \sum_{i_k=0}^{M_k} \cdots \sum_{i_N=0}^{M_N} b_{0 \cdots 0 i_k 0 \cdots 0} \hat{z}_{i_1 \cdots i_N} \\
&= A(\hat{z}) - \sum_{k=1}^{N} \sum_{i_k=1}^{M_k} \left(\sum_{i_1=0}^{M_1} \cdots \sum_{i_{k-1}=0}^{M_{k-1}} \sum_{i_{k+1}=0}^{M_{k+1}} \cdots \sum_{i_N=0}^{M_N} b_{0 \cdots 0 i_k 0 \cdots 0} \hat{z}_{i_1 \cdots i_N} \right) \\
&= A(\hat{z}) - \sum_{k=1}^{N} \sum_{i_k=0}^{M_k} b_{0 \cdots 0 i_k 0 \cdots 0}.
\end{aligned}$$

Similarly, $B(\hat{y}) = A(\hat{y}) - \sum_{k=1}^{N} \sum_{i_k=0}^{M_k} b_{0 \cdots 0 i_k 0 \cdots 0}$. Thus, $B(\hat{z}) - B(\hat{y}) = A(\hat{z}) - A(\hat{y})$, which implies the result. □

The *third derivation* begins with the optimization problem

$$\text{Maximize } \left\{ P(\Gamma = \gamma \mid Z^N) \mid \gamma \in \Gamma^* \right\}$$

where the objective function is not normalized and $P(\Gamma = \gamma \mid Z^N)$ is given by (2.17a). Since $\ln \left[\frac{C}{p(Z^N)} \right]$ in (2.17a) is, by assumption, independent of any partition, the multidimensional assignment problem is posed just as in (2.16) except with c replaced by $\hat{c}$. The above Invariance property is then used to zero out all cost coefficients of the form $\hat{c}_{0 \cdots i_k \cdots 0}$ by setting $b_{0 \cdots i_k \cdots 0} = \hat{c}_{0 \cdots i_k \cdots 0}$.

Having completed the derivation of the assignment problem (2.16), several remarks are in order. The definition of a partition in (2.5) and (2.6) implies that each actual report belongs to at most one track of reports Z_{γ_i} in a partition Z_γ of the cumulative data set. One can modify this to allow multi-assignments of one, some, or all the actual reports. The assignment problem (2.16) is changed accordingly. For example, if $z_{i_k}^k$ is to be assigned no more than, exactly, or no less than $n_{i_k}^k$ times, then the "$= 1$" in the constraint (2.16) is changed to "$\leq, =, \geq n_{i_k}^k$," respectively. In making these changes, one must pay careful attention to the independence assumptions (2.7). (These assumptions need not be valid in many potential applications such as sensor level tracking [9, 27], and this may lead to the possibility of nonlinearity in the objective function in (2.16).)

Next, the normalization of $P(\Gamma = \gamma \mid Z^N)$ by dividing by $P(\Gamma = \gamma^0 \mid Z^N)$ where the partition γ_0 of all false reports is useful for at least two reasons. First it allows the cancellation of the term $\frac{C}{p(Z^N)}$ in (2.8c). Since $p(Z^N)$ can have dimensions, the result is that the ratio $L_{i_1\cdots i_N}$ in (2.13) is dimensionless. This normalization also yields a convenient preprocessing technique called fine gating as discussed in the next paragraph.

The complexity of this problem (2.16) makes its formulation and solution formidable. We have already mentioned that it is NP-hard [13]. Note also that the computation of all the cost coefficients in (2.16) can be a time consuming task. For example, six scans of measurements with one hundred measurements per scan requires the computation of one trillion cost coefficients for track initiation. Thus preprocessing is essential for reducing the complexity. One class of methods for tracking and sensor fusion problems is called gating [4, 7], which is generally composed two parts-coarse and fine gating. The philosophy of the former is to compute only those cost coefficients that are feasible for the underlying physical problem, thereby removing unlikely pairings of measurements. Techniques vary widely and are discussed in the book by Blackman [7]. If a sequence of points $\{z_{i_1}^1, \ldots, z_{i_k}^N\}$ passes this coarse gating procedure, then the corresponding coefficients $c_{i_1\cdots i_N}$ is computed. If $c_{i_1\cdots i_N}$ is positive, then one preassigns (before solving (2.16)) the corresponding 0-1 variable $z_{i_1\cdots i_N} = 0$, since the assignment of the reports in the corresponding $\{z_{i_1}^1, \ldots, z_{i_k}^N\}$ as false reports leads to a lower cost. (The normalization in (2.2) where γ^0 is the partition of all false reports is fundamental to this conclusion.) We emphasize that this does not mean that each of the actual reports in $\{z_{i_1}^1, \ldots, z_{i_k}^N\}$ is declared to be a false report in advance of solving the problem. Another commonly used complexity reducing technique is that of clustering [2, 8] which is used to decompose the problem into a number of smaller independent problems. Thus all preprocessing techniques used to reduce the computational costs associated with formulating and solving data association problems in tracking and sensor fusion can also be used to reduce the complexity of the assignment problem (2.16).

3. Assignment formulation of multiple hypothesis tracking for multitarget tracking

The principal method for multiscan processing in multitarget tracking and for sensor fusion is an explicit enumeration scheme called multiple hypothesis tracking (MHT) [7, 25]. The goal in this section is to demonstrate that a data association problem formulated for multitarget tracking in terms of multiple hypothesis tracking can be formulated as a multidimensional assignment problem as developed in the previous section. In particular, we convert the general data association problem as posed by Reid [25] to a multidimensional assignment problem in Subsection 3.1. This scenario does not include maneuvering targets; however, a similar formula for maneuvering targets can be converted from the MHT formulation to the assignment formulation by using the formulas derived by Kurien [16]. The underlying assumptions in the work of Reid [25], Kurien [16], and Blackman [7] are the same as those independence assumptions expressed in

(2.7). Track initiation and a sliding window formulation for track maintenance, as developed in our previous work [22] for a special case, are presented in the second and third subsections, respectively, for this more general formulation.

Before beginning this conversion, we briefly discuss some similarities and differences in the multiple hypothesis tracking and assignment formulations. First, one can view the multiple hypothesis tracking enumeration algorithm as a method of solving the multidimensional assignment problem (2.16). As discussed at the end of the previous section, all the preprocessing that is generally used for multiple hypothesis tracking can be used for other methods for solving the assignment problem (2.16). The particular algorithms that we have developed for obtaining near-optimal solutions of the assignment problem (2.16) [20, 23, 24] make use of all of these preprocessing techniques, except the explicit pruning of noncompetitive partitions. We do prune tracks of reports from all partitions via the gating techniques discussed at the end of the previous section. Since the Lagrangian relaxation based algorithms [20, 24] are not enumerative in nature, pruning of whole partitions is not done explicitly. The algorithms [24] instead compute both the correct partition and score during the course of solving the problem. Thus the pruning of the unlikely partitions is done implicitly in the course of solving the problem.

3.1. Likelihood calculations

The development of an expression for the posterior probability $P(\Gamma = \gamma \mid Z^N)$ in (2.2) for the problem described by Reid [25] is quite tedious. The final expression, modified to include terminating tracks, is given in (3.4), and it is this expression that is needed to convert the problem to an assignment problem. However, for completeness we briefly outline the derivation and relegate to the appendix the derivation of some of the formulas. The expression for this posterior probability is developed sequentially (scan by scan) [25] and then converted to the form (2.13).

Recall from (2.3) that the k^{th} data set of reports is $Z(k) = \{z_i^k\}_{i_k=1}^{M_k}$ and the cumulative set of k such data sets is $Z^k = \{Z(1), \dots, Z(k)\}$ so that $Z^k = Z(k) \cup Z^{k-1}$. The customary notation here [25] is to define $\Omega^k = \{\Omega_j^k \mid j = 0, 1, \dots, I_k\}$ to be the set of hypotheses about the feasible partitions of the cumulative set of measurements Z^k into tracks and false alarms. In the notation of the previous section, each Ω_j^k represents the event that a partition $\gamma^k \in \Gamma^{k*}$ is true. (Here, the dependence of the γ on the number of scans k has been explicitly included as a superscript.) We shall follow this customary notation. The hypothesis Ω_0^k is that all reports are false. Let $\Omega_{\psi_l}^{k-1}$ denote that specific hypothesis in Ω^{k-1} that produces Ω_l^k, and let $\psi_l(k)$ denote the hypothesis that indicates the specific status of all targets postulated by $\Omega_{\psi_l}^{k-1}$ at the scan time t_k and the specific origin of all reports received at scan time t_k. Thus

$$\Omega_l^k = \psi_l(k) \cup \Omega_{\psi_l}^{k-1} \text{ and } Z^k = Z(k) \cup Z^{k-1} \tag{3.1}$$

Using this and Bayes' rule [19], one can write

$$P(\Omega_l^k \mid Z^k) = p(Z(k) \mid \Omega_l^k, Z^{k-1}) P(\psi_l(k) \mid \Omega_{\psi_l}^{k-1}, Z^{k-1}) \times P(\Omega_{\psi_l}^{k-1} \mid Z^{k-1}) \left[\frac{p(Z^{k-1})}{p(Z^k)} \right] \tag{3.2}$$

The term $\left[\frac{p(Z^{k-1})}{p(Z^k)}\right]$ is a normalizing factor, which will be removed when $P(\Omega_l^k \mid Z^k)$ is normalized by dividing this expression by the probability of all false reports, i.e., $P(\Omega_0^k|Z^k)$.

