Fast Mutual Exclusion

Simon Maskell[a], Mark Briers[ab], and Robert Wright[a]

[a]QinetiQ Ltd, Malvern Technology Centre, St Andrews Road, Malvern, Worcs, UK
[b]Cambridge University Engineering Department, Cambridge, UK

ABSTRACT

A method is presented that circumvents the combinatorial explosion often assumed to exist when summing probabilities of joint association events in a multiple target tracking context. The approach involves no approximations in the summation and while the number of joint events grows exponentially with the number of targets, the computational complexity of the approach is substantially less than exponential. Multiple target tracking algorithms that use this summation include mutual exclusion[11, 21] in a particle filtering context and the Joint Probabilistic Data Association Filter,[7] a Kalman Filter based algorithm. The perceived computational expense associated with this combinatorial explosion has meant that such algorithms have been restricted to applications involving only a handful of targets. The approach presented here makes it possible to use such algorithms with a large number of targets.

1. INTRODUCTION

Kalman filtering[10] is used extensively in the tracking literature. The Kalman filter is optimal in a limited set of scenarios when the models used for the measurement process and the dynamics are linear and Gaussian. In real scenarios, models are rarely linear and Gaussian. However, one often finds that one can approximate the models as linear and Gaussian, and so use an Extended Kalman filter, or approximate the models with a sum of Gaussians and so use algorithms such as the popular Interacting Multiple Model[4] (IMM) and Track-Oriented Multiple Hypothesis Tracker[5] (TO-MHT), which respectively resolve ambiguity relating to manoeuvring targets and data association. Such algorithms achieve impressive performance given the succinct representation of the probability distribution that they employ.

The particle filter[8] is a state-of-the-art algorithm for tracking, which uses a set of samples to represent the probability distributions of interest. The Monte-Carlo nature of the algorithm enables particle filters to use complex nonlinear and non-Gaussian models. When approximating the models as linear and Gaussian loses vital information, the use of particle filters can offer a significant improvement in performance over Kalman Filter based approaches. However, particle filters are often perceived as computationally expensive for high dimensional problems. As a result, there has been substantial interest in approaches that can reduce this computational expense.

One high dimensional problem attracting interest is the tracking of multiple targets. One approach to tackling this problem is to consider the state of the system to be the state of all the targets stacked on top of one another. In the context of Kalman Filtering this would result in the number of elements in the covariance matrix parameterising the Gaussian distributions used growing quadratically with the number of targets. In the context of Particle Filtering, this results in a need for a large number of samples.[2] The authors assert that in both cases, the real quantities of interest are each of the targets' states; the use of the joint state space is simply a mechanism for accurately representing the targets' individual distributions. Such an approach is equivalent to approximating the joint distribution as the product of marginal distributions.

Data association is the primary source of complexity in multi-target tracking and provides the main barrier to this representation in terms of the product of marginals; at each time step one receives several candidate measurements and one needs to update all the targets' distributions while enforcing the constraint that each measurement relates to zero or one targets and each target relates to zero or one measurements. If one defines a *joint event* as a valid combination of measurements being associated with targets, then the number of such joint events grows exponentially with the number of targets. Ideally one would treat the joint event as a nuisance parameter and integrate it out and so include the effect of all the joint events on each targets' distribution. This is commonly referred to as Mutual Exclusion[11, 21] in the particle filtering literature. The same idea is used in the Kalman Filter tracking literature under the name of the Joint Probabilistic Data Association Filter, JPDAF.[7] Since the JPDAF was proposed, the approach has also been generalised so that each target's distribution is modelled as a mixture of Gaussians rather than a single Gaussian.[16]

The exponential growth of the number of joint events with the number of targets and number of measurements has meant that these approaches have hitherto been presumed too computationally expensive to be used with large numbers of targets. As a result, research has either approximated the soft decision (averaging over the joint event) with a hard decision[17, 18, 20] (finding best joint event, usually at some point in the past) or used Monte-Carlo sampling to approximate the summation involved in the soft decision.[9, 14, 15, 21] The approach described here is to exactly calculate the sum over the exponentially growing number of joint events in sub-exponential time.[12, 13]

The format of the paper is as follows: in Section 2 the concept of mutual exclusion is described using an explicit sum over the associations of targets with measurements.[21] In Section 3, a fast method for calculating the sum is discussed. In Section 4, a worked example of the tree construction, which is deemed to be the difficult conceptual element to this algorithm is provided, and then two illustrative examples are used to demonstrate the significant reduction in computational cost of implementing mutual exclusion using the approach. Finally, Section 5 concludes.

2. MULTI-TARGET PARTICLE FILTERING

The aim here is to formulate the problem faced by mutual exclusion as it relates to the summation over joint events. Hence, only a concise description is given of particle filtering, of the association of targets with measurements and of how the two relate.

2.1. Particle Filtering

While the reader unfamiliar with particle filtering is referred to the various tutorials and books available on the subject,[1, 6] the aim here is to provide some context and introduce notation.

