An Algorithm for Tracking Multiple Targets

DONALD B. REID, MEMBER, IEEE

Abstract—An algorithm for tracking multiple targets in a cluttered environment is developed. The algorithm is capable of initiating tracks, accounting for false or missing reports, and processing sets of dependent reports. As each measurement is received, probabilities are calculated for the hypotheses that the measurement came from previously known targets in a target file, or from a new target, or that the measurement is false. Target states are estimated from each such data-association hypothesis using a Kalman filter. As more measurements are received, the probabilities of joint hypotheses are calculated recursively using all available information such as density of unknown targets, density of false targets, probability of detection, and location uncertainty. This branching technique allows correlation of a measurement with its source based on subsequent, as well as previous, data. To keep the number of hypotheses reasonable, unlikely hypotheses are eliminated and hypotheses with similar target estimates are combined. To minimize computational requirements, the entire set of targets and measurements is divided into clusters that are solved independently. In an illustrative example of aircraft tracking, the algorithm successfully tracks targets over a wide range of conditions.

I. INTRODUCTION

THE SUBJECT of multitarget tracking has application in both military and civilian areas. For instance, application areas include ballistic missile defense (reentry vehicles), air defense (enemy aircraft), air traffic control (civil air traffic), ocean surveillance (surface ships and submarines), and battlefield surveillance (ground vehicles and military units). The foremost difficulty in the application of multiple-target tracking involves the problem of associating measurements with the appropriate tracks, especially when there are missing reports (probability of detection less than unity), unknown targets (requiring track initiation), and false reports (from clutter). The key development of this paper is a method for calculating the probabilities of various data-association hypotheses. With this development, the synthesis of a number of other features becomes possible.

In addition to the above data-association capabilities, the algorithm developed in this paper contains the desirable features of multiple-scan correlation, clustering, and recursiveness. Multiple-scan correlation is the capability to use later measurements to aid in prior correlations (associations) of measurements with targets. This feature is usually found in batch-processing or track-splitting

algorithms. Clustering is the process of dividing the entire set of targets and measurements into independent groups (or clusters). Instead of solving one large problem, a number of smaller problems are solved in parallel. Finally, it is desirable for an algorithm to be recursive so that all the previous data do not have to be reprocessed whenever a new data set is received.

The algorithm can use measurements from two different generic types of sensors. The first type is capable of sending information which can be used to infer the number of targets within the area of coverage of the sensor. Radar is an example of this type of sensor. This type of sensor generates a data set consisting of one or more reports, and no target can originate more than one report per data set. (The terms "data set" and "scan" are used interchangeably in this paper to mean a set of measurements at the same time. It is not necessary that they come from a sensor that scans.) The second type of sensor does not contain this "number-of-targets" type of information. A radar detector, for example, would not detect a target unless the target's radar were on. In this case, very little can be implied about a target's existence or nonexistence by the fact that the target is not detected. Also, for the second type of sensor, individual reports are transmitted and processed one at a time, instead of in a data set. The multiple-target tracking algorithm developed here accounts for these factors by using the detection and false-alarm statistics of the sensors, the expected density of unknown targets, and the accuracy of the target estimates.

A number of complicating factors not considered in this paper include nonlinear measurements, nonlinear dynamics, maneuvering targets (abrupt and unknown change in target dynamics), requirement for an adaptive algorithm (to account for unknown statistics), some aspects of multiple sensors (problems of sensor configurations, registration, and different types of information), and time-delayed or out-of-sequence measurements. The first four factors have already been investigated extensively in the single-target case, and do not aid in illuminating the multiple-target problem. The inclusion of the last two factors would greatly increase the scope of this paper. In addition, the real-world constraints involved in implementing this algorithm are not explicitly considered.

References [1]–[8] are the basic reference papers that illustrate previously known techniques for solving the multiple-target tracking problem. The scope of each of

TABLE I
SCOPE OF CURRENT PAPERS IN MULTIPLE-TARGET TRACKING

Algorithm Characteristics	Reference							
	1	2	3	4	5	6	7	8
Multiple Targets	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Missing Measurements	Yes	Yes	Yes	No	Yes	Yes	No	Yes
False Alarms (e.g., Clutter)	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Track Initiation	No	No	No	No	Yes	Yes	No	Yes
Sensor Data Sets (e.g., Number-of-Targets Information)	Yes	Yes	Yes	No	Yes	Yes	No	No
Multiple-Scan Correlation	Yes	No	No	Yes	Yes	No	Yes	Yes
Clustering	No	No	Yes	No	No	No	No	No
Recursive (i.e., Filter)	Yes	Yes	Yes	Yes	No	Yes	Yes	No

these papers is summarized in Table I. In addition, there are a number of good papers incorporated into and referenced by these eight references which are not repeated here. A more comprehensive set of papers is included in the recent survey paper by Bar-Shalom [9]. The algorithm developed in this paper includes all the characteristics shown in Table I.

Reference [1], by Singer et al., is the culmination of several previous papers by the authors. In this reference, they develop an "N-scan filter" for one target. Whenever a set of measurements is received, a set of hypotheses is formed as to the measurement that was originated by the target. This branching process is repeated whenever a new set of measurements is received so that an ever-expanding tree is formed. To keep the number of branches to a reasonable number, all branches which have the last N measurements in common are combined. A simulation of their filter was included in the paper. The significant finding of their simulation was that, for $N = 1$, the N-scan filter gave near-optimal performance. This is significant in that the concept of track-splitting has been immediately discounted by others as being too expensive.

In [2], Bar-Shalom and Tse also treat a single target with clutter measurements. They develop the "probabilistic data association" filter that updates one target with every measurement, in proportion to the likelihood that the target originated the measurement. The filter is suboptimal in that track-splitting is not allowed (i.e., it is an $N = 0$ scan filter). In [3], Bar-Shalom extends this filter to the multiple-target case. He separates all the targets and measurements into "clusters" which can then be processed independently. He then suggests a rather complicated technique to calculate the probability of each measurement originating from each target. Compared to his technique, the derivation in this paper reduces to a relatively simple expression, as discussed in Section IV.

In [4], Alspach applies his nonlinear "Gaussian sum" filter [10] to the multitarget problem. His concepts are similar to those above; however, it should be noted that his filter is not optimal in that the density function he is trying to estimate "does not contain all the information available in the measurements since at each stage the state of the target giving rise to the nth measurement is conditioned on only the nth measurement..." and not all

measurements up to this stage. This paper is unique in estimating the type of target (a discrete state) as well as the target's continuous-valued states.

Reference [5], by Sittler, was published ahead of its time and is included here even though it is ten years older than any of the other basic references. By using very simple processes, Sittler illustrated most of the major concepts in multitarget tracking. In addition to track initiation, false alarms, and missing measurements, he included the possibility of a target ceasing to exist (track termination), a factor not covered in this paper. This possibility results in several other concepts, such as track status. If data are being received that eliminate the possibility of the track being dropped, then the track status is defined to be good.

In [6], Stein and Blackman implement and modernize most of the concepts suggested in [5]. As in [5], they retain the concept of track dropping, as well as track initiation, and derive two gates around each target. In their implementation, they choose a suboptimal sequential method of processing the data. As each set of data is received, only the most likely assignment of targets and measurements is selected.