We partition the reports just as in the work of Reid [25] in the following way. Assume the parent hypothesis $\Omega_{\psi_l}^{k-1}$ postulates that τ^k targets were extended from scan $k-1$ to scan k. Of the total number of measurements M_k received on scan k, ν^k reports originate from new targets with an associated probability mass function $\mu_\nu^k(\nu^k)$, f_k reports are false reports with probability mass function $\mu_f^k(f^k)$, and the remaining δ_k reports are associated with existing targets. Of the τ^k targets that exist after scan $k-1$, χ^k of these τ^k targets are terminated and are not observed (on scan k), δ^k of the $\tau^k - \chi^k$ nonterminated targets are detected, and $\tau^k - \chi^k - \delta^k$ nonterminated targets are not detected. (Thus, the total number of targets that exist after scan k is $\tau^{k+1} = \tau^k - \chi^k + \nu^k$.) Next let P_χ^k denote the probability oftermination and P_d^k, the probability of detection on scan k.

For ease of reference, here are these numbers, probabilities, and the definition of some required indicator functions:

P_χ^k is the probability of termination on scan k;
P_d^k is the probability of detection on scan k;
$z_{i_k}^k$ is measurement i_k from scan k;
δ^k is the number of measurements on scan k originating from previously established tracks;
ν^k is the number of new targets detected on scan k;
f^k is the number of false alarms on scan k;
M_k is the total number of measurements on scan k;
τ^k is the number of targets that were extended from scan $k-1$ to scan k;
χ^k is the number of terminated (and nondetected) targets on scan k;

$$f_i^k = \begin{cases} 1, & \text{if } z_i^k \text{ is a false alarm;} \\ 0, & \text{otherwise;} \end{cases}$$

$$\nu_i^k = \begin{cases} 1, & \text{if } z_i^k \text{ is a new target;} \\ 0, & \text{otherwise;} \end{cases}$$

$$\delta_i^k = \begin{cases} 1, & \text{if } z_i^k \text{ belongs to an existing track;} \\ 0, & \text{otherwise;} \end{cases}$$

$$\Delta_{ij} = \begin{cases} 1, & \text{if } i = j; \\ 0, & \text{otherwise.} \end{cases} \tag{3.3}$$

In addition, let $p_t^k(z_{i_k}^k|\Omega_l^k, Z^{k-1})$ represent the likelihood that the report $z_{i_k}^k$ originated from a target modeled by the state space model (2.1), $p_\nu^k(z_{i_k}^k \mid \Omega_l^k, Z^{k-1})$ represent the likelihood that the report originated from a new source; and, $p_f^k(z_{i_k}^k \mid \Omega_l^k, Z^{k-1})$ represent the likelihood the report represents a false alarm.

The development of the specific formulas for the first and second factors in (3.2) are developed in the appendix. The development of these formulas requires the same independence assumptions in (2.7). The substitution of (A.1) and (A.5) in Appendix A into (3.2) yields (see also equation (A.6))

$$\begin{aligned} P(\Omega_l^k \mid Z^k) = &\left\{ \frac{\nu^k! f^k!}{M_k!} \mu_f^k(f^k) \mu_\nu^k(\nu^k) \right\} \left\{ (P_\chi^k)^{\chi^k} \left[(1 - P_\chi^k)(1 - P_d^k)\right]^{\tau^k - \delta^k - \chi^k} \right\} \\ &\times \prod_{i=1}^{M_k} \left\{ \left[(1 - P_\chi^k) P_d^k p_t^k(z_{i_k}^k \mid \Omega_l^k, Z^{k-1})\right]^{\delta_{i_k}^k} \left[p_\nu^k(z_{i_k}^k \mid \Omega_l^k, Z^{k-1})\right]^{\nu_{i_k}^k} \right. \\ &\left. \times \left[p_f^k(z_{i_k}^k \mid \Omega_l^k, Z^{k-1})\right]^{f_{i_k}^k} \right\} P(\Omega_{\psi_l}^{k-1} \mid Z^{k-1}) \left[\frac{p(Z^{k-1})}{p(Z^k)}\right] \end{aligned} \tag{3.4}$$

where the indicator functions f_i^k, ν_i^k, δ_i^k, and Δ_{ij} are defined in equation (3.3).

Normalizing this expression by dividing by $P(\Omega_l^0 \mid Z^k)$ where Ω_0^k = {All False Alarms} yields

$$\begin{aligned} \frac{P(\Omega_l^k \mid Z^k)}{P(\Omega_0^k \mid Z^k)} = &\left\{ \frac{\nu^k! f^k!}{M_k!} \frac{\mu_f^k(f^k) \mu_\nu^k(\nu^k)}{\mu_f^k(M_k) \mu_\nu^k(0)} \right\} \\ &\times \left\{ (P_\chi^k)^{\chi^k} \left[(1 - P_\chi^k)(1 - P_d^k)(1 - P_m^k)\right]^{\tau^k - \delta^k - \chi^k} \right\} \\ &\times \prod_{i=1}^{M_k} \left\{ \left[(1 - P_\chi^k) P_d^k \frac{p_t^k(z_{i_k}^k \mid \Omega_l^k, Z^{k-1})}{p_f^k(z_{i_k}^k \mid \Omega_0^k, Z^{k-1})}\right]^{\delta_{i_k}^k} \right. \\ &\left. \times \left[\frac{p_\nu^k(z_{i_k}^k \mid \Omega_l^k, Z^{k-1})}{p_f^k(z_{i_k}^k \mid \Omega_0^k, Z^{k-1})}\right]^{\nu_{i_k}^k} \right\} \frac{P(\Omega_{\psi_l}^{k-1} \mid Z^{k-1})}{P(\Omega_0^{k-1} \mid Z^{k-1})}. \end{aligned} \tag{3.5}$$

If, in addition, $\mu_f^k(f^k) = \exp(-\lambda_f^k) \frac{(\lambda_f^k)^{f^k}}{f^k!}$ and $\mu_\nu^k(\nu^k) = \exp(-\lambda_\nu^k) \frac{(\lambda_\nu^k)^{\nu^k}}{\nu^k!}$ are Poisson probability mass functions, then

$$\begin{aligned} \frac{P(\Omega_l^k \mid Z^k)}{P(\Omega_0^k \mid Z^k)} = &\frac{P(\Omega_{\psi_l}^{k-1} \mid Z^{k-1})}{P(\Omega_0^{k-1} \mid Z^{k-1})} \left\{ (P_\chi^k)^{\chi^k} \left[(1 - P_\chi^k)(1 - P_d^k)\right]^{\tau^k - \delta^k - \chi^k} \right\} \\ &\times \prod_{i=1}^{M_k} \left\{ \left[\frac{(1 - P_\chi^k) P_d^k p_t^k(z_{i_k}^k \mid \Omega_l^k, Z^{k-1})}{\lambda_f^k p_f^k(z_{i_k}^k \mid \Omega_0^k, Z^{k-1})}\right]^{\delta_{i_k}^k} \right. \end{aligned}$$

$$\times \left[\frac{\lambda_\nu^k p_\nu^k(z_{i_k}^k | \Omega_l^k, Z^{k-1})}{\lambda_f^k p_f^k(z_{i_k}^k | \Omega_0^k, Z^{k-1})} \right]^{\nu_{i_k}^k} \Bigg\} \tag{3.6}$$

Note that the Poisson assumption allows one to satisfy (2.7).