A single target particle filter uses the diversity of a set of N samples to represent the uncertainty of the state of the target. The samples are hypotheses for the evolution of the target state, $x_{1:t} = x_1 \dots x_t$, consistent with the history of measurements, $y_{1:t} = y_1 \dots y_t$. The kth particle consists of a hypothesis for the state sequence, $x^k_{1:t}$, and an associated weight, w^k_t. At each iteration, this state sequence is extended using a *proposal* distribution which often takes the convenient form $q\left(x_t|x_{t-1}, y_t\right)$, such that the particles only need to remember the hypothesis for the current state.

$$x^k_t \sim q\left(x_t|x^k_{t-1}, y_t\right) \tag{1}$$

The weights are then updated to reflect the disparity between this proposal and the posterior from which one is (generally) unable to draw samples from, $p\left(x_t|x_{t-1}, y_t\right)$.

$$w^k_t \propto w^k_{t-1} \frac{p\left(y_t|x^k_t\right) p\left(x^k_t|x^k_{t-1}\right)}{q\left(x^k_t|x^k_{t-1}, y_t\right)} \tag{2}$$

The weights are normalised to sum to unity, but the fact that the proposal and posterior are different means that the weights necessarily become skewed over time. To counter this, a *resampling* step is used to replicate those particles with large weights and discard those particles with small weights such that the total number of samples remains constant, but the computational load is spread more evenly across the posterior. The simplest approach to resampling is to sample independently N times from the Multinomial distribution constituting the weights; more advanced approaches can reduce the variance of the weights across the different results of the resampling operation while maintaining the property that the resampling step does not introduce any bias.[1]

To enable this paper to focus on mutual exclusion and avoid unnecessary obfuscation, we will assume that the proposal used is the prior. Other proposals, such as that used in the Unscented Particle Filter,[22] are well suited to environments when the likelihood is highly peaked with respect to the prior; such proposals make each particle more computationally expensive in the hope that the overall computational cost can be reduced by having far fewer particles. However, the prior is very frequently used as a proposal and is often an expedient choice:

$$q\left(x_t|x_{t-1}, y_t\right) \triangleq p\left(x_t|x_{t-1}\right) \tag{3}$$

The reason that using the prior can often constitute a convenient choice of proposal distribution is that the weight update simplifies considerably:

$$w_t^k \propto w_{t-1}^k p\left(y_t|x_t^k\right) \tag{4}$$

2.2. Association of Targets to Measurements

There are typically multiple measurements at each time step, the ith of the N_M such measurements being y_t^i. To model the fact that each target may be undetected, another hypothesis is used. Hence, a dummy measurement is introduced so that when $i = 0$, the target is undetected and when $1 \leq i \leq N_M$, the target is assigned to the ith measurement.

The assignment of targets to measurements is assumed to be decomposed into an assignment for each of the N_T targets. The measurement to which the jth target is assigned is denoted i_j. Gating is used to select a subset of measurement hypotheses for each target which are considered candidates for the assignment; the subset is made up of the dummy and those measurements that lie in high probability regions of the target's likelihood. In any case, this assignment can be represented as an indicator vector, Ω_j, with $N_M + 1$ elements ω_{ji}:

$$\Omega_j = [\omega_{ji}] \tag{5}$$

where

$$\omega_{ji} = \begin{cases} 1 & i = i_j \\ 0 & i \neq i_j \end{cases} \tag{6}$$

The assignment for all the targets can then be represented as an indicator matrix, Ω, where each row is such an indicator vector:

$$\Omega = \left[{\Omega_1}^T \ldots {\Omega_{N_T}}^T\right]^T = [\omega_{ji}] \tag{7}$$

where A^T is the transpose of matrix A and, as before:

$$\omega_{ji} = \begin{cases} 1 & i = i_j \\ 0 & i \neq i_j \end{cases} \tag{8}$$

Since it is assumed that each measurement is assigned to zero or one targets and that each target is assigned to zero or one measurements, there are constraints on the elements of this matrix:

$$\sum_{j=1}^{N_T} \omega_{ji} \leq 1 \qquad i = 1 \ldots N_M \tag{9}$$

since each measurement is assigned to at most one target and

$$\sum_{i=0}^{N_M} \omega_{ji} = 1 \qquad \forall j \tag{10}$$

since each target is assigned to exactly one measurement hypothesis (either a hypothesis relating to a real measurement or to the dummy measurement).

It is easy to impose the second constraint by assigning one measurement hypothesis to each of the N_T targets in turn. The resulting matrix representation of the assignment will then potentially violate the first constraint. V is defined to be the set of valid assignments which do not violate the constraint in this way. So, if $\Omega \in V$, then the constraints are satisfied. Since it is possible (but computationally expensive) to enumerate all the possible assignment matrices, and since it is possible to test each resulting matrix for being in the set V, it is possible to enumerate the members of the set of valid assignments V.