In [7], Smith and Buechler very briefly present a branching algorithm for multiple-target tracking. By calculating the relative likelihood of each branch, they are able to eliminate unlikely branches. In calculating the likelihoods, they assume that each target is present ($P_D = 1$) and do not account for false-alarm statistics. More seriously, however, they apparently allow a target to be associated with every measurement within its gate. If measurements are within several gates, this leads to sets of data-association hypotheses that are not mutually exclusive. On the other hand, the ad hoc procedure of eliminating branches whose estimates are less than a specified distance away partially remedies this problem.

In [8], Morefield solves for the most likely data-association hypothesis (as opposed to calculating the probabilities of all the data-association hypotheses). He does so by formulating the problem in the framework of integer linear programming; as such, it is an interesting approach. His algorithm is basically a batch-processing technique. Even though a recursive version is included, it does not guarantee optimality over several time intervals as the batch-processing version does.

For the remainder of this paper, it is assumed that each target is represented by a vector x of n state variables which evolves with time according to known laws of the form

$$x(k+1) = \Phi x(k) + \Gamma w(k) \tag{1}$$

where

Φ = the state transition matrix
Γ = the disturbance matrix
w = a white noise sequence of normal random variables with zero mean and covariance Q.

These state variables are related to measurements z according to

$$z(k) = Hx(k) + v(k) \qquad (2)$$

where

H = a measurement matrix

v = a white noise sequence of normal random variables with zero mean and covariance R.

If the measurements could be uniquely associated with each target, then the conditional probability distribution of the state variables of each target is a multivariate normal distribution given by the Kalman filter [11]. The mean $\bar{x}$ and covariance $\bar{P}$ of this distribution evolve with time between measurements according to the following "time update" equations:

$$
\begin{aligned}
\bar{x}(k+1) &= \Phi \hat{x}(k) \\
\bar{P}(k+1) &= \Phi \hat{P}(k) \Phi^T + \Gamma Q \Gamma^T.
\end{aligned} \qquad (3)
$$

When a measurement is received, the conditional mean $\hat{x}$ and covariance $\hat{P}$ are given by the following "measurement update" equations:

$$
\begin{aligned}
\hat{x}(k) &= \bar{x}(k) + K \big[z(k) - H\bar{x}(k) \big] \\
\hat{P}(k) &= \bar{P} - \bar{P} H^T (H \bar{P} H^T + R)^{-1} H \bar{P} \\
K &= \hat{P} H^T R^{-1}.
\end{aligned} \qquad (4)
$$

II. OVERVIEW OF THE ALGORITHM

A flow diagram of the tracking algorithm is shown in Fig. 1. Most of the processing is done within the four subroutines shown in the figure. The CLUSTR subroutine associates measurements with previous clusters. If two or more previously independent clusters are associated because of a measurement, then the two clusters are combined into a "super cluster." A new cluster is formed for any measurement not associated with a prior cluster. As part of the initialization program, previously known targets form their own individual clusters.

The HYPGEN subroutine forms new data-association hypotheses for the set of measurements associated with each cluster. The probability of each such hypothesis is calculated and target estimates are updated for each hypothesis of each cluster.

Both the CLUSTR and HYPGEN subroutines use the REDUCE subroutine for eliminating unlikely hypotheses or combining hypotheses with similar target estimates. Once the set of hypotheses is simplified by this procedure, uniquely associated measurements are eliminated from the hypothesis matrix by the MASH subroutine. Tentative targets become confirmed targets if they were the unique origin of the eliminated measurement.
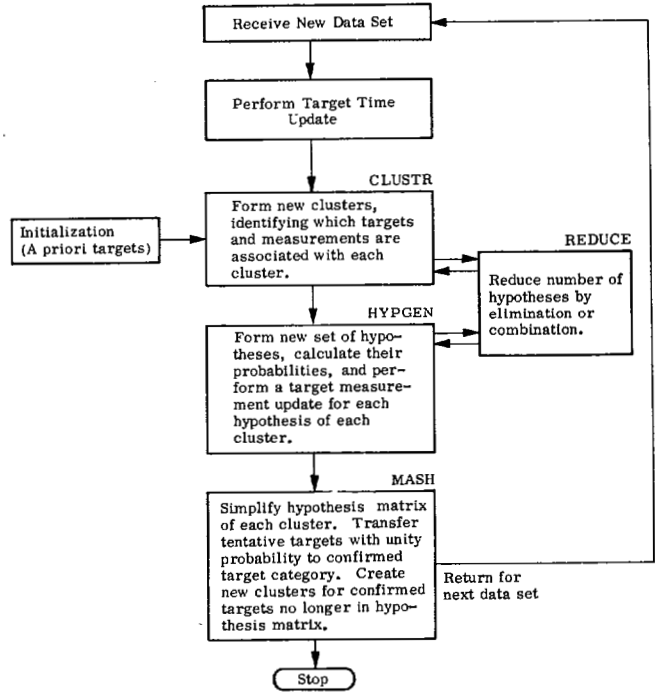

Fig. 1. Flow diagram of multiple-target tracking algorithm.

III. HYPOTHESIS GENERATION TECHNIQUE

The basic approach used in this paper is to generate a set of data-association hypotheses to account for all possible origins of every measurement. The filter in this paper generates *measurement*-oriented hypotheses, in contrast to *target*-oriented hypotheses developed in [2] and [3]. In the target-oriented approach, every possible measurement is listed for each target and vice versa for the measurement-oriented approach. Although both approaches are equivalent, a simpler representation is possible with the target-oriented approach *if* there is no requirement for track initiation. However, with a track initiation requirement, the measurement-oriented approach appears simpler.

Let $Z(k) \triangleq \{ Z_m(k), \ m = 1, 2, \cdots, M_k \}$ denote the set of measurements in data set k; $Z^k \triangleq \{ Z(1), Z(2), \cdots, Z(k) \}$ denote the cumulative set of measurements up through data set k; $\Omega^k \triangleq \{ \Omega_i^k, \ i = 1, 2, \cdots, I_k \}$ denote the set of all hypotheses at the time of data set k which associate the cumulative set of measurements Z^k with targets or clutter; and $\bar{\Omega}^m$ denote the set of hypotheses after the mth measurement of a data set has been processed. As a new set of measurements $Z(k+1)$ is received, a new set of hypotheses Ω^{k+1} is formed as follows: $\bar{\Omega}^0$ is initialized by setting $\bar{\Omega}^0 = \Omega^k$. A new set of hypotheses $\bar{\Omega}^m$ is repetitively formed for each prior hypothesis $\bar{\Omega}_i^{m-1}$ and each measurement $Z_m(k+1)$. Each hypothesis in this new set is the joint hypothesis that $\bar{\Omega}_i^{m-1}$ is true and that measurement $Z_m(k+1)$ came from target j. The values which j may assume are 0 (to denote that the measurement is a false alarm), the value of any prior target, or a value one greater than the current number of tentative targets (to denote that the measurement originated from a new

target). This technique is repeated for every measurement in the new data set until the set of hypotheses $\Omega^{k+1} = \overline{\Omega}^{M_{k+1}}$ is formed.