The formula for $\frac{P(\Omega_l^k | Z^k)}{P(\Omega_0^k|Z^k)}$ has now been developed recursively for $k \geq 1$ as in the work of Reid [25]. The task now is to convert this expression to the form (2.13). For this purpose, consider N sets of reports as a single entity. For a track $Z_{i_1 \cdots i_N} \equiv \{z_{i_1}^1, \ldots, z_{i_N}^N\}$ define

$$P_\phi^k = \begin{cases} P_\chi^k, & \text{if track } Z_{i_1\cdots i_N} \text{ terminates at scan } k; \\ (1 - P_\chi^k)(1 - P_d^k), & \text{if track } Z_{i_1\cdots i_N} \text{ has a missed detection} \\ & \quad \text{on scan } k; \\ 1, & \text{otherwise.} \end{cases} \tag{3.7}$$

Each track of reports $Z_{i_1 \cdots i_N}$ in a feasible partition $\gamma \in \Gamma^*$ is scored as

$$\begin{aligned} L_{i_1 i_2 \cdots i_N} &\equiv L(Z_{i_1 i_2 \cdots i_N}) \equiv L(z_{i_1}^1, \ldots, z_{i_N}^N) \\ &= \prod_{k=1}^{N} \{P_\phi^k\}^{\Delta_{0 i_k}} \left\{ \left[\frac{(1 - P_\chi^k) P_d^k p_t^k(z_{i_k}^k \mid Z_{i_1 \cdots i_N})}{\lambda_f^k p_f^k(z_{i_k}^k \mid Z_{0\cdots 0 i_k 0 \cdots 0})} \right]^{\delta_{i_k}^k} \right. \\ &\quad \left. \times \left[\frac{\lambda_\nu^k p_\nu^k(z_{i_k}^k \mid Z_{i_1 \cdots i_N})}{\lambda_f^k p_f^k(z_{i_k}^k | Z_{0 \cdots 0 i_k 0 \cdots 0})} \right]^{\nu_{i_k}^k} \right\}^{(1-\Delta_{0 i_k})} \\ &\quad \text{provided at least two of the indices in } \{i_1, i_2, \ldots, i_N\} \\ &\quad\quad \text{are nonzero} \end{aligned} \tag{3.8}$$

and

$$L_{0\cdots 0 i_k 0 \cdots 0} \equiv 1 \text{ provided } \{0, \ldots, 0, i_k, 0, \ldots, 0\} \in \gamma$$

where the conditioning now is only on the assumption that the collection of reports in $Z_{i_1 \cdots i_N} \equiv \{z_{i_1}^1, \ldots, z_{i_N}^N\}$ represent a true event. The conditioning in the probability density for a false report $p_f^k(z_{i_k}^k \mid Z_{0\cdots 0 i_k 0 \cdots 0})$ in the denominator is on the assumption that $z_{i_k}^k$ is a false report which is represented by $Z_{0\cdots 0 i_k 0 \cdots 0}$. Thus, an equivalent expression for (3.6) is

$$\frac{P(\Omega_l^k \mid Z^k)}{P(\Omega_0^k \mid Z^k)} = \prod_{\{i_1 i_2 \cdots i_N\} \in \gamma} L_{i_1 i_2 \cdots i_N}. \tag{3.9}$$

Several remarks about this expression are in order. Each track of reports $Z_{i_1 \cdots i_N} = \{z_{i_1}, \ldots, z_{i_N}\}$ carries with it such information as the time at which the track initiates or terminates. For example, if this particular target initiates on scan $r > 1$ and terminates on scan $s < N$, then $i_k = 0$ and $\Delta_{0i_k} = 1$ for $k = 1, \ldots, r-1$ and for $k = s, \ldots, N$, so that

$$\{P_\phi^k\}^{\Delta_{0i_k}} = \begin{cases} 1 & \text{if } k = 1, \ldots, r-1 \text{ and } k = s+1, \ldots, N; \\ (1 - P_\chi^k)(1 - P_d^k) & \text{if } r < k < s \text{ and there is a missed} \\ & \quad \text{detection on scan } k \\ P_\chi^k & \text{if } k = s \\ 1 & \text{otherwise.} \end{cases} \tag{3.10}$$

Next, consider a particular sequence of reports $\{z_{i_1}, \ldots, z_{i_N}\}$. If there is only one nonzero index, then this point must be a false report by definition. If there are two or more points assigned together as a track, then the underlying assumption is that there are no false reports in the list. Thus each point z_{i_k} is either a dummy, associated with a the track defined by the collection, or may correspond to a new target. This last situation may cause some ambiguity in the definition of $L_{i_1 i_2 \cdots i_N}$. This is generally decided from auxiliary information. Here is one example of how such auxiliary reasoning is frequently used. If z_{i_k} is the first non-dummy report and the subsequent reports in the list $\{z_{i_1}, \ldots, z_{i_N}\}$ are dynamically feasible with it, then $\nu_{i_k}^k = 1$ and $\delta_{i_k}^k = 0$. For this same list of reports, one would not consider any subsequent points in the list as new sources. Finally, track terminations can also be based on a track length model as in the work of Stein and Blackman [7].

3.2. *Track initiation*

If one specializes the development in §3.1 to the situation in which each object in the data set $Z(k)$ is a measurement or observation for $k = 1, \ldots, N$, then the problem of partitioning the observations into tracks and false alarms can be posed as the multidimensional assignment problem (2.16) where

$$c_{i_1 \cdots i_N} = -\ln L_{i_1 \cdots i_N} \tag{3.11}$$

and $L_{i_1 \cdots i_N}$ is defined by equation (3.8).

3.3. *Track maintenance using a sliding window*

Suppose now that the observations on P *previous* scans (of observations) have been partitioned into tracks and false alarms and that K *new* scans of observations are to be added. One approach to solving the resulting data association problem is formulate the problem as a track initiation problem with $P + K$ scans. This is the previously mentioned *batch* approach. The *deferred logic approach* adopted here is to treat the track extension problem within the framework of a window sliding over the observation sets. First assume that the scans of observations are partitioned into three components: D *discarded* scans of observations, R *retained* scans of observations from the P *previously* processed scans, and K *new* scans of observations. Thus the number of scans in the sliding window is $N = R + K$ while the number of *discarded* scans is $D = P - R$.

Let M_0 denote the number of confirmed tracks previously constructed from the discarded and retained regions that are present at the beginning of the window of N scans. Tracks that initiate within the window are not included in this list. Tracks that have been declared as terminated at some scan within the present window are still included in this number M_0. Thus a tentative decision about a terminated track in the window may be changed given new information. A track is actually terminated only if the termination point occurs prior to the current window, i.e., the termination occurs in the discarded region. Other than these restrictions, the M_0 tracks are obtained as the solution of a previous problem assignment problem. Another possibility is to include all tracks in the best K such solutions or even all feasible tracks. For $i_0 = 1, \ldots, M_0$ the i_0^{th} such track is denoted by T_{i_0} and the $(N+1)$-tuple $\{T_{i_0}, z_{i_1}^1, \ldots, z_{i_N}^N\}$ will denote a track T_{i_0} plus a set of observations or measurements $\{z_{i_1}^1, \ldots, z_{i_N}^N\}$, actual or dummy, that are feasible with the track T_{i_0}. The $(N+1)$-tuple $\{T_0, z_{i_1}^1, \ldots, z_{i_N}^N\}$ will denote a track that initiates in the sliding window. Note the use once again of a dummy variable T_0. A false report in the sliding window is one with all but one non-zero index in the $(N+1)$-tuple $\{T_0, z_{i_1}^1, \ldots, z_{i_N}^N\}$

The hypothesis about a partition $\gamma \in \Gamma^*$ being true is now conditioned on the truth of the M_0 tracks entering the N-scan window. (Thus the assignments prior to this sliding window are fixed.) The likelihood function is given by $L_\gamma = \prod_{\{T_{i_0}, z_{i_1}, \ldots, z_{i_N}\} \in \gamma} L_{i_0 i_1 \cdots i_N}$, where $L_{i_0 i_1 \cdots i_N} = L_{T_{i_0}} L_{i_1 \cdots i_N}$, $L_{T_{i_0}}$ is the composite likelihood from the discarded scans just prior to the first scan in the window for $i_0 > 0$, $L_{T_0} = 1$, and $L_{i_1 \cdots i_N}$ is defined as in (3.8) for the N-scan window. ($L_{T_0} = 1$ is used for any tracks that initiate in the sliding window.) Thus the track extension problem can be formulated as Maximize $\{L_\gamma \mid \gamma \in \Gamma^*\}$. With the same convention as in Section 2, a feasible partition is one which is defined by the properties in (2.5) and (2.9). Analogously, the definition of the zero-one variable

$$z_{i_0 i_1 \cdots i_N} = \begin{cases} 1 & \text{if } \{T_{i_0}, z_{i_1}^1, \ldots, z_{i_N}^N\} \text{ is assigned as a unit,} \\ 0 & \text{otherwise,} \end{cases}$$

