Abstract – This paper introduces cardinality tracking, a special case of the more general multi-target tracking problem for which measurements do not provide any target state information. That is, each scan only provides information as to how many targets are present. We address the problem with a modified form of the multiple-hypothesis tracking formalism using equivalence classes. Structural results exist which enable optimal track extraction to be achieved. We introduce as well some variations, approximate approaches that introduce further hypothesis aggregation. We show that we are able to improve significantly over a straightforward MHT approach to the problem. Similar results can be obtained by considering the problem as one of Kalman filtering over the aggregation of targets.

Keywords: multiple-hypothesis tracking, cardinality tracking, equivalence classes, greedy target problem, group tracking.

1 Introduction

In the limit of large measurement errors, we are faced with a cardinality tracking problem in which we seek only to identify how many targets are present. We study this problem starting from a track-oriented MHT formalism [1]. The relevant tracking metric is track cardinality error, which measures how well on average we can identify the number of targets present. Our initial study on cardinality tracking is documented in [2], and was motivated by the need to introduce preference for confirmed tracks in the MHT assignment process; in the large-measurement-error limit, this preference leads to optimal tracking.

As shown in Section 2, given the interchangeability of sensor measurements and the lack of gating conditions, there are many identical solutions to the MHT optimization task. This can be exploited to result in significant computational savings.

Unfortunately, cardinality tracking highlights a weakness in the MHT formalism, in particular that fact that the approach return a single global hypothesis that satisfies a maximum a posteriori (MAP) or a maximum likelihood condition.

The limitation in returning the likeliest solution is that it may not capture statistics of interest. There is an interesting analogy to a simple coin-tossing experiment. Consider a biased coin with p(heads)=0.6. Toss the coin 100 times. What is the a priori likeliest outcome? The answer, 100 heads, provides a very poor estimate of the expected number of heads (60).

The fact is, in cardinality tracking, the null solution (no tracks) will often be identified as the MAP solution, particularly for non-trivial false alarm rates. This follows from the lack of filter innovation terms, consistent with the lack of target state information in cardinality measurements.

To circumvent this difficulty, in Sections 3-4 we introduce an equivalence class approach whereby all indistinguishable global hypotheses are treated together. As such, the MHT formalism identifies the likeliest class of hypotheses. We call this the cardinality tracker.

Some modifications to the cardinality tracker, called the psi tracker and aggregate psi tracker (for reasons that will be clear in the sequel) are discussed in Section 5. Finally, if we relax the requirement of integer estimates for the number of tracks, we can formulate the problem via as one of Kalman filtering, as discussed in Section 6. (Interestingly, in the single-scan case, it is straightforward to consider the exact form of the posterior distribution [3].) Section 7 provides numerical results, and concluding remarks are in section 8.

We believe our approach is relevant to large-scale aggregate surveillance tasks. The approach could be fruitfully applied to the so-called cardinality estimation problem, which is studied in the anonymous RFID community (see [4] and related references) and in other large-scale counting problems [5]. Note that approximate counting is done when privacy or speed (or both) are of concern.

2 Cardinality tracking: standard and modified MHT formulations

In the case of cardinality-only information, the key recursive track-oriented MHT equation is the following:
\[
p(q^k | Z^k) = p_Z^X \left[ \left( 1 - p_Z^X \right)^{X_d - d} \left( 1 - p_d \right)^{p_d^{\beta \theta} \lambda^k_{j0}} \right] \left( \frac{p(q^k | Z^{k-1})}{\bar{c}_k} \right),
\]

where for all scans \( \lambda_{jh} \) is the mean for Poisson-distributed target births, \( \lambda_{ja} \) is the mean for Poisson-distributed false returns, \( p_d \) is the object detection probability, \( p_Z \) is the object termination probability, \( \tau \) is the number of active tracks, \( q^k = (q_1^k, \ldots, q_k^k) \) is the discrete state that accounts for target births and terminations as well as all data associations, \( \bar{c}_k \) is a normalizing factor that is independent of \( q^k \), and \( Z^k = (Z_1^k, \ldots, Z_k^k) \) is the observation sequence with \( Z_i \) denoting the measurement set at scan \( i \).