Before a new hypothesis is created, the candidate target must satisfy a set of conditions. First, if the target is a tentative target, its existence must be implied by the prior hypothesis from which it is branching. Second, a check is made to ensure that each target is not associated with more than one measurement in the current data set. Finally, a target is only associated with a measurement if the measurement lies within the gate or validation region of the target. If $\bar{x}$ and $\bar{P}$ are the mean and covariance of the target estimate for the prior hypothesis Ω_i^k, then the covariance of $\nu = Z_m - H\bar{x}$ is given by

$$B = H\bar{P}H^T + R \qquad (5)$$

and the measurement Z_m lies within an "η-sigma" validation region if

$$(Z_m - H\bar{x})^T B^{-1} (Z_m - H\bar{x}) \leq \eta^2. \qquad (6)$$

Note that the validation region also depends on the measurements since R is included in (5); however, for simplicity, it is assumed that all observations in the same data set have the same covariance R.

The representation of the hypotheses as a tree and as stored in the computer is shown in Fig. 2 for a representative cluster of two targets and three measurements. For the example, the prior targets are numbered 1 and 2, and the new tentative targets are numbered 3, 4, and 5. The three measurements in the data set are numbered 11, 12, and 13. Notice, for example, that if target 2 is already assigned to either measurement 11 or 12, a branch assigning it to measurement 13 will not be formed since it is assumed that one target cannot generate more than one measurement in a data set. The set of hypotheses is represented in a computer by a two-dimensional array, the "hypothesis matrix" which has a row for each hypothesis, and a column for each measurement. The entry in the array is the hypothesized origin of the measurement for that particular hypothesis. In programming the automatic hypothesis generation routine, a simplification of the hypothesis matrix occurs if the "prior hypothesis loop" is placed inside the "measurement loop." In this case, the hypothesis matrix at one stage is just a subset of the hypothesis matrix for the next stage as shown in the figure. This follows the numbering scheme for hypotheses used in [1].

Although there may be many hypotheses in a cluster, as far as each target in the cluster is concerned there are relatively few hypotheses. As an example, the cluster shown in Fig. 2 has 28 hypotheses; however, as far as target 1 is concerned, it only has two hypotheses: either it is associated with measurement 11 or it is not associated with any measurement. Similarly, targets 2, 3, 4, and 5 have 4, 2, 2, and 2 target hypotheses, respectively. A
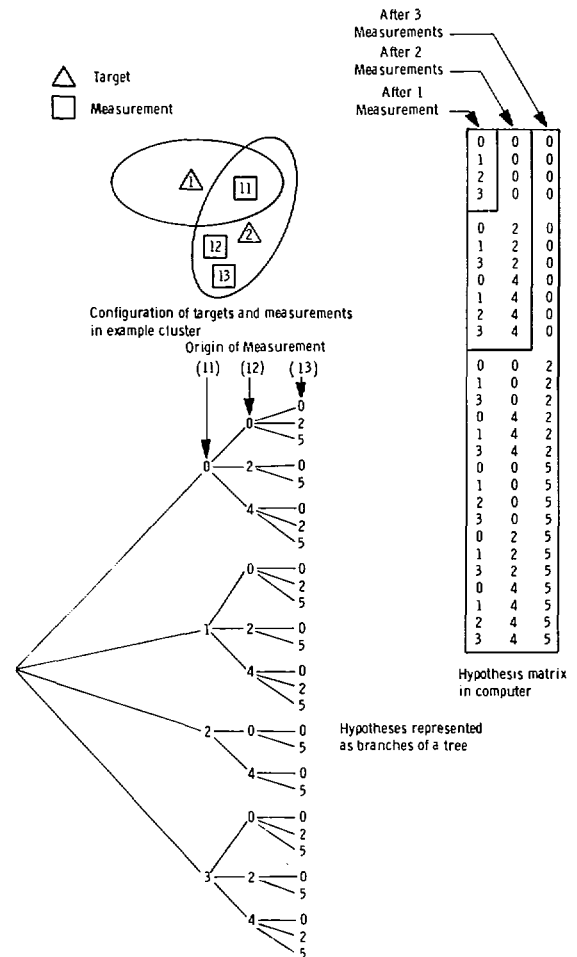

Fig. 2. Representation of hypothesis matrix.

"hypothesis relationship matrix" is created for each target, listing those cluster hypotheses which correspond to each target hypothesis. Alternative target states are then estimated for each target hypothesis and not each cluster hypothesis. The target estimates for each hypothesis are calculated by using a Kalman filter. The conditional probability distribution for the target states is then the sum of the individual estimates for each hypothesis, weighted by the probability of each hypothesis.

IV. PROBABILITY OF EACH HYPOTHESIS

The derivation for determining the probability of each hypothesis depends on whether the measurements are from a type 1 sensor or type 2 sensor. A type 1 sensor is one that includes numbers-of-targets type information as well as information on the location of each target. All the measurements in such a data set are considered together. In addition, an estimate of the new target density must be maintained to process measurements from this type of sensor. A type 2 sensor sends only positive reports. One measurement at a time is processed for this type of sensor and the new target density is not changed after each report.

A. Type 1 Sensor

Let P_i^k denote the probability of hypothesis Ω_i^k, given measurements up through time k, i.e.,

$$P_i^k \triangleq P(\Omega_i^k | Z^k). \tag{7}$$

We may view Ω_i^k as the joint hypothesis formed from the prior hypothesis Ω_g^{k-1} and the association hypothesis for the current data set ψ_h. The hypothesis ψ_h involves the hypothetical assignment of every measurement in the data set $Z(k)$ with a target. We may write a recursive relationship for P_i^k by use of Bayes' equation

$$P(\Omega_g^{k-1}, \psi_h | Z(k)) = \frac{1}{c} P(Z(k) | \Omega_g^{k-1}, \psi_h)$$
$$\cdot P(\psi_h | \Omega_g^{k-1}) P(\Omega_g^{k-1}) \tag{8}$$

where for brevity we have dropped the conditioning on past data through data set $k-1$. The factor c is a normalizing factor found by summing the numerator over the values of g and h. The first two terms on the right-hand side (RHS) of the above equation will now be evaluated. The first term is the likelihood of the measurements $Z(k)$, given the association hypothesis, and is given by

$$P(Z(k) | \Omega_g^{k-1}, \psi_h) = \prod_{m=1}^{M_K} f(m) \tag{9}$$

where

$f(m) = 1/V$ if the mth measurement is from clutter or a new target

$\quad = N(Z_m - H\bar{x}, B)$ if the measurement is from a confirmed target or a tentative target whose existence is implied by the prior hypothesis Ω_g^{k-1}.

V is the volume (or area) of the region covered by the sensor and $N(x, P)$ denotes the normal distribution $\exp[-\frac{1}{2}x^T P^{-1}x]/\sqrt{(2\pi)^n |P|}$. The values of x and B [through (5)] are those appropriate for the prior hypothesis Ω_g^{k-1}.

The second term on the RHS of (8) is the probability of a current data-association hypothesis given the prior hypothesis Ω_g^{k-1}. Each current data-association hypothesis ψ_h associates each measurement in the data set with a specific source; as such, it includes the following information.

Number: The number of measurements associated with the prior targets $N_{DT}(h)$, the number of measurements associated with false targets $N_{FT}(h)$, and the number of measurements associated with new targets $N_{NT}(h)$.

Configuration: Those measurements which are from previously known targets, those measurements which are from false targets, and those measurements which are from new targets.

Assignment: The specific source of each measurement which has been assigned to be from some previously known target.

Also, it is worth noting that the prior hypothesis Ω_g^{k-1} includes information as to the number of previously known targets $N_{TGT}(g)$ within the area of coverage of the sensor. This number includes any tentative targets whose existence is implied by that prior hypothesis, as well as the confirmed targets for that cluster. However, according to the current data-association hypothesis, only N_{DT} of these targets are detected by the sensor.