and the corresponding cost for the assignment of the sequence$\{T_{i_0}, z_{i_1}, \ldots, z_{i_N}\}$ to a track by $c_{i_0 i_1 \cdots i_N} = -\ln L_{i_0 i_1 \cdots i_N}$ yield the following multidimensional assignment formulation of the data association problem for track maintenance:

$$\text{Minimize} \sum_{i_0=0}^{M_0} \cdots \sum_{i_N=0}^{M_N} c_{i_0 \cdots i_N} z_{i_0 \cdots i_N}$$

$$\text{Subject To} \sum_{i_1=0}^{M_1} \cdots \sum_{i_N=0}^{M_N} z_{i_0 \cdots i_N} = 1, \quad i_0 = 1, \ldots, M_0,$$

$$\sum_{i_0=0}^{M_0}\sum_{i_2=0}^{M_2}\cdots\sum_{i_N=0}^{M_N} z_{i_0 i_1\cdots i_N} = 1, \quad i_1 = 1,\ldots,M_1,$$

$$\sum_{i_0=0}^{M_0}\cdots\sum_{i_{k-1}=0}^{M_{k-1}}\sum_{i_{k+1}=0}^{M_{k+1}}\cdots\sum_{i_N=0}^{M_N} z_{i_0\cdots i_N} = 1,$$

$$\text{for} \quad i_k = 1,\ldots,M_N \text{ and } k = 2,\ldots,N-1,$$

$$\sum_{i_0=0}^{M_0}\cdots\sum_{i_{N-1}=0}^{M_{N-1}} z_{i_0\cdots i_N} = 1, \quad i_N = 1,\ldots,M_N,$$

$$z_{i_0\cdots i_N} \in \{0,1\} \text{ for all } i_0,\ldots,i_N. \tag{3.12}$$

Note that the association problem involving N scans of observations is an N-dimensional assignment problem for track initiation and an $(N+1)$-dimensional one for track maintenance.

4. An example problem

Consider the problem of many targets traveling in two dimensional space according to the constant acceleration model $x(t,\alpha) = (x_1(t,\alpha), x_2(t,\alpha))$ where

$$\begin{aligned} x_1(t,\alpha) &= x_{01} + tv_1 + \frac{t^2}{2}a_1 \\ x_2(t,\alpha) &= x_{02} + tv_2 + \frac{t^2}{2}a_2 \end{aligned} \tag{4.1}$$

where the parameters $\alpha = (x_{01}, v_1, a_1, x_{02}, v_2, a_2)$ identify a particular target. At a discrete set of *scan times* $\{t_k\}_{k=1}^N$ $(t_1 \le t_2 \le \cdots \le t_N)$ a radar located at the origin in this Cartesian space observes error contaminated ranges and angles of the targets in the measurement space which is a circle with radius R. Some measurements are spurious and some measurements of true targets are missed. At time t_k, the radar is assumed to return the set of noise contaminated measurements $\{z_{i_k}^k\}_{k=1}^{M_k}$, where M_k is the number of measurements and $z_{i_k}^k = (r_{i_k}^k, \theta_{i_k}^k)$. To every scan, a dummy measurement z_0^k is added to represent missed detections. Each measurement $(r_{i_k}^k, \theta_{i_k}^k)$ is related to the true observable $H(x(t_k,\alpha))$ by

$$\begin{pmatrix} r_{i_k}^k \\ \theta_{i_k}^k \end{pmatrix} = H(x(t_k,\alpha)) + \begin{pmatrix} e_r^k \\ e_\theta^k \end{pmatrix} \tag{4.2}$$

where e_r^k and e_θ^k are independent zero-mean Gaussian random variables with standard deviations σ_r^k and σ_θ^k, respectively. The true observable $H(x(t,\alpha))$ is related to the track $x(t,\alpha) \equiv (x_1(t,\alpha), x_2(t,\alpha))$ by

$$H(x(t,\alpha)) = \begin{pmatrix} \sqrt{x_1(t,\alpha)^2 + x_2(t,\alpha)^2} \\ \arctan\left[\frac{x_2(t,\alpha)}{x_1(t,\alpha)}\right] \end{pmatrix}. \tag{4.3}$$

(If the definition of arctan is based on the principal angle, then the appropriate shifts in $\theta^k_{i_k}$ must be made.) If the measurement corresponds to a false alarm or new target, then

$$\begin{pmatrix} r^k_{i_k} \\ \theta^k_{i_k} \end{pmatrix} = \begin{cases} \begin{pmatrix} w^k_r \\ w^k_\theta \end{pmatrix} & \text{if the measurement is spurious,} \\ \begin{pmatrix} u^k_r \\ u^k_\theta \end{pmatrix} & \text{if the measurement arises from a new source,} \end{cases} \tag{4.4}$$

where the random sequences w^k and u^k have some assumed densities p^k_f and p^k_ν, respectively. A common assumption is that each of w^k_r, w^k_θ, u^k_r, u^k_θ is uniformly distributed so that

$$p^k_f(w^k) = \begin{cases} \frac{1}{2\pi R} & \text{if } 0 \le w^k_r \le R \text{ and } -\pi < w^k_\theta \le \pi \\ 0 & \text{otherwise} \end{cases} \tag{4.5}$$

A similar expression would hold for p^k_ν. The number of false alarms and new targets are assumed to be generated at each time interval $[t_{k-1}, t_k]$ according to a Poisson distributions with expected numbers λ^k_f and λ^k_ν, respectively.

Under these assumptions the likelihood expression in (3.8) simplifies to

$$\begin{aligned} L_{i_1 i_2 \cdots i_N} = \prod_{k=1}^{N} \{P^k_\phi\}^{\Delta_{0 i_k}} & \left\{ \left[\frac{(1-P^k_\chi)P^k_d p^k_t(z^k_{i_k} \mid Z_{i_1\cdots i_n})}{\lambda^k_f p^k_f(z^k_{i_k} \mid Z_{i_1\cdots i_n})} \right]^{\delta^k_{i_k}} \right. \\ & \times \left. \left[\frac{\lambda^k_\nu p^k_\nu(z^k_{i_k} \mid Z_{i_1\cdots i_n})}{\lambda^k_f p^k_f(z^k_{i_k} \mid Z_{i_1\cdots i_n})} \right]^{\nu^k_{i_k}} \right\}^{(1-\Delta_{0 i_k})} \\ & \text{provided at least two of the indices in } \{i_1, i_2, \ldots, i_N\} \\ & \quad \text{are nonzero} \end{aligned} \tag{4.6}$$

and

$$L_{0\cdots 0 i_k 0\cdots 0} \equiv 1 \text{ provided } \{0,\ldots,0,i_k,0,\ldots,0\} \in \gamma.$$

where

$$P^k_\phi = \begin{cases} P^k_\chi, & \text{if track } Z_{i_1\cdots i_N} \text{ terminates at scan } k; \\ (1-P^k_\chi)(1-P^k_d), & \text{if track } Z_{i_1\cdots i_N} \text{ has a missed detection} \\ & \quad \text{on scan } k; \\ 1, & \text{otherwise.} \end{cases} \tag{4.7}$$

In writing down this score, one assumes that the targets dynamics (4.1) are known, so that the measurement errors can be computed using the densities

$$p^k_t(z^k_{i_k} \mid Z_{i_1\cdots i_n}) = \frac{1}{2\pi\sigma_r(t_k)\sigma_\theta(t_k)} \exp\left\{ -\frac{1}{2}\left[z^k_{i_k} - H_k(x(t_k,\alpha))\right]^T \right.$$

$$\times\, \Sigma_{r\theta}^{-1}(t_k)\left[z_{i_k}^k - H_k(x(t_k,\alpha))\right] \Bigg\} \quad (4.8)$$

However, it is precisely these tracks that are to be identified, so that they must be estimated from the given the measurements $\{z_{i_1}^1, \ldots, z_{i_k}^N\}$. We do this by replacing the parameters α by the maximum likelihood estimate

$$\hat{\alpha} = \text{Arg Max } L_{i_1 \cdots i_n}(x(\cdot,\alpha)) = \text{Arg Min } c_{i_1 \cdots i_n}(x(\cdot,\alpha)) \quad (4.9)$$

which can be equivalently characterized as the solution to the nonlinear least squares problem

$$\begin{aligned}\hat{\alpha} &= \text{Arg Min } c_{i_1 \cdots i_N}(x(\cdot,\alpha)) \\ &= \text{Arg Min } \sum_{k=1}^{N}(1-\Delta_{0i_k})\left[z_{i_k}^k - H_k(x(t_k,\alpha))\right]^T \\ &\quad \times \Sigma_{r\theta}^{-1}(t_k)\left[z_{i_k}^k - H_k(x(t_k,\alpha))\right], \end{aligned} \quad (4.10)$$

where $\Sigma_{r\theta}(t) = diag(\sigma_r^2(t), \sigma_\theta^2(t))$.