In the standard track-oriented MHT formulation, we seek the maximum a posteriori (MAP) \( q^k \) given measurement data (which we denote by \( q^k \)), and then condition on \( q^k \) to estimate a set of target trajectories, \( X^k = (X_1^k, \ldots, X_k^k) \), that constitutes the solution to the tracking problem. We denote \( \hat{X}^k \) as the MMSE solution conditioned on \( q^k \).

\[
q^k = q_{\text{MAP}}(Z^k) = \arg \max q^k | Z^k, \quad X^k = X_{\text{MMSE}}(Z^k, q^k).
\]

Note that, while (1) expresses the global hypothesis score, this expression factors directly into local track hypothesis scores. Indeed, track-oriented MHT avoids the explicit enumeration of global hypotheses.

In cardinality tracking, each target trajectory is solely characterized by a track initiation time and a track termination time.

The use of (1) for cardinality tracking is problematic. A first difficulty is a computational one: since all measurements gate with all tracks, there is a large number of track hypotheses.

A partial reduction in the complexity of the optimization problem (1) can be achieved by recognizing that measurements are interchangeable. Assume we have \( N \) resolved tracks, \( n \)-scan delayed hypothesis resolution, and a sequence of sets of measurements of cardinality \( \{Z_1, \ldots, Z_{n-\text{scan}+1}\} \). We recast the hypothesis generation logic such that a single measurement exists in each scan, but it must be used multiple times. Accordingly, hypothesis branching is more limited than in standard MHT, though no measurement gating is applied. It is easy to see that the number of track hypotheses \( n \) at level \( i \) is bounded by the following, with the inequality due to track termination of some branches (inability to reach track confirmation or repeated missed detections). Note that \( n_0 = N \).

\[
n_i \leq 2^{n_i} + |Z|, i = 1, \ldots, n - \text{scan} + 1.
\]

The solution to eqn. (4) is given by:

\[
n_i \leq 2^n N + \sum_{j=1}^{i} 2^{j-i} |Z|, i = 1, n - \text{scan} + 1.
\]

Having generated \( n_{n-\text{scan}+1} \) track hypotheses characterized by a vector of track scores \( c \), we have the following optimization problem with \( \xi_{n-\text{scan}+1} = \sum_{i=0}^{n-\text{scan}+1} \).

\[
J = c^T x, \quad Ax = b,
\]

where the matrix \( A \) is a \( \xi_{n-\text{scan}+1} \)-by- \( n_{n-\text{scan}+1} \) matrix with structure defined by the form of the set of track hypothesis trees, \( x \) is a vector of ones and zeros, and the vector \( b \) is given by \( N \) ones (resolved track constraints) and \( n-\text{scan}+1 \) entries given by \( \{Z_1, \ldots, Z_{n-\text{scan}+1}\} \) (measurement constraints). We seek to minimize (6) while satisfying (7).

The reduced-complexity optimization problem formulation does not address a more serious concern with cardinality tracking: a large number of comparably-scoring global hypotheses in large-scale surveillance applications, leading to the null solution as the optimal choice under (2). Indeed, for large surveillance problems, the posterior probability \( p(q^k | Z^k) \) will be very small for all non-trivial choice of \( q^k \) (i.e. for all but the null solution).

Further, unlike conventional tracking, measurements are much less informative and we lack kinematic filter residuals that lead to relatively large hypothesis scores for some association decisions and relatively small scores for others. This leads to a need for a small sensor revisit time, further exacerbating these difficulties.

The difficulties associated with identifying a single global hypothesis point to the fundamental limitation in tracking solutions (like MHT) that do not provide the rich solution that optimal Bayesian tracking would provide: the full posterior probability distribution \( p(q^k | Z^k) \). Nonetheless, it would be beneficial to identify a set of global hypotheses that are indistinguishable (due to measurement equivalence) and that provide significant probability mass. This motivates the development in the next section.
3 Cardinality tracking: The equivalence-classes approach

We wish to identify a most probably set of discrete states \( \{y^k\} \), where all member of the set are equivalent under measurement re-ordering. To do so, we must revisit the derivation of track-oriented MHT equations and introduce suitable modifications.