It is assumed that the number of previously known targets that are detected is given by a binomial distribution, the number of false targets follows a Poisson distribution and the number of new targets also follows a Poisson distribution. With these assumptions, the probability of the numbers N_{DT}, N_{FT}, and N_{NT} given Ω_g^{k-1} is

$$P(N_{DT}, N_{FT}, N_{NT} | \Omega_g^{k-1})$$
$$= \binom{N_{TGT}}{N_{DT}} P_D^{N_{DT}} (1 - P_D)^{(N_{TGT} - N_{DT})} \tag{10}$$
$$\times F_{N_{FT}}(\beta_{FT}V) F_{N_{NT}}(\beta_{NT}V)$$

where

P_D = probability of detection

β_{FT} = density of false targets

β_{NT} = density of previously unknown targets that have been detected (i.e., the P_D term has already been included in it)

$F_n(\lambda)$ = the Poisson probability distribution for n events when the average rate of events is λ.

The total number of measurements is given by

$$M_K = N_{DT} + N_{FT} + N_{NT}. \tag{11}$$

Of the M_K measurements, there are many configurations or ways in which we may assign N_{DT} of them to prior targets, N_{FT} of them to false targets, and N_{NT} of them to new targets. The number of configurations is given by

$$\binom{M_K}{N_{DT}} \binom{M_K - N_{DT}}{N_{FT}} \binom{M_K - N_{DT} - N_{FT}}{N_{NT}}.$$

The probability of a specific configuration, given N_{DT}, N_{FT}, and N_{NT}, is then

$$P(\text{Configuration} | N_{DT}, N_{FT}, N_{NT})$$
$$= \frac{1}{\binom{M_K}{N_{DT}} \binom{M_K - N_{DT}}{N_{FT}}}. \tag{12}$$

For a given configuration, there are many ways to assign the N_{DT} designated measurements to the N_{TGT}

targets. The number of possible assignments is given by

$$\frac{N_{TGT}!}{(N_{TGT}-N_{DT})!}.$$

The probability of an assignment for a given configuration is therefore

$$P(\text{Assignment}|\text{Configuration})=\frac{(N_{TGT}-N_{DT})!}{N_{TGT}!}. \quad (13)$$

Combining these last three equations and simplifying, we find that the probability of

$$
\begin{aligned}
P\big(\psi_h|\Omega_g^{k-1}\big)=&\frac{N_{FT}!N_{NT}!}{M_K!}\\
&\times P_D^{N_{DT}}(1-P_D)^{(N_{TGT}-N_{DT})}\\
&\times F_{N_{FT}}(\beta_{FT}V)F_{N_{NT}}(\beta_{NT}V). \quad (14)
\end{aligned}
$$

Substituting this and (9) into (8), we find that

$$
\begin{aligned}
P_i^k=&\frac{1}{c}\frac{N_{FT}!N_{NT}!}{M_K!}P_D^{N_{DT}}(1-P_D)^{(N_{TGT}-N_{DT})}\\
&\times F_{N_{FT}}(\beta_{FT}V)F_{N_{NT}}(\beta_{NT}V)\\
&\times\left[\prod_{m=1}^{N_{DT}}N(Z_m-H\bar{x},B)\right]\frac{1}{V^{N_{FT}+N_{NT}}}P_g^{k-1} \quad (15)
\end{aligned}
$$

where for ease of notation the measurements have been reordered so that the first N_{DT} measurements correspond to measurements from prior targets. Substituting for the Poisson processes, the dependence on V is eliminated! Simplifying and combining constants into c, we finally have

$$
\begin{aligned}
P_i^k=&\frac{1}{c}P_D^{N_{DT}}(1-P_D)^{(N_{TGT}-N_{DT})}\beta_{FT}^{N_{FT}}\beta_{NT}^{N_{NT}}\\
&\times\left[\prod_{m=1}^{N_{DT}}N(Z_m-H\bar{x},B)\right]P_g^{k-1}. \quad (16)
\end{aligned}
$$

This equation is the key development presented in this paper. It is similar to (12) in the paper by Singer, Sea, and Housewright [1], except it has been extended to the multiple-target case. They have a slightly different approach in that they are only concerned with sensor returns *within a target validation region*. If this approach is extended to the multiple-target case (as suggested by Bar-Shalom in [3]), considerable difficulty ensues in the derivation. Also, by considering area outside validation regions, we now have a track initiation capability.

This equation is used iteratively within the hypothesis generation routine to calculate the probability of each data-association hypothesis. Although it appears complicated, it is relatively easy to implement. If all the prior hypotheses are first multiplied by $(1-P_D)^{N_{TGT}}$, then as a branch is created for each measurement and its hypothesized origin, the likelihood of the branch is found by either multiplying the prior probability by either β_{FT}, β_{NT}, or $P_D N(Z_m-H\bar{x},B)/(1-P_D)$ as appropriate. After all such branches are generated the likelihoods are then normalized.

Concurrently with the above calculations, a calculation of β_{NT}, the density of new (i.e., unknown) targets, is performed whenever a data set from a type 1 sensor is received. The density of new targets β_{NT} depends upon the number of times the area has been observed by a type 1 sensor and the possible flux of undetected targets into and out of the area.

The development of this paper has implicitly assumed that the probability distribution of the target state would be given by or approximated by a normal distribution after one measurement. If the measurement vector contains all the target state variables, then the initial state and covariance of a target are given by $x=Z_m$ and $P=R$. However, in general this will not be true but the normal distribution assumption might nevertheless be made. For example, if the target state is position and velocity and only position is given in the measurement, then the velocity might be assumed to be normally distributed with zero mean and a standard deviation equal to one-third the maximum velocity of the target.

If the measurements or other factors are such that the assumption of a normal distribution after one measurement is not a good assumption, then appropriate modifications would have to be made to the gate criterion, the hypothesis probability calculations and the Kalman filter equations. As an example, consider the case where N targets on a plane surface, generate two sets of line-of-bearing (LOB) reports each containing exactly N LOB's (i.e., $P_D=1$, $\beta_{FT}=0$). The LOB's intersect in N^2 points, corresponding to the N real targets and N^2-N "ghosts." Since all the statistical degrees-of-freedom in the measurements are necessary just to determine location, there are no additional degrees-of-freedom remaining for correlating one LOB with another. Therefore, in this case, there is no gate criterion for the second data set and each of the N^2 pairs are equally likely.

B. Type 2 Sensor

To calculate the probability that a single measurement from a type 2 sensor is from a false target, a previously known target, or a new target, let us assume that it is selected at random from a set of $N_{DT}+N_{FT}+N_{NT}$ possible measurements, where the probability of N_{DT}, N_{FT}, and N_{NT} is given by the RHS of (10). For a given N_{DT}, N_{FT}, and N_{NT}, the probability of the measurement being from clutter, a previous target, or a new target is given by the ratio of N_{FT}, N_{DT}, and N_{NT} to their sum. Given that a measurement is from some previous target, the probability it is from a particular target is $1/N_{TGT}$. Finally, the likelihood of the measurement, given the target which originated it, is $1/V$ if it is from a false or new target and $N(Z_m-H\bar{x},B_j)$ if from a previous target.