In a given problem one may need to solve thousands or even hundreds of thousands of these time consuming problems. To obtain a good initial estimate, one could transform the errors and measurements from polar coordinates to Cartesian coordinates. The measurements are easily transformed via $(x_{i_k}^k, y_{i_k}^k) = (r_{i_k}^k \cos(\theta_{i_k}^k), r_{i_k}^k \sin(\theta_{i_k}^k))$. Similarly, $e_x = (r_t + e_r)\cos(\theta_t + e_\theta) - r_t\cos(\theta_t)$ and $e_y = (r_t + e_r)\sin(\theta_t + e_\theta) - r_t\sin(\theta_t)$. Assuming the errors are small, one can show [7] that the means are zero to first order in the Taylor series expansion of e_r and e_θ about zero. Furthermore, the covariance matrix can be approximated to within this approximation by

$$\begin{aligned}\Sigma_{xy}(t) &= \begin{pmatrix} \sigma_x^2(t) & \sigma_{xy}(t) \\ \sigma_{xy}(t) & \sigma_y^2(t) \end{pmatrix} \\ &= \begin{pmatrix} \sigma_r^2\cos^2\theta_t + r_t^2\sigma_\theta^2\sin^2\theta_t & \frac{1}{2}(\sigma_r^2 - r_t^2\sigma_\theta^2)\sin 2\theta_t \\ \frac{1}{2}(\sigma_r^2 - r_t^2\sigma_\theta^2)\sin 2\theta_t & \sigma_r^2\sin^2\theta_t + r_t^2\sigma_\theta^2\cos^2\theta_t \end{pmatrix}\end{aligned} \quad (4.11)$$

where all quantities are functions of time and $r_t = r(x(t,\alpha))$ and $\theta_t = \theta(x(t,\alpha))$. Unfortunately, r_t and θ_t are not known since the parameters in α are unknown. As another approximation, these are replaced by the corresponding measured quantities at $t = t_k$. Then the parameters in α are estimated by the solution of (linear) least squares problem

$$\begin{aligned}\hat{\alpha} = \text{Arg Min } &\sum_{k=1}^{N}(1-\Delta_{0i_k})\begin{pmatrix} x_{i_k}^k - x_1(t_k,\alpha) \\ y_{i_k}^k - x_2(_k,\alpha) \end{pmatrix}^T \\ &\times \Sigma_{xy}^{-1}(t_k)\begin{pmatrix} x_{i_k}^k - x_1(t_k,\alpha) \\ y_{i_k}^k - x_2(t_k,\alpha) \end{pmatrix}\end{aligned} \quad (4.12)$$

In many applications one does not refine these estimates using (4.10) [7].

A couple of final remarks are in order. First, the estimation problem (4.9) or (4.10) is expressed as a batch nonlinear least squares problem; however, the parameters could be estimated recursively (i.e., sequentially) as the scans of measurements become available. Next, note that this model fits the state space representation (2.1). If let the vector $y(k) = (y_1(k), y_2(k), y_3(k), y_4(k), y_5(k), y_6(k))^T$ denote the position x_1, velocity $\frac{dx_1}{dt}$, acceleration $\frac{d^2x_1}{dt^2}$, position x_2, velocity $\frac{dx_2}{dt}$, and acceleration $\frac{d^2x_2}{dt^2}$, respectively. Then (3.1) can be written as

$$\begin{bmatrix} y_1(k+1) \\ y_2(k+1) \\ y_3(k+1) \\ y_4(k+1) \\ y_5(k+1) \\ y_6(k+1) \end{bmatrix} = \begin{bmatrix} 1 & \Delta t_k & \frac{\Delta t_k^2}{2} & 0 & 0 & 0 \\ 0 & 1 & \Delta t_k & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \Delta t_k & \frac{\Delta t_k^2}{2} \\ 0 & 0 & 0 & 0 & 1 & \Delta t_k \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_1(k) \\ y_2(k) \\ y_3(k) \\ y_4(k) \\ y_5(k) \\ y_6(k) \end{bmatrix}$$

$$\begin{bmatrix} r_{i_k}^k \\ \theta_{i_k}^k \end{bmatrix} = \begin{bmatrix} \sqrt{y_1(t_k,\alpha)^2 + y_4(t_k,\alpha)^2} \\ \arctan\left[\frac{y_4(t_k,\alpha)}{y_1(t_k,\alpha)}\right] \end{bmatrix} + \begin{bmatrix} e_r^k \\ e_\theta^k \end{bmatrix} \tag{4.13}$$

The dynamics in the first equation (4.13) does not contain noise. If one includes noise as in (2.1), then an extended Kalman filter, iterated extended Kalman filter, or some other nonlinear estimation technique [1] must be used to estimate the state of the system.

5. Multisensor fusion

The use of multiple sensors, through more varied information, has the potential to greatly enhance target identification and state estimation. Again, the central problem is that of data association, i.e., that of determining which measurements emanate from common targets or sources and which measurements are false. When one considers techniques for solving the corresponding data association problems, we note once again that MHT is a method that permeates the field [2, 7, 9, 27]. This section considers two examples, but before proceeding, a brief review of some of the issues involved in the design of multisensor fusion algorithms is in order.

Two important decisions involve sensor location, i.e., distributed or collocated, and the level of data association, i.e., central level, sensor level, or hybrids of these two. Sensors spatially distributed enhances geographical diversity and survivability. What's more, the combination of active and passive sensors can use the geometry of separation to identify targets. For example, several passive sensors can be used to great benefit; however, the ghosting problem [8] now appears. Other difficulties include the communication complexity and registration or alignment errors [2]. In collocated sensor placement, i.e., in the same place or on the same platform, sensor diversity is generally chosen to provide complementary information. For example, a two dimensional radar provides accurate range and moderately accurate azimuth measurements, whereas an infrared sensor provides highly accurate azimuth. The combination can yield highly accurate range and azimuth measurements. Communication complexities and registration errors can be greatly reduced for collocated sensors.

The choices of the level of association range from sensor- to central-level tracking with hybrids in between. In the former case, each sensor forms tracks from its own measurements and then the tracks from N sensors are fused in a central location. (One may think of N-dimensional assignments here as assigning tracks from the first sensor to tracks from the second sensor and so forth to tracks from the N^{th} sensor.) Once the matching is complete, one then combines the tracks with appropriate modification in the statistics. One of the difficulties here is that the usual error independence assumptions (2.7) are not valid and this introduces additional complexity. Another problem is that the combined track estimates tend to be worse than in central-level fusion of measurements only. Arguments in favor of this method are reduced communications costs and higher survivability since the sensors maintain their own tracks. At the other extreme is centralized fusion in which sensors send measurements to a central processing unit where they are combined to give superior position measurements. From the point of view of track estimation, this method seems to be superior. The difficulties are data association problem, communication costs between the sensor and central processing unit, and the loss of the tracking capability if the central processing unit becomes inoperative.

In the next two subsections examples of non-collocated and collocated sensors are discussed. These examples assume synchronous measurements. The problem of asynchronous measurements will be addressed in future work.

5.1. Non-collocated sensors

In this first subsection we closely follow the work of Deb, Pattipati, and Bar-Shalom [11]. We consider N non-collocated sensors (i.e., all sensors are spatially separated from one another) with the number N and the locations $\{x_k\}_{k=1}^{N}$ known. The number and location of the targets are unknownlocations with the unknown number M and locations $\{x^t\}_{t=1}^{M}$. For a three dimensional scenario the notation $x_k = (x_{1k}, x_{2k}, x_{3k})$ and $x^t = (x_1^t, x_2^t, x_3^t)$ will be used to denote the x, y, and z cartesian components of the location of the sensor and the target, respectively. Each of the sensors may be one of three types: a 3D radar, a 2D radar, or 2D passive sensor. The passive sensor measures the azimuth angle and elevation angle of each potential target t, i.e. $z_{i_k}^k = [\theta_{kt}, \phi_{kt}[$; the 2D radar measures azimuth and range, i.e., $z_{i_k}^k = [r_{kt}, \theta_{kt}]$; and, a 3D radar measures all three, i.e., $z_{i_k}^k = [r_{kt}, \theta_{kt}, \phi_{kt}[$. The k^{th} sensor makes M_k actual measurements and we add a dummy variable for a missed detection and denote the corresponding data set by $Z(k) = \{z_{i_k}^k\}_{i_k=0}^{M_k}$ where z_0^k is the dummy variable. The measurements are made synchronously with the following statistical properties

$$z_{i_k}^k = \begin{cases} H(x_k, x^t) + v_{i_k}^k, & \text{if } z_{i_k}^k \text{ is from a true target;} \\ w_{i_k}^k, & \text{if } z_{i_k}^k \text{ is a spurious measurement;} \end{cases} \tag{5.1}$$

where $H(x_k, x^t)$ is the true observable, $v_{i_k}^k \sim N(0, \Sigma_k)$, and the density of the spurious measurement is assumed uniform and is given by

$$p_{w_{i_k}^k}(w) = \frac{1}{\Psi_k}$$

where Ψ_k is the field of view. Finally, let P_d^k denote the probability of detection of the k^{th} sensor.