We are interested in a recursive and computationally efficient expression for \( p[y^k] | Z^k \) that lends itself to functional optimization without the need for explicit enumeration of global hypotheses. We do so through repeated use of Bayes’ rule. Note that, for notational simplicity, we use \( p(\cdot) \) for both probability density and probability mass functions.

\[
p[y^k] | Z^k = p[Z_k | Z^{k-1}, \{y^k\}] p[y^k] | Z^{k-1} = \sum_{\{y^k\}} p[Z_k | Z^{k-1}, \{y^k\}] p[y^k] | Z^{k-1}.
\]

(8)

The recursive expression (8) involves two factors that we consider in turn, with the discrete state probability one first. It will be useful to introduce the aggregate variable \( \psi_k \) that accounts for the number of detections \( d \) for the \( \tau \) existing tracks, the number of track deaths \( \chi \), the number of new tracks \( b \), and the number of false returns \( r - d - b \), where \( r \) is the number of contacts in the current scan.

\[
p[y^k] | Z^k = \sum_{\{y^k\}} p[Z_k | Z^{k-1}, \{y^k\}] p[y^k] | Z^{k-1} \psi_k
\]

(10)

\[
p[y^k] | Z^k = \sum_{\{y^k\}} p[Z_k | Z^{k-1}, \{y^k\}] p[y^k] | Z^{k-1} \psi_k
\]

(11)

\[
p[y^k] | Z^k = \sum_{\{y^k\}} p[Z_k | Z^{k-1}, \{y^k\}] p[y^k] | Z^{k-1} \psi_k
\]

(12)

The key difference with respect to standard track-oriented MHT is in (12), since we must not account for differences in which measurements are taken to be track updates, which are taken to be track births, and how measurements are assigned to tracks. Substituting (11-12) into (10) and simplifying yields the following.

\[
p[y^k] | Z^k = \sum_{\{y^k\}} p[Z_k | Z^{k-1}, \{y^k\}] p[y^k] | Z^{k-1} \psi_k
\]

(13)

\[
\psi_k = \sum_{\{y^k\}} p[Z_k | Z^{k-1}, \{y^k\}] p[y^k] | Z^{k-1} \psi_k
\]

(14)

(15)

Note that the multiplicative weights in (14) that are not present in (1) obviate in a natural way the need for the modified MHT formalism that addresses the greedy-target problem in the general tracking setting.

Cardinality tracking expressed via equivalence classes leads both to computational efficiency (there are much fewer equivalence classes over global hypothesis than there are global hypotheses) and to a well-posed formulation whereby the MAP equivalence class is of interest. Further, the weights in (14) imply structural results regarding the form of the optimal equivalence class and optimal track-extraction rules. We address this structure in the next section.

4 Structural results

It can be shown that an optimal tracking solution can be extracted from a so-called Tetris structure, where measurements are arranged as shown in Figure 1.

![Tetris structure](image)

Fig. 1. Each column corresponds to a scan of data. In each column, measurements are stacked on top of one another, with no empty slots. Tracks are extracted out of rows.
We now characterize the optimal track extraction criteria. First, for each row \( z \) in the Tetris structure, consider separately sequences of measurements divided by more that \( K \) misses, where \( K = 0 \) if \( p_d = 1 \); otherwise, \( K \) is given by the largest non-negative integer that satisfies:

\[
K < \frac{\log\left(\frac{\lambda_b \rho_X}{1 - \rho_X}\right)}{\log\left(\left|1 - \rho_X\right|\right)}.
\]  

(16)

For a sequence of length \( N \) that includes \( M \) measurements, the track score is given by (17). Note that, for tracks that start at the beginning of the data stream, the steady-state number of targets, \( \frac{\lambda_b}{\rho_X} \), applies in lieu of the birth rate \( \lambda_b \); likewise, for tracks that reach the end of the data stream, the factor \( \rho_X \) is not to be included in (17).

\[
f(M, N, z, \mathbb{Z}^N) = \frac{\lambda_b \rho_X}{1 - \rho_X} \left(\frac{\rho_d}{\lambda_b}\right)^{N-1} \left(1 - \rho_d\right)^{N-M} \prod_{i=1}^{N} \left|Z_i - z + 1\right|.
\]  

(17)

The pseudo-code below extracts the optimal track (the one with the largest track score) from a sequence of measurements of length \( N \), provided that a track exists with score \( f(\cdot, \cdot, \cdot) \) greater than unity.

Note that score \( s \) in the pseudo-code is in log space. Also, note that \( K \) can be defined a priori, independent of track ordering and the input data realization, while the track extraction methodology depends both on track ordering and the input data cardinality sequence. Measurement sequences are only extracted if they contribute to the overall posterior probability of the hypothesis equivalence class. Multiple track segments are only extracted as a single track when it is advantageous, again in terms of overall posterior probability of the hypothesis equivalence class. The procedure below is applied to each row of data, \( z \in \{1, \ldots, \max\{Z_i\}\} \).