Combining these effects, the likelihood of the measure-

ment given N_{DT}, N_{FT}, and N_{NT} is given by

$$\mathcal{L}(M1 = j | N_{DT}, N_{FT}, N_{NT})$$

$$= \frac{N_{FT}}{V}, \qquad j = 0$$

$$= \frac{N_{DT}}{N_{TGT}} N(Z_m - Hx_j, B_j), \qquad 1 \leqslant j \leqslant N_{TGT}$$

$$= \frac{N_{NT}}{V}, \qquad j = N_{TGT} + 1. \tag{17}$$

The unconditional likelihood of the measurement is found by taking the expected value of (17), namely,

$$\mathcal{L}(M1 = j) = \frac{\overline{N}_{FT}}{V} = \beta_{FT}, \qquad j = 0$$

$$= \frac{\overline{N}_{DT}}{N_{TGT}} N(Z_m - Hx_j, B_j)$$

$$= P_D N(Z_m - Hx_j, B_j), \qquad 1 \leqslant j \leqslant N_{TGT}$$

$$= \frac{\overline{N}_{NT}}{V} = \beta_{NT}, \qquad j = N_{TGT} + 1. \tag{18}$$

If these likelihoods are normalized, one obtains the probability for each possible origin of the measurement. The implementation is the same as for (16) except that only one measurement at a time is processed for a type 2 sensor and there are no $(1 - P_D)$ terms.

V. HYPOTHESIS REDUCTION TECHNIQUES

The optimal filter developed in the previous section requires an ever-expanding memory as more data are processed. For this reason, techniques are needed to limit the number of hypotheses so that a practical version can be implemented. The goal is an algorithm which requires a minimum amount of computer memory and execution time while retaining nearly all the accuracy of the optimal filter. All the hypotheses may be considered as branches of a tree: the hypothesis reduction techniques may be viewed as methods of either pruning these branches or binding together branches.

A. Zero-Scan Algorithms

A zero-scan filter allows only one hypothesis to remain after processing each data set. The simplest method (and that probably most representative of current practice) is to choose the most likely data association hypothesis and use a standard Kalman filter to estimate target states. This is strictly a pruning operation. An improved variation of this is to still choose the maximum likelihood hypothesis but to increase the covariance in the Kalman filter to account for the possibility of miscorrelations. Another approach, developed in [2] and [3] and denoted the probabilistic data association (PDA) filter, is equivalent to combining all the hypotheses by making the target estimates depend on all the measurements.

B. Multiple-Scan Algorithms

In multiple-scan algorithms, several hypotheses remain after processing a data set. The advantage of this is that subsequent measurements are used to aid in the correlation of prior measurements. Hypotheses whose probabilities are increased correspond to the case in which subsequent measurements increase the likelihood of that data association. The simplest technique is again to prune all the unlikely hypotheses but keep all the hypotheses with a probability above a specified threshold. In [1], an N-scan filter for the single-target case was developed in which hypotheses that have the last N data scans in common were combined. A remarkable conclusion of their simulation was that with N only equal to one, near-optimal performance was achieved.

An alternative criterion for binding branches together (i.e., combining hypotheses) is to combine those hypotheses which have similar effects. Generally, this criterion would correspond to the N-scan criterion, but not always. If hypotheses with the last N data scans in common are combined, then hypotheses that differentiate between measurements in earlier scans are eliminated. Examples can be conceived [12] in which it is more important to preserve earlier rather than later hypotheses. For this reason, this paper uses the criterion of combining those hypotheses with similar effects concurrently with the criterion to eliminate hypotheses with a probability less than a specified amount α. For two hypotheses to be similar, they must have the same number of tentative targets and the estimates for all targets in each hypothesis must be similar, i.e., both the means and the variances of each estimate must be sufficiently close. The mean and covariance of the resulting estimate is a combination of the individual estimates.

C. Simplifying the Hypothesis Matrix and Initiating Confirmed Targets

By eliminating hypotheses, as in the previous section, the number of rows in the hypothesis matrix is reduced. This reduction may also allow us to reduce the number of columns in the hypothesis matrix. If all the entries in a column of the hypothesis matrix are the same, then that measurement has a unique origin and that column may be eliminated. This simple procedure is the only technique used to simplify the hypothesis matrix of each cluster. If the unique origin of the measurement is a tentative target, then that target is transferred to the confirmed target file. In other words, the criterion for initiating a new confirmed target is that a tentative target has a probability of existing equal to one (after negligible hypotheses have been dropped). Once the hypothesis matrix has been simplified as much as possible, many of the confirmed targets for that cluster may no longer be in the hypothesis matrix. These targets may then be removed from that cluster to form new clusters of their own. In this way, clusters are decomposed and prevented from becoming ever larger

and larger through collisions. The features in this paragraph have been incorporated into the MASH subroutine.

VI. Cluster Formation

If the entire set of targets and measurements can be divided into sets of independent clusters [3], then a great deal of simplification may result. Instead of one large tracking problem, a number of smaller tracking problems can be solved independently. Since the amount of computer storage and computation time grows exponentially with the number of targets, this can have an important effect in reducing computer requirements. If every target could be separated into its own individual cluster, these requirements would only grow linearly with the number of targets.

A cluster is completely defined by specifying the set of targets and measurements in the cluster, and the alternative data-association hypothesis (in the form of the hypothesis matrix) which relates the targets and measurements. Included in this description is the probability of each hypothesis and a target file for each hypothesis.

As part of the program initialization, one cluster is created for each confirmed target whose existence is known *a priori*. Each measurement of the data set is associated with a cluster if it falls within the validation region [(6)] of any target of that cluster for any prior data-association hypothesis of that cluster. A new cluster if formed for each measurement which cannot be associated with any prior cluster. If any measurement is associated with two or more clusters, then those clusters are combined into a "supercluster." The set of targets and measurements of the supercluster is the sum of those in the associated prior clusters. The number of data-association hypotheses of the supercluster is the product of the number of hypotheses for the associated prior clusters. The hypothesis matrix, probabilities of hypotheses, and target files must be created from those of their constituent prior clusters.

It can be verified that the probabilities of a set of joint hypotheses formed by combining two or more clusters is the same whether calculated by (16) for the combined clusters, or by multiplying the probabilities calculated by this equation for each separate cluster. This property, in fact, was one of the factors for choosing the Poisson and binomial distributions for describing the number of targets in (10).

VII. Example to Illustrate Filter Characteristics

A simple aircraft tracking problem from [1] was chosen for illustrating and evaluating the filter derived in the previous section. The state of the aircraft is its position and velocity in the X and Y directions. Measurements of position only are taken. Each dimension is assumed independent, with identical equations of motion, measurement errors, and process noise, i.e.,

$$\Phi = \begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \Gamma = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix},$$
$$H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},$$
$$Q = \begin{bmatrix} qT & 0 \\ 0 & qT \end{bmatrix}, \quad R = \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}. \tag{19}$$

A. Track Initiation

In the first example, a set of five measurements at five different times is used to illustrate track initiation. For this example, there are no initially known targets, the initial density of unknown targets is 0.5, the density of false reports is 0.1, the probability of detection is 0.9, and both the process and measurement noise have variances of 0.04. The five measurements are shown as triangles in Fig. 3. The most likely hypothesis after processing each measurement is that there is one target. The estimated position, velocity, and 1σ error circle of the target for that hypothesis is also shown in the figure. As expected, the estimated position at each time lies between the previously projected position and the measured position. After the first measurement is processed, there is a 5/6 probability the measurement came from a target and a 1/6 probability it came from a false report, since the relative densities are initially 5:1. The probability there is at least one target increases with every measurement to 99+ percent after five measurements, at which point a confirmed target is created. There is an interesting effect after four measurements are processed. The most likely hypothesis is that all four measurements came from the same target ($p=88$ percent); the second most likely hypothesis is that the first measurement came from a false report, and the remaining three measurements came from the same target ($p=4$ percent). Both of these hypotheses declare that there is one target, and since the estimated state of the target in both cases is nearly equal the two hypotheses are automatically combined.