Assuming equal prior probabilities (see (2.7) and (2.8)), the likelihood ratio $L_{i_1 \cdots i_N}$ in (2.13) can be written as [11]

$$\begin{aligned} L_{i_1 \cdots i_n}(x^t) &= \prod_{k=1}^{N} \frac{\left[P_d^k p(z_{i_k}^k \mid x^t)\right]^{1-\Delta_{0i_k}} \left[1 - P_d^k\right]^{\Delta_{0i_k}}}{\prod_{k=1, i_k \neq 0}^{N} \Psi_k^{-1}} \\ &= \prod_{k=1}^{N} \left[P_d^k \Psi_k p(z_{i_k}^k \mid x^t)\right]^{1-\Delta_{0i_k}} \left[1 - P_d^k\right]^{\Delta_{0i_k}} \end{aligned} \tag{5.2}$$

Here the conditioning in the probability density $p(z_{i_k}^k \mid x^t)$ is upon the measurement $z_{i_k}^k$ arising from a target at location x^t.

The data association problem of matching the measurements from all sensors to the targets is precisely the multidimensional assignment problem formulated in (2.16) where $c_{i_1 \cdots i_N} = -\ln L_{i_1 \cdots i_n}(x^t)$. Since x^t is unknown, it is replaced by its maximum likelihood estimate

$$\hat{x}^t = \text{Arg Max} L_{i_1 \cdots i_n}(x^t). \tag{5.3a}$$

This $\hat{x}^t$ is the solution of the statistically weighted nonlinear least squares problem

$$\text{Minimize } \sum_{k=1}^{N} (1 - \Delta_{0i_k}) \left[z_{i_k}^k - H(x_k, x^t)\right]^T \Sigma_k^{-1} \left[z_{i_k}^k - H(x_k, x^t)\right] \tag{5.3b}$$

As an example, $H(x_k, x^t)$ for a 3D-radar is given by

$$H(x_k, x^t) = \begin{bmatrix} r_{kt} \\ \theta_{kt} \\ \phi_{kt} \end{bmatrix} = \begin{bmatrix} \sqrt{\Delta x_{1k}^2 + \Delta x_{2k}^2 + \Delta x_{3k}^2} \\ \arctan\left[\frac{\Delta x_{2k}}{\Delta x_{1k}}\right] \\ \arctan \frac{\Delta x_{3k}}{\sqrt{\Delta x_{1k}^2 + \Delta x_{2k}^2}} \end{bmatrix} \tag{5.4}$$

where $(\Delta x_{1k}, \Delta x_{2k}, \Delta x_{3k}) = (x^t - x_{1k}, y^t - x_{2k}, z^t - x_{3k})$ and (x^t, y^t, z^t) is determined in the course of solving the nonlinear least squares problem (5.3b).

5.2. Multiple platforms: an example

Consider $2N$ spatially separated platforms such that on each platform one as a 2D radar measuring range and azimuth and a passive sensor measuring azimuth and elevation. The location $x_k = (x_{1k}, x_{2k}, x_{3k})$ of each platform is known; however, the group of M targets (M unknown) and unknown locations $\{x^t\}_{t=1}^{M}$ are observed by the sensors on the various platforms. For a three dimensional scenario the notation $x_k = (x_{1k}, x_{2k}, x_{3k})$ and $x^t = (x^t, y^t, z^t)$ will be used. The passive sensor measures the azimuth angle and elevation angle of each potential target t, i.e. $z_{i_k}^k = [\theta_{kt}, \phi_{kt}]$; the 2D radar measures azimuth and range, i.e., $z_{i_k}^k = [r_{kt}, \theta_{kt}]$. The k^{th} sensor makes $M_k$actual measurements and we add a dummy variable for a missed detection and denote the corresponding data set $Z(k) = \{z_{i_k}^k\}_{i_k=0}^{M_k}$ where z_0^k is the dummy variable. The statistical properties of these measurements are defined by as in (5.1), and the problem is formulated as in (5.2). A minor difference occurs in the computation of x^t. To be precise, let sensors $2k-1$ and $2k$ denote the 2D radar and passive sensors on platform k. Then $\hat{x}^t$ is the solution of the statistically weighted nonlinear least squares problem

$$\text{Minimize} \sum_{k=1}^{2N} (1 - \Delta_{0i_k}) \left[z_{i_k}^k - H(x_k, x^t)\right]^T \Sigma_k^{-1} \left[z_{i_k}^k - H(x_k, x^t)\right] \tag{5.5}$$

As an example, $H(x_{2k-1}, x^t)$ for a 2D-radar is given by

$$H(x_{2k-1}, x^t) = \begin{bmatrix} \sqrt{\Delta x_{1k}^2 + \Delta x_{2k}^2 + \Delta x_{3k}^2} \\ \arctan\left[\frac{\Delta x_{2k}}{\Delta x_{1k}}\right] \end{bmatrix} \tag{5.6a}$$

and that for the passive sensor by

$$H(x_{2k}, x^t) = \begin{bmatrix} \arctan\left[\frac{\Delta x_{2k}}{\Delta x_{1k}}\right] \\ \arctan \frac{\Delta x_{3k}}{\sqrt{\Delta x_{1k}^2 + \Delta x_{2k}^2}} \end{bmatrix} \tag{5.6b}$$

where $(\Delta x_{1k}, \Delta x_{2k}, \Delta x_{3k}) = (x^t - x_{1k}, y^t - x_{2k}, z^t - x_{3k})$ and (x^t, y^t, z^t) is determined in the course of solving the nonlinear least squares problem (5.5).

6. Concluding remarks

In this work we have formulated some general classes data association problems in multitarget tracking and multisensor data fusion as multidimensional assignment problems. The objective function is derived from a composite negative log posterior or likelihood function for each of the tracks of reports. This formulation is of sufficient generality to cover our earlier work [20–24] which used the scoring of Stein and Blackman [7], the popular multiple hypothesis tracking method introduced by Reid [25] and modified by Kurien [17] to include maneuvering targets and terminations, and the work of Deb, Pattipati, Somnath, Bar-Shalom [11] on centralized multisensor data fusion. Other sensor fusion problems that employ multiple hypothesis tracking methods, e.g., the work of C.-Y. Chong, S. Mori, and K.-C. Chang [9] on distributed multitarget multi-sensor tracking and other sensor level data fusion problems [8], remain to be investigated. The key problem that needs development is the formulation of the data association problem that incorporates asynchronous measurements from heterogeneous sensors into a system for tracking multiple objects.

Although these multidimensional assignment problems are NP-hard [14], they are fundamentally important and central to multitarget and multisensor tracking systems. The only known method for solving these problems optimally is branch and bound; however, such a method is too computationally intensive to produce real-time solutions, especially in dense scenarios. Since the objective function is generally noisy due to various sources of errors (e.g., plant noise, observation errors, and uncertainty about exact values of various probability parameters), one only needs to solve these problems to or just below the noise level in the problem. Thus high quality and near optimal solutions computed to the noise level in the problem is the goal of algorithm development. Motivated by the work of Frieze and Yadegar [13], several algorithms based on Lagrangian relaxation have been and continue to be developed [20, 24]. These algorithms are near-optimal, provide a lower and upper bound on the optimal value, and are real-time for many scenarios. Future work will be devoted to the status of the current algorithms.

Acknowledgment

I wish to thank to the referees for several helpful suggestions.