---

For \( t=1:N \) % define score sequence
If \( |Z_t|\geq z \)
\[
l_t=\frac{|Z_t|-z+1}{p_d/\lambda_{fa}}
\]
Else\[
l_t=1-p_d
\]
End
\[
s=\log((1-p_{\text{death}}\times l_t)+s; \text{ % update absolute score}
\]
\[
t=t+1;
\]
While \( t<N \) \% extend track\[
s_star=s; \text{ % track score}
\]
\[
\omega=t; \text{ % (temporary) track end}
\]
\[
f=1; \text{ % set extraction flag to unity}
\]
\[
\text{If } s>0
\]
\[
s=\log((1-p_{\text{death}}\times l_t); \text{ % tentative track score}
\]
\[
\alpha=t; \text{ % tentative track start}
\]
\[
l=\log((1-p_{\text{death}})* l_t)+l; \text{ % update relative score}
\]
\[
s=\log((1-p_{\text{death}})* l_t)+s; \text{ % update absolute score}
\]
End\[
\text{If } s<0 \text{ % tentative track is good}
\]
\[
f=1; \text{ % set extraction flag to unity}
\]
\[
\text{omeg}=t; \text{ % (temporary) track end}
\]
\[
s_star=s; \text{ % track score}
\]
End\[
\text{While } t<N \text{ % extend good track}
\]
End\[
\text{While } f=0 \text{ & } t<N \text{ % search for a good track}
\]
End\[
\text{While } f=0 \text{ & } t<N \text{ % search for tentative track start}
\]
End\[
\text{If } f=1 \text{ % extract optimal track}
\]
Return \( \alpha, \omega, s \text{ star} \);

---

5 The psi tracker

It is possible to introduce further hypothesis aggregation in the cardinality tracker, by recognizing that we need not distinguish between which tracks are updated: we only care about how many tracks there are at any time. Note that the amount of aggregation is characterized by (12). Thus, we may replace (18) by (19).

\[
\{\hat{k}\} = q_{\text{MAP}}(Z^k) = \arg \max \rho(\hat{k} | Z^k).
\]  

(18)

\[
\hat{k}^* = \psi_{\text{MAP}}(Z^k, \psi^{-1}) = \arg \max \rho(\psi^* | Z^k, \psi^{-1}).
\]  

(19)

Use of (19) requires scan-based optimization: no structural results are available as we had for the cardinality tracker that is based on equivalence classes of hypotheses. Thus, we introduce further beneficial aggregation, at the cost of an inability to solve the problem optimally. We will see that this is advantageous; further improvements might be had with a multiple-hypothesis approach to solving (19). We refer to this approach as psi tracking.

A further level of aggregation can be introduced, recognizing that we care only about the number of tracks at a given time without distinguishing between sets of hypotheses with different numbers of births and deaths at the current time, so long as the \( \text{difference} \) remains the same. That is, we replace (19) by (20). We refer to this approach as aggregate psi tracking.

\[
\{\psi^k\} = \psi_{\text{MAP}}(Z^k, \psi^{-1}) = \arg \max \rho(\psi | Z^k, \psi^{-1}).
\]  

(19)
6 The Kalman filtering approach

We can reformulate the cardinality tracking problem by relaxing the requirement for integer numbers of tracks and using a simple Kalman filtering approach. We initialize the scalar state based on the steady-state distribution on the number of targets, and proceed via recursive estimation. We are interested in the minimizing the minimum mean absolute error (MMAE); this is achieved by identifying the median of the posterior distribution [6]. The Kalman filter provides the optimal linear filter for MMSE estimation; a more expensive numerical approach would be required to synthesize the optimal MMAE filter.

\[
\chi_{\lambda p b} = 0 | 1 , \quad (6.1)
\]

\[
\chi_{\lambda p b} = \Sigma 0 | 1 , \quad (6.2)
\]

\[
\Sigma(k + 1 | k) = [I - p_\lambda X] \Sigma(k | k) + \lambda_b , \quad (6.3)
\]

\[
L(k + 1) = -p_d \Sigma(k + 1 | k) , \quad (6.4)
\]

\[
x(k + 1 | k + 1) = x(k + 1 | k) + L(k + 1) [Z_{k+1} - p_d X(k + 1 | k) - \lambda_{fa}] , \quad (6.5)
\]

\[
\Sigma(k + 1 | k + 1) = [I - L(k + 1)p_d \Sigma(k + 1 | k) . \quad (6.6)
\]

One could instantiate a Kalman smoother for (slightly) improved performance [7]. Indeed, since the cardinality tracker utilizes batch extraction of tracks, we need not impose a latency requirement on the filtering approach here.