B. Crossing Tracks

In the next example, we examine the capability of the filter to process measurements from two crossing targets. One target starts in the upper left corner and moves downward to the lower right while the other target starts in the lower left corner and moves upward to the upper right corner. The existence of just one of the targets is known *a priori*. The set of measurements and the target estimates corresponding to the most likely hypothesis are shown in Fig. 4. The first two measurements at the top of the figure are processed as in the track initiation example, and the first two measurements at the bottom are processed as a track maintenance problem. At $k=3$, however, the two clusters "collide" and a supercluster made of both clusters and both measurements is formed. This
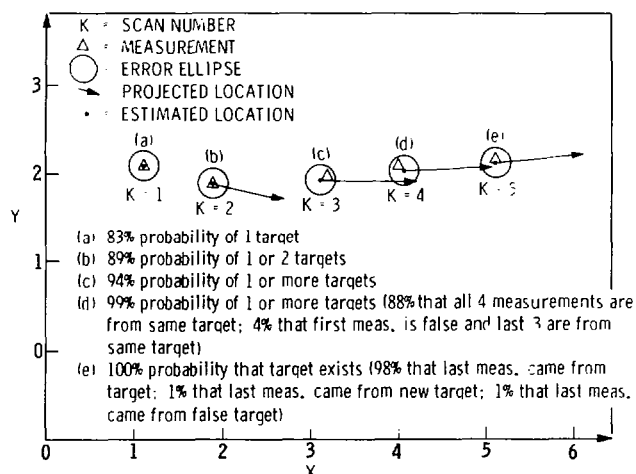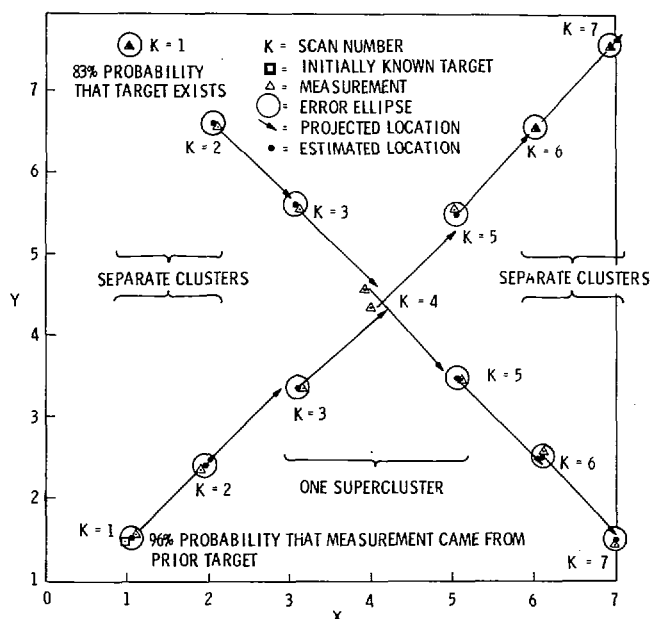
Fig. 3. Example of track initiation.

Fig. 4. Example of crossing tracks.

collision is due to the fact that one of the hypotheses in the top cluster is that the top measurement at $k = 1$ was from a target, but the next measurement was from clutter. Since we assumed an initial variance in velocity of 1.0, the above target could have originated the measurement at $(3.2, 2.9)$. After the measurements are processed, however, this possibility is so remote that it is eliminated. After eliminating all the other negligible hypotheses at $k = 3$, the supercluster is separated into two clusters, corresponding to the two targets. To process the two measurements at $k = 4$, the supercluster has to be formed again. At this time, the tentative target in the top of the figure becomes a confirmed target. Two hypotheses remain after processing the measurements at $k = 4$; that the lower target originated the lower measurement and the higher target originated the higher measurement ($p_1 = 60$ percent) or vice versa ($p_2 = 40$ percent). The measurements at $k = 5$ are such that they reduce the difference in probabilities of these two hypotheses (to $p_1 = 54$ percent, $p_2 = 46$ percent).

This is one case in which later data did not help resolve a prior data-association hypothesis; in fact, the ambiguity was increased. At $k = 5$, the data-association hypotheses at $k = 4$ are the most significant and are preserved. (If the $N = 1$ scan filter criterion was used, the hypotheses at $k = 4$ would have been eliminated.) By the time measurements at $k = 6$ are processed, the difference in the hypotheses at $k = 4$ is no longer important since the target estimates are now so similar. From then on, we have two separate track maintenance problems.

C. High Target Density

The last example illustrates the difficulty of associating measurements into tracks for a more complicated arrangement of measurements. This example is a single run from the Monte Carlo program described in the next section. In this example there are five real targets; the existence of four of them is initially known by the filter. The *a priori* location and velocity estimates of these four targets as well as measurements from the first six scans are shown in Fig. 5. Both the measurement noise and the process noise are relatively large ($q = r = 0.40$). The data points are shown grouped according to the maximum likelihood data association hypotheses (except as noted below). In addition, there are approximately 15 other feasible groupings of targets that are also possible arrangements. As measurements are processed, the probabilities of these different groupings change. For example, at scan 4 the most likely hypothesis is that measurement 19 is associated with target 1 and measurement 18 is associated with target 2; however, on scan 5 and subsequent scans, another hypothesis becomes the most likely and reverses this assignment. The one target unknown by the filter is being formed by measurements 2, 8, and 14. Even at scan 4 when there is only one measurement for either target 3 or the new target, the most likely hypothesis is that measurements 2, 8, and 14 are a new target and measurements 5, 10, 16, and 20 are from target 3. At scan 5, however, the most likely hypothesis is that measurements 2, 10, and 14 (as well as 23) are false, and that measurements 5, 8, 16, and 20 are associated with target 3. At scan 6 the likelihood that measurements 2, 8, 14, 23, and 30 form a new target is increased and by scan 7 it is part of the most likely hypothesis. This target does not become a confirmed target until scan 9. In each case, the general grouping of measurements corresponds to one of the actual targets.