Appendix A: Probability calculations

The objective in this appendix is to derive specific formulas for the factors $p(Z(k) \mid \Omega_l^k, Z^{k-1})$ and $P(\psi_l(k) \mid \Omega_{\psi_l}^{k-1}, Z^{k-1})$ in equation (3.2). For the first factor, note that each report is assigned to a previous track, to a new source, or to a false alarm. Using the symbols defined in (3.3), the likelihood of the reports $Z(k)$ given the association hypothesis is

$$p(Z(k) \mid \Omega_l^k, Z^{k-1}) = \prod_{i_k=1}^{M_k} \left\{ \left[p_t^k(z_{i_k}^k \mid \Omega_l^k, Z^{k-1})\right]^{\delta_{i_k}^k} \left[p_\nu^k(z_{i_k}^k \mid \Omega_l^k, Z^{k-1})\right]^{\nu_{i_k}^k} \times \left[p_f^k(z_{i_k}^k \mid \Omega_l^k, Z^{k-1})\right]^{f_{i_k}^k} \right\} \tag{A.1}$$

where $p_t^k(z_{i_k}^k \mid \Omega_l^k, Z^{k-1})$ represents the likelihood that the report originated from a target maintaining the state space model (2.1), $p_\nu^k(z_{i_k}^k \mid \Omega_l^k, Z^{k-1})$ represents the likelihood that the report originated from a new source, and $p_f^k(z_{i_k}^k \mid \Omega_l^k, Z^{k-1})$ represents the likelihood the report represents a false alarm.

The development of an expression for $P(\psi_l(k) \mid \Omega_{\psi_l}^{k-1}, Z^{k-1})$ makes extensive use of the notation and definitions in (3.3). Let $\psi_{lN}(k)$ denote the event that postulates the *numbers* $\{f^k, \nu^k, \delta^k, \chi^k\}$ defined in (3.3). Of the total number of reports M_k, ν^k reports originate from new targets with an associated probability mass function $\mu_\nu^k(\nu^k)$, f_k reports are false reports with probability mass function $\mu_f^k(f^k)$, and the remaining δ_k reports are associated with existing targets. Of the τ^k targets that exist after scan $k-1$, χ^k of these τ^k targets are terminated and are not observed (on scan k), δ^k of the $\tau^k - \chi^k$ nonterminated targets are detected, and $\tau^k - \chi^k - \delta^k$ nonterminated targets are not detected. (Thus, the total number of targets that exist after scan k is $\tau_{k+1} = \tau^k - \chi^k + \nu^k$.) Recalling that the binomial coefficient $\binom{n}{m} \equiv \frac{n!}{(n-m)!m!}$ gives the number of ways to arrange n elements taken m at a time, one can show [25]

$$P(\psi_{lN}(k) \mid \Omega_{\psi_l}^{k-1}, Z^{k-1}) = \left\{\mu_f^k(f^k)\mu_\nu^k(\nu^k)\right\}\left\{\binom{\tau^k}{\chi^k}(P_\chi^k)^{\chi^k}(1-P_\chi^k)^{\tau^k-\chi^k}\right\}$$
$$\times\left\{\binom{\tau^k-\chi^k}{\delta^k}(P_d^k)^{\delta^k}(1-P_d^k)^{\tau^k-\chi^k-\delta^k}\right\}. \tag{A.2}$$

The number of ways M_k observations can be divided into δ^k detected targets, f^k false reports, and ν^k new targets is $\binom{M_k}{\delta^k}\binom{M_k-\delta^k}{\nu^k}\binom{M_k-\delta^k-\nu^k}{f^k} = \binom{M_k}{\delta^k}\binom{M_k-\delta^k}{\nu^k}$. Also, the number of ways τ^k targets can be divided into χ^k terminated targets, δ^k detected targets, and $\tau^k-\chi^k-\delta^k$ missed targets is $\binom{\tau^k}{\chi^k}\binom{\tau^k-\chi^k}{\delta^k}\binom{\tau^k-\chi^k-\delta^k}{\tau^k-\chi^k-\delta^k} = \binom{\tau^k}{\chi^k}\binom{\tau^k-\chi^k}{\delta^k}$. Let $\psi_{lC}(k)$ denote the event within $\psi_{lN}(k)$ that designates a specific set of δ^k detected targets, f^k false reports, ν^k new targets, χ^k terminated targets, and $\tau^k-\chi^k-\delta^k$ missed targets. Assuming each such event is equally likely,

$$P(\psi_{lC}(k) \mid \psi_{lN}(k), \Omega_{\psi_l}^{k-1}, Z^{k-1}) = \left\{\binom{\tau^k}{\chi^k}\binom{\tau^k-\chi^k}{\delta^k}\right.$$
$$\left.\times\binom{M_k}{\delta^k}\binom{M_k-\delta^k}{\nu^k}\right\}^{-1} \tag{A.3}$$

From the event $\psi_{lC}(k)$, $\psi_l(k)$ represents a specific hypothesis that assigns a specific set of reports to the detected targets. The number of ways to assign δ^k detections to δ^k targets is $\delta^k!$. Thus, assuming the probability for each such assignment is the same, the probability of the hypothesis $\psi_l(k)$ given $\psi_{lC}(k)$ is

$$P(\psi_l(k) \mid \psi_{lC}(k), \Omega_{\psi_l}^{k-1}, Z^{k-1}) = \frac{1}{\delta^k!} \tag{A.4}$$

The product of the expressions in (A.2)–(A.4) yields

$$\begin{aligned} P(\psi_l(k) \mid \Omega_{\psi_l}^{k-1}, Z^{k-1}) &= P(\psi_{lN}(k) \mid \Omega_{\psi_l}^{k-1}, Z^{k-1}) P(\psi_{lC}(k) \mid \psi_{lN}(k), \Omega_{\psi_l}^{k-1}, Z^{k-1}) \\ &\quad \times P(\psi_l(k) \mid \psi_{lC}(k), \Omega_{\psi_l}^{k-1}, Z^{k-1}) \\ &= \left\{ \frac{\nu^k! f^k!}{M_k!} \mu_f^k(f^k) \mu_\nu^k(\nu^k) \right\} \left\{ (P_\chi^k)^{\chi^k} \right\} \\ &\quad \times \left\{ \left[(1 - P_\chi^k)(1 - P_d^k)\right]^{\tau^k - \delta^k - \chi^k} \left[(1 - P_\chi^k) P_d^k\right]^{\delta^k} \right\} \end{aligned} \tag{A.5}$$

The substitution of (A.1) and (A.5) into (3.2) yields the following expression for $P(\Omega_l^k \mid Z^k)$:

$$\begin{aligned} P(\Omega_l^k \mid Z^k) &= P(\Omega_{\psi_l}^{k-1} \mid Z^{k-1}) \left\{ \frac{p(Z^{k-1})}{p(Z^k)} \right\} \left\{ \frac{\nu^k! f^k!}{M_k!} \mu_f^k(f^k) \mu_\nu^k(\nu^k) \right\} \\ &\quad \times \left\{ (P_\chi^k)^{\chi^k} \left[(1 - P_\chi^k)(1 - P_d^k)\right]^{\tau^k - \delta^k - \chi^k} \right\} \\ &\quad \times \prod_{i=1}^{M_k} \left\{ \left[(1 - P_\chi^k) P_d^k p_t^k(z_{i_k}^k \mid \Omega_l^k, Z^{k-1})\right]^{\delta_{i_k}^k} \left[p_\nu^k(z_{i_k}^k \mid \Omega_l^k, Z^{k-1})\right]^{\nu_{i_k}^k} \right. \\ &\qquad \left. \times \left[p_f^k(z_{i_k}^k \mid \Omega_l^k, Z^{k-1})\right]^{f_{i_k}^k} \right\} \end{aligned} \tag{A.6}$$

Appendix B: Probabilistic framework

The objective in this appendix is to provide a probabilistic framework for the assumptions in equation (2.7). Assume $(\hat{Z}^N, \Gamma)$ is a random element defined on the space $(S \times \Gamma^*)$ with probability "density" $p(Z^N, \Gamma = \gamma)$ where Γ^* is defined in (2.5), S is the range space of the data Z^N (2.3), $\hat{Z}^N$ is a continuous random variable and Γ is discrete. (The customary notation here is to use z^N and Z^N in place of Z^N and $\hat{Z}^N$, respectively.) The joint density between $\hat{Z}^N$ and Γ is defined by

$$p(Z^N, \Gamma = \gamma) = \left[\prod_{i=1}^{n(\gamma)} p(Z_{\gamma_i}) \right] P_\Gamma(\Gamma = \gamma)$$

where we have used the decomposition $Z^N = \{Z_{\gamma_1}, \ldots, Z_{\gamma_{n(\gamma)}}\}$. Here, p is used as a generic density function defined on the tracks of data Z_{γ_i}. This is a valid probability density function since $\sum_{\gamma \in \Gamma^*} \int p(Z^N, \Gamma = \gamma) dZ^N = \sum_{\gamma \in \Gamma^*} \int \left[\prod_{i=1}^{n(\gamma)} p(Z_{\gamma_i})\right] dZ^N P_\Gamma(\Gamma = \gamma) = 1$. Note that this density is defined relative to the tracks of data independent of a partition in which this track of data occurs. (Thus assumptions (2.7a) and (2.7b) are satisfied with $p(Z_{\gamma_i} \mid \Gamma = \gamma) = p(Z_{\gamma_i})$, regardless of the partition γ in which γ_i occurs.) Then

$$P(\Gamma = \gamma \mid Z^N) = \frac{1}{p(Z^N)} p(Z^N, \Gamma = \gamma) = \frac{1}{p(Z^N)} \left[\prod_{i=1}^{n(\gamma)} p(Z_{\gamma_i}) \right] P_\Gamma(\Gamma = \gamma).$$