7 Numerical results

The following study provides initial Monte Carlo performance results that illustrate the performance of the cardinality tracking approaches considered in this paper. Parameter settings are as follows:

- Monte of Monte Carlo iterations: 10;
- Number of scans of data per realization: \( N = 250 \);
- Probability of death: \( p_X = 0.01 \);
- Birth rate: \( \lambda_b = 1 \);
- False alarm rate: \( \lambda_{fa} = 10 \).

Based on these settings, the *a priori* estimate for the number of targets is given by the steady-state distribution, or \( \frac{\lambda_b}{p_X} = 100 \). In addition to the *a priori* estimate, we have *a posteriori* estimation sequences from the various approaches discussed in this paper. The results, measured in terms of average discrepancy between the true and estimated number of targets, the *track cardinality error*, are given in Table 1.

It is worth noting that the baseline MHT returns no tracks for this data, thus exhibiting an extremely large track cardinality error. Interestingly, the cardinality tracker does not perform very: it identifies lengthy tracks, and in a sense exhibits the same difficulty as the baseline MHT, though in a much less pronounced fashion. That is, the likeliest equivalence class of hypotheses is still not representative of the statistic of interest (i.e. track cardinality). The psi tracker performs well, and nearly matches the performance of the Kalman filter. Surprisingly, the aggregate psi tracker performs worse than the psi tracker: the reason for this is not clear, and will require further investigation. Indeed, the aggregate psi tracker exhibits an unexplained lag in its estimation results.

<table>
<thead>
<tr>
<th>estimation scheme</th>
<th>track cardinality error</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>a priori</em> estimate</td>
<td>9.9716</td>
</tr>
<tr>
<td>cardinality tracker</td>
<td>11.1536</td>
</tr>
<tr>
<td>psi tracker</td>
<td>2.5020</td>
</tr>
<tr>
<td>aggregate psi tracker</td>
<td>3.3508</td>
</tr>
<tr>
<td>Kalman filter</td>
<td>2.0072</td>
</tr>
</tbody>
</table>

Table 1. Estimation errors for cardinality estimation schemes.

Sample realizations are shown in Figures 3-4.
Note that, unlike the exact cardinality sequence, it is a simple matter to achieve an accurate a priori estimate of the total number of targets $N_T$ over a given time period. This number can be estimated quite precisely by the following:

\[(7.1) \ E[N_T] = \frac{\lambda_t}{P_z} + (N - 1)\lambda_t.\]

**8 Conclusions**

This paper introduces an interesting generalization of the MHT formalism, extending it to hypothesis equivalence classes. The approach provides partial remediation for a fundamental limitation of MHT processing, whereby a single MAP solution is given rather than the (infeasible) full posterior probability distribution on target solutions. For the cardinality tracking problem considered here, this limitation is of fundamental importance.

We consider as well further extensions to the cardinality tracker, namely the psi tracker and the aggregate psi tracker, for which structural results are not available and sub-optimal processing is required. Nonetheless, we find that the psi tracker provides the best alternative to the Kalman filtering approach that allows fractional track number estimates.

While the Kalman filter performs well here, we believe cardinality tracking holds significant potential for use in large-scale surveillance problems, where localization information of varying quality exists relative to the density of targets. In such a setting, a hybrid conventional-cardinality tracking scheme might be adopted as an effective approach to group tracking [8].

As a final note, much has been documented in recent years concerning the (Cardinalized) Probability Hypothesis Density or (C)PHD filter; an accessible physical-space reinterpretation may be found in [9]. We have not compared our cardinality tracking approaches, which are based on a limiting form of the MHT recursion, to the analogous limiting form of the CPHD. Nonetheless, we note that our psi tracker, quite reassuringly, leads to similar results as the Kalman filtering formulation of the problem that allows for fractional targets. We believe the CPHD equations also ought to provide similar results as the Kalman filter in this limiting case.

**9 References**