VIII. MONTE CARLO SIMULATION

The independent factors affecting filter performance include both filter characteristics (e.g., the filter criteria for eliminating or combining hypotheses) and environmental variables, such as target density β_{TT}, measurement accuracy R, target perturbances Q, false report densities β_{FT}, and data rate T, P_D.
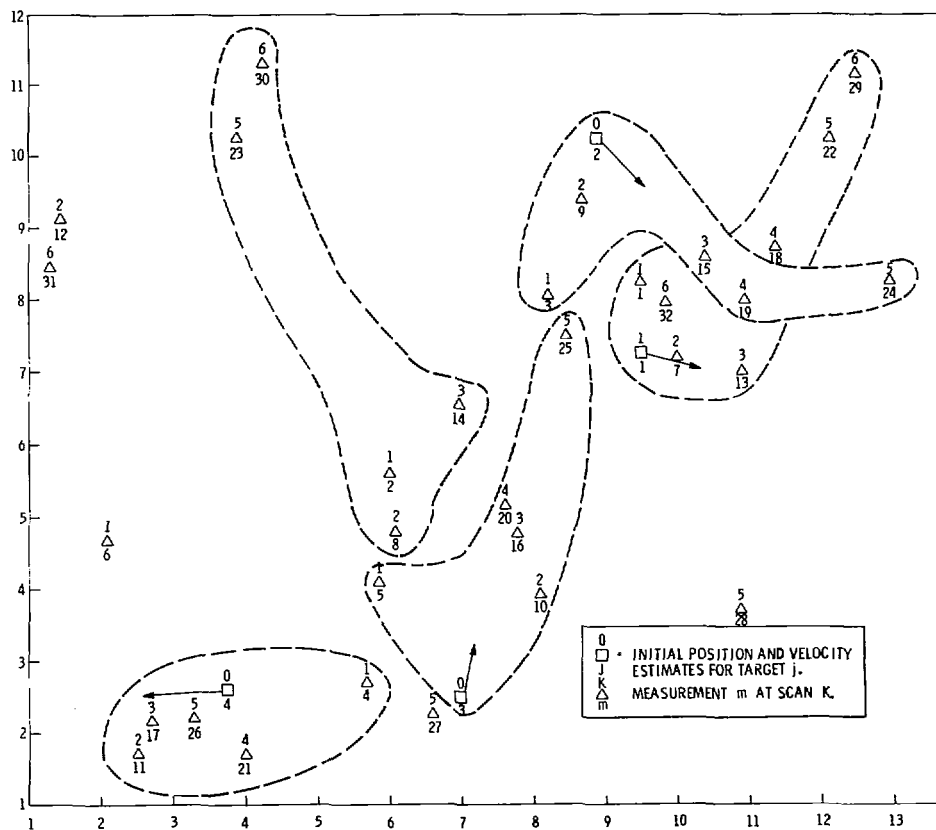
Fig. 5. Tracking in a dense target environment.

The probability or ease of making correct associations between targets and measurements appears to be a key factor in determining filter performance. As such, the primary effect of many of the environmental variables mentioned above is in affecting this factor. The more dense the targets and measurements or the larger the uncertainty in their locations, the more difficult it is to make the correct associations. The variables β_{TT}, β_{FT}, and P_D determine the density of measurements according to

$$\beta_M = \beta_{FT} + P_D \beta_{TT} \qquad (20)$$

and the variance between measurements and eliminated target locations (just before an association is to be made) is given by

$$\sigma^2 = \bar{P}_{11} + r. \qquad (21)$$

If we make the approximation that $\bar{P}_{11}$ reaches its steady-state value as given by solving the Kalman filter equations, then $\bar{P}_{11}$ can be related to q, r, and T. P_D also affects P_{11} by affecting the average time between measurements. In addition, the actual value of $\bar{P}_{11}$ would be affected by combining hypotheses so that the relationship between $\bar{P}_{11}$ and q, r, and T must be considered an approximation.

An indication of the increasing difficulty of association with increasing target and measurement densities and uncertainties is given by the association probabilities shown in Fig. 6. The figure shows the probability of not associating a target with a measurement, and the probabilities of associating a target with a measurement correctly and incorrectly, as the density-variance product increases. These curves may be viewed as the association probabilities of an $N=0$ scan filter in processing one scan of target information when the filter knows the number of true targets and their prior locations from the previous scan, and uses the following association rules: a target is associated with the closest measurement within its validation region; if no measurements are within validation region (or such a measurement is assigned to another target), then no measurement is associated with the target. The curves do not reflect the performance of any filter over more than one scan.

The curves in Fig. 6 were generated for the case where a proportion $P_D = 0.9$ of the true targets generated a measurement, the relative density of true targets to false measurements was $5/1$, and the target gate size was $\eta = 3$. Since only 90 percent of targets generated measurements, the probability of correctly associating a target with a measurement is no greater than 90 percent. As the density-variance product increases, this probability decreases. An even worse condition is the related increase in the probability of incorrectly associating a target and a measurement. Incorrect association can have a cascading effect on tracking algorithms which do not account for it. The probability a target is not associated with any measurement is initially 10 percent for this set of conditions and also decreases with increasing density-variance.
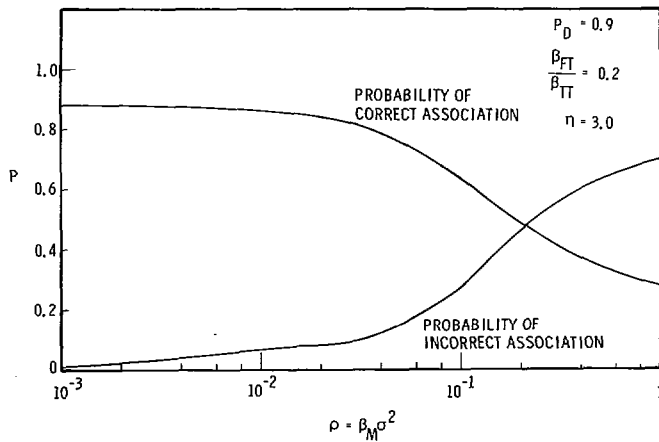
Fig. 6. The probability of correct and incorrect data association versus measurement density × variance.

A number of Monte Carlo runs were made to validate the filter and examine its performance under various conditions. The area under surveillance was 10 units by 10 units. Unless otherwise noted, values of $\beta_{TT} = 0.05$, $\beta_{FT} = 0.01$, $P_D = 0.9$, $q = 0.04$, $r = 0.04$, and $T = 1.0$ were used in generating the number of targets and the measurements. The fraction of true targets initially known by the filter was $f_{NT} = 0.80$. The initial value of the density of unknown targets in the filter was given by

$$\beta_{NT} = \beta_{TT}(1 - f_{NT})P_D \qquad (22)$$

with a lower limit of $\beta_{FT}/4$.

In the first set of simulations, three versions of the filter, corresponding to values of the hypothesis elimination criterion $\alpha = 0.01$, $\alpha = 0.1$, and $\alpha = 0.5$, were run with the density-variance product ranging from $\rho = 0.01$ to $\rho = 0.1$. The filter is mechanized so that at least one data-association hypothesis is retained for each cluster so that $\alpha = 0.5$ corresponds to the maximum likelihood zero-scan filter. The density-variance product was varied by increasing both q and r the same amount. The results of this simulation are summarized in Table II. Each data point summarizes results over 10 Monte Carlo runs at 10 time intervals, or a total of 100 comparisons between the true target locations and the estimated target locations. There did not appear to be any error trends with time for these runs for the above values of P_D, f_{NT}, etc. except for an increase in false targets with time.