A framework for the assumption $P_\Gamma(\Gamma = \gamma) = C \prod_{\gamma_i \in \gamma} G(\gamma_i)$ where C is a constant independent of the partition γ is developed next. This begins with definition of a probability measure on each of the tracks γ_i occurring in the definition of the partitions γ in (2.5)

Let $T = \{$ All Tracks γ_i in (2.5)$\}$ and let G denote a probability measure on T, so that $G(\gamma_i) \geq 0$ and $\sum_{i=1}^{|T|} G(\gamma_i) = 1$ where $|T|$ denotes the cardinality of the set T. On the Cartesian product T^n, let $P^n = G \times \cdots \times G$ be the product measure defined by

$$P^n(\gamma_1, \ldots, \gamma_n) = \prod_{i=1}^{n} G(\gamma_i) \quad \text{where} \quad \gamma_i \in T.$$

On the space $\Omega = \cup_{n=1}^{N^*} T^n$ define the probability measure

$$P(\omega) = \frac{P^n(\omega)}{N^*} \quad \text{if} \quad \omega \in T^n$$

where $N^* = \sum_{k=1}^{N} M_k$ is the number of "tracks" in the largest partition γ, i.e., the partition of false reports $\gamma^0 = \{i_k \ : \ i_k = 1, \ldots, M_k; k = 1, \ldots, N\}$. Recalling that Γ^* is the collection of all partitions, we have

$$P(\Gamma^*) = \sum_{n=1}^{N^*} \frac{P^n(\Gamma^* \cap T^n)}{N^*}.$$

Let X be a random element of Ω with distribution P, i.e.,

$$P(X = x) = \frac{P^n(x)}{N^*} \text{ if } x \in T^n$$

and

$$P(X = x \mid X \in \Gamma^*) = \frac{P(X = x, X \in \Gamma^*)}{P(\Gamma^*)} = \begin{cases} \frac{P(X=x)}{P(\Gamma^*)} & \text{if } x \in \Gamma^*, \\ 0 & \text{if } x \notin \Gamma^*. \end{cases}$$

Finally, let Γ be a Γ^*-valued random element with distribution given by

$$P_\Gamma(\Gamma = \gamma) = P(X = \gamma \mid X \in \Gamma^*) = \frac{G(\gamma_1) \cdots G(\gamma_n)}{N^* P(\Gamma^*)} \text{ if } \gamma = (\gamma_1, \ldots, \gamma_n) \in T^n.$$

(This is an example of a situation in which (2.7c) is satisfied.) Then

$$P(\Gamma = \gamma \mid Z^N) = \frac{1}{p(Z^N)} p(Z^N, \Gamma = \gamma) = \frac{1}{p(Z^N)} \left[\prod_{i=1}^{n(\gamma)} p(Z_{\gamma_i}) \right] P(\Gamma = \gamma)$$

$$= \frac{1}{p(Z^N)} \left[\prod_{i=1}^{n(\gamma)} p(Z_{\gamma_i}) G(\gamma_i) \right] \frac{1}{N^* P(\Gamma^*)}.$$

References

1. B. D. O. Anderson and J. B. Moore, *Optimal Filtering*, Prentice-Hall, Englewood Cliffs, New Jersey, 1979.
2. Y. Bar-Shalom, ed., *Multitarget-Multisensor Tracking: Advanced Applications*, Artech House, Dedham, MA., 1990.
3. Y. Bar-Shalom, ed., *Multitarget-Multisensor Tracking: Applications and Advances*, Artech House, Dedham, MA., 1992.
4. Y. Bar-Shalom and T. E. Fortmann, *Tracking and Data Association*, Academic Press, Boston, 1988.
5. D. P. Bertsekas, *Linear Network Optimization:Algorithms and Codes*, The MIT Press, Cambridge, Mass., 1991.
6. D. P. Bertsekas and D. A. Castañon, "A Forward/Reverse Auction Algorithm for Asymmetric Assignment Problems," *Computational Optimization and Applications*, Vol. 1, No. 3, pp 277–298, 1992.
7. S. S. Blackman, *Multiple TargetTracking with Radar Applications*, Artech House, Dedham, MA., 1986.
8. S. S. Blackman, "Association and fusion of multiple sensor data," in [2]
9. C.-Y. Chong, S. Mori, and K.-C. Chang, "Distributed multitarget multisensor tracking," in [2].
10. I. J. Cox, J. M. Rehg, and S. Hingorani, "A Bayesian Multiple Hypothesis Approach to Contour Grouping and Segmentation," *Int. J. of Computer Vision*, Vol. 11, No. 1, pp. 55–24, 1993.
11. S. Deb, K. R. Pattipati, and Y. Bar-Shalom, "A multisensor-multitarget data association algorithm for heterogeneous systems," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 29, No. 2, pp, 560–568, 1993.
12. S. Deb, K. R. Pattipati, Y. Bar-Shalom, and H. Tsaknakis, "A new algorithm for the generalized multidimensional assignment problem," to appear in *Proc. IEEE International Conference on Systems, Man, and Cybernetics.*
13. M. R. Garvey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W. H. Freeman & Co., CA, 1979.
14. A. M. Geoffrion, "Lagrangean relaxation for integer programming," in M. L. Balinski, ed., *Mathematical Programming Study 2: Approaches to Integer Programming*, North Holland Publishing Company, Amsterdam, 1974.
15. R. Jonker and T. Volgenant "A Shortest Augmenting Path Algorithm for Dense and Sparse Linear Assignment Problems," *Computing 38*, pp. 325–340, 1987.
16. T. Kurien, "Issues in the designing of practical multitarget tracking algorithms," in [2].
17. C. L. Morefield, "Application of 0-1 integer programming to multitarget tracking problems," *IEEE Transactions on Automatic Control*, Vol. AC-22, No. 3, pp. 302–312, June, 1977.
18. G. L. Nemhauser and L. A. Wolsey, *Integer and Combinatorial Optimization*, Wiley-Interscience, New York, 1988.
19. A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, Second Edition, McGraw-Hill Book Company, New York, 1984.

20. A. B. Poore and N. Rijavec, "A Lagrangian Relaxation Algorithm for Multi-dimensional Assignment Problems Arising from Multi-target Tracking," *SIAM Journal on Optimization*, Vol. 3, No. 3, pp. 545–563, 1993.
21. A. B. Poore and N. Rijavec, "Multitarget Tracking and Multidimensional Assignment Problems," in Oliver E. Drummond, Editor, *Proceedings of the 1991 SPIE Conference on Signal and Data Processing of Small Targets*, Vol. 1481, pp. 345–356, 1991.
22. A. B. Poore, N. Rijavec, and T. Barker, "Data association for track initiation and extension using multiscan windows," in Oliver E. Drummond, Editor, *Signal and Data Processing of Small Targets 1992*, Proc. SPIE, Vol. 1698, pp. 432–441, 1992.
23. A. B. Poore and N. Rijavec, "A Numerical Study of Some Data Association Problems Arising in Multitarget Tracking," in W. W. Hager, D. W. Hearn and P. M Pardalos, editors, *Large Scale Optimization: State of the Art*, Kluwer Academic Publishers, Boston, pp. 347–370, 1994.
24. A. B. Poore and N. Rijavec, "Partitioning multiple data sets: multidimensional assignments and Lagrangian relaxation," to appear in DIMACS Series in Discrete Mathematics and Theoretical Computer Science, 1994.
25. D. B. Reid, "An algorithm for tracking multiple targets," IEEE Transactions on Automatic Control, Vol. AC-24, No. 6, pp. 843-854, December, 1979.
26. R. W. Sittler, "An Optimal Data Association Problem in Surveillance Theory," *IEEE Transactions on Military Electronics*, Vol. MIL-8, pp. 125–139, April, 1964.
27. E. Waltz and J. Llinas, *Multisensor Data Fusion*, Artech House, Boston, 1990.