As the density-variance product increases, the difficulty of making data-to-target association increases. This difficulty causes an increase in the number of hypotheses for the $\alpha = 0.01$ filter. The effect of more drastic pruning ($\alpha = 0.1$) reduces the number of hypotheses. Increases in the density-variance product do not appear to increase the number of hypotheses for the heavily pruned cases ($\alpha = 0.1$ and $\alpha = 0.5$). The percent of targets tracked (ratio of correct tracks to true number of targets) appeared quite good for the first case; however, as the density-variance product increased or as more hypotheses were pruned, this percentage dropped. The percent of false targets (ratio of false tracks to total number of tracks) remained mod-

TABLE II
RESULTS OF MONTE CARLO SIMULATION

Case	Number of Hypotheses	Percent Targets Tracked	Percent False Targets	Normalized Error
$q = r = 0.04$, $\alpha = 0.01$ ($\rho = 0.01$)	10.52	97.3	2.9	0.920
$q = r = 0.04$, $\alpha = 0.1$ ($\rho = 0.01$)	7.67	97.1	3.8	0.922
$q = r = 0.04$, $\alpha = 0.5$ ($\rho = 0.01$)	4.40	81.0	0.6	0.870
$q = r = 0.12$, $\alpha = 0.01$ ($\rho = 0.03$)	13.65	93.6	6.7	1.053
$q = r = 0.12$, $\alpha = 0.1$ ($\rho = 0.03$)	8.04	92.6	6.0	1.052
$q = r = 0.12$, $\alpha = 0.5$ ($\rho = 0.03$)	4.40	75.4	7.0	1.062
$q = r = 0.12$, $\alpha = 0.01$ ($\rho = 0.1$)	16.41	89.1	7.7	1.055
$q = r = 0.4$, $\alpha = 0.1$ ($\rho = 0.1$)	7.90	85.2	11.8	1.071
$q = r = 0.4$, $\alpha = 0.5$ ($\rho = 0.1$)	4.40	75.0	6.5	1.043
$q = r = 0.04$, $\alpha = 0.1$ ($\rho = 0.01$) (High β_{NT})	7.76	99.3	5.8	0.937
$q = r = 0.12$, $\alpha = 0.1$ ($\rho = 0.03$) (High β_{NT})	8.05	94.3	6.2	1.052
$q = r = 0.4$, $\alpha = 0.1$ ($\rho = 0.1$) (High β_{NT})	7.81	90.8	12.1	1.098

erate for all conditions. Under all conditions, the normalized error (half the average squared position error inversely weighted by the filter's estimated variance in position) was approximately equal to one, indicating that the actual accuracy of the filter agreed with the accuracy predicted by the filter. The filter was neither overconfident nor underconfident.

All of the filter parameters are well-defined and measurable quantities except the new target density β_{NT}, which is scenario-dependent. A value of β_{NT} should be chosen based upon the expected range of target densities and the relative importance the user assigns to missing real tracks versus creating false tracks. The higher β_{NT}, the less likely the algorithm will miss a real target but the more likely it will create a false track. The last three entries in Table II are for the case where the value of β_{NT} used by the filter is twice what it should have been according to (22).

To present results of other factors which affect performance (e.g., P_D, f_{NT}, and β_{FT}/β_{TT}) would unduly increase the size of this paper and will not be done except to say that decreases in P_D and f_{NT} and increases in β_{FT}/β_{TT} reduce performance. Also, the percentage of targets tracked and the number of false tracks are functions of the scan number for values of P_D, f_{NT}, and β_{FT}/β_{TT} not used in this simulation.

The algorithm was coded in 1 500 lines of Fortran and executed on a UNIVAC 1100 computer. Each of the subroutines shown in Fig. 1 took approximately 300 lines of Fortran, with the main program and other subroutines taking another 300 lines. The core memory requirements for an algorithm capable of handling 10 clusters and 30 targets (and including the Monte Carlo program for gen-

erating measurements and evaluating the algorithm) was approximately 64K words. Ten Monte Carlo runs of 10 scans each were executed in 25–45 s depending upon the particular case. To handle more clusters and targets, or to reduce memory requirements, the cluster and target information could be put on a disk file. However, disk access time would then cause a large increase in the overall execution time of the program.

IX. CONCLUSIONS

This paper has developed a multiple-target tracking filter incorporating a wide range of capabilities not previously synthesized, including track initiation, multiple-scan correlation, and the ability to process data sets with false or missing measurements. The primary contribution is a Bayesian formulation for determining the probabilities of alternative data-to-target association hypotheses, which permitted this synthesis. In simulations of a simple aircraft tracking problem, the filter demonstrated its capabilities over a wide range of target densities and measurement uncertainties. The filter proved to be robust to errors in the given filter parameters (e.g., unknown target density).

ACKNOWLEDGMENT

The author would like to thank Dr. H. R. Rauch, Mr. R. G. Bryson, and the reviewers for their helpful suggestions.

REFERENCES

[1] R. A. Singer, R. G. Sea, and K. B. Housewright, "Derivation and evaluation of improved tracking filters for use in dense multi-target environments," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 423–432, July 1974.

[2] Y. Bar-Shalom and E. Tse, "Tracking in a cluttered environment with probabilistic data association," *Automatica*, vol. 11, pp. 451–460, 1975.

[3] Y. Bar-Shalom, "Extension of the probabilisitc data association filter in multi-target tracking," in *Proc. 5th Symp. on Nonlinear Estimation*, Sept. 1974, pp. 16–21.

[4] D. L. Alspach, "A Gaussian sum approach to the multi-target identification-tracking problem," *Automatica*, vol. 11, pp. 285–296, 1975.

[5] R. W. Sittler, "An optimal data association problem in surveillance theory," *IEEE Trans. Mil. Electron.*, vol. MIL-8, pp. 125–139, Apr. 1964.

[6] J. J. Stein and S. S. Blackman, "Generalized correlation of multi-target track data," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-11, pp. 1207–1217, Nov. 1975.

[7] P. Smith and G. Buechler, "A branching algorithm for discriminating and tracking multiple objects," *IEEE Trans. Automat. Contr.*, vol. AC-20. pp. 101–104, Feb. 1975.

[8] C. L. Morefield, "Application of 0–1 integer programming to multitarget tracking problems," *IEEE Trans. Automat. Contr.*, vol. AC-22, pp. 302–311, June 1977.

[9] Y. Bar-Shalom, "Tracking methods in a multitarget environment," *IEEE Trans. Automat. Contr.*, vol. AC-23, pp. 618–626, Aug. 1978.

[10] D. L. Alspach and H. W. Sorenson, "Recursive Bayesian estimation using Gaussian sums," *Automatica*, vol. 7, 1971.

[11] R. E. Kalman, "A new approach to linear filtering and prediction problems," *J. Basic Eng.*, vol. 82-D, pp. 35–45, 1960.

[12] D. B. Reid, "A multiple hypothesis filter for tracking multiple targets in a cluttered environment," LMSC Rep. D-560254, Sept. 1977.

Donald B. Reid (S'69–M'72) was born on March 29, 1941, in Washington, DC. He received the B.S. degree from the U.S. Military Academy, West Point, NY, in 1963 and the M.S. and Ph.D. degrees in aero & astronautical engineering from Stanford University, Palo Alto, CA in 1965 and 1972, respectively.

From 1965 to 1969 he served in the U.S. Air Force as an Astronautical Engineer in the 6595th Aerospace Test Wing at Vandenberg AFB and participated in the testing and launch operations of military space programs. He was a member of the technical staff at the Institute for Defense Analyses in Arlington, VA, from 1972 to 1976. Since July 1976 he has been a scientist with the Palo Alto Research Laboratory of Lockheed Missiles & Space Company. His current interests include multiple target tracking, orbital rendezvous and station keeping, system identification, and military command and control systems.