An alternative approach to including assignment matrices that violate the constraints is to build up the assignment matrices using a tree of associations possible for each target. So, the topmost nodes in the tree represent the valid choices of assignments for the first target. Similarly, the second layer of nodes represent the valid assignments for the first and second targets. The jth layer of nodes then represents the valid assignments for the first to jth targets. Each node then needs to store the set of measurements that have been *used* by the targets down to the target under consideration.

To enumerate the set of valid assignment matrices, each target is considered in turn and the assignment matrices built up target-by-target. For each node for the previous target, a new node for the current target is created for each gated measurement hypotheses that does not result in any violation of the constraints. So, the tree represents an enumeration of the valid sequences of assignments for the targets by considering each target in turn.

It should be observed that although only enumerating those association matrices that satisfy the constraints does avoid some computations, the computational expense is still prohibitive for large numbers of targets.

2.2.1. Example

Consider the simple case of a set of four measurements being assigned to four tracks, where the process of gating has been conducted and the *dummy* measurement is given an index of 0. The set of valid associations for each track is summarised below in table 1.

Track number	Gated Measurement Indices
1	0,1,2,3
2	0,2,3
3	0,3,4
4	0,4

The resulting tree, which displays the enumerated hypotheses for the four tracks, can be found in Figure 1. Note that this has been placed later in the paper to allow a direct comparison with the proposed approach. Furthermore, for clarity only the first two tracks have been labelled with the measurement indices.

2.3. Probability Calculations (Integrating out the Association Matrix)

When considering multiple targets, one would like to have a set of particles for each target so the kth particle for the jth target then consists of a hypothesis, $x_t^{k_j}$, and a weight, $w_t^{k_j}$. If the assignment is known, we could update all the particle weights using the approach outlined in the previous section. For the jth target, the associated measurement, i_j, would be used in equation (4):

$$w_t^{k_j} \propto w_{t-1}^{k_j} p\left(y_t^{i_j} | x_t^{k_j}\right) \tag{11}$$

where k_j is the kth particle for the jth target.

However, the assignment is typically unknown and so can either be estimated jointly with the state or integrated out. Here, the assignment is integrated out. Hence for the jth target, (4) becomes:

$$w_t^{k_j} \propto w_{t-1}^{k_j} \sum_{i \in G_j} p\left(y_t^i, i | x_t^{k_j}\right) \quad (12)$$

$$= w_{t-1}^{k_j} \sum_{i \in G_j} p\left(y_t^i | x_t^{k_j}, i\right) p\left(i|j\right) \quad (13)$$

where G_j is the gated set of measurement hypotheses for the jth target.

$p(i|j)$ is the probability of the ith measurement being assigned to the jth target, so these probabilities sum to unity over the measurements and are calculated by marginalising over the valid set of assignment matrices:

$$\sum_{i=0}^{N_M} p(i|j) = 1 \quad (14)$$

$$p(i|j) = \frac{\sum_{\Omega \in V, \omega_{ji}=1} p(\Omega)}{\sum_{\Omega \in V} p(\Omega)} \quad (15)$$

$$= \frac{\sum_{\Omega \in V, \omega_{ji}=1} p(\Omega)}{\sum_{i=0}^{N_M} \left(\sum_{\Omega \in V, \omega_{ji}=1} p(\Omega)\right)} \quad (16)$$

where the numerator is a sum over all valid association matrices that have $\omega_{ji} = 1$ and the denominator is a sum of this quantity over all the gated association hypotheses (which ensures that the normalisation condition of (14) holds).

It is assumed that (for valid association matrices such that $\Omega \in V$) the joint association likelihood can be expressed as a product of terms for each target; this is the same approach taken by popular hard decision logic algorithms such as the auction algorithm[3]:

$$p(\Omega) = \prod_{j=1}^{N_T} p\left(\omega_{ji_j}\right) \quad (17)$$

$p\left(\omega_{ji_j}\right)$ is the likelihood based on a single-target model where $i_j = 1$, which can be calculated by marginalising over the particles:

$$p\left(\omega_{ji_j}\right) = \sum_{k_j} w_t^{k_j} p\left(y_t^{i_j} | M^{k_j}\right) \quad (18)$$

where M^{k_j} is the parameter set for the kth particle for the jth target.

In the context of particles, $M^{k_j} \triangleq x_t^{k_j}$; the parameter set is just the sampled state of the particle. In an approach based on a Gaussian mixture model, such as that in,[16] the sum would be over the mixture components, so the parameter set for each component is the mean and covariance of the relevant Gaussian component. In the case of the JPDAF, there is only a single Gaussian in the mixture, so only a single term in the summation. The idea common to all these algorithms is to sum over all the valid association matrices.

It is worth noting that the formulation of the JPDAF often described in the literature does not involve joint likelihoods with the same form as (18), but joint likelihoods that include another term, which is dependent on the number of false alarms (and so number of targets that are associated with the dummy measurement).[4] The assertion made here is that this restriction is justifiable on the basis that its use is widespread in other algorithms (such as the auction algorithm). Furthermore, the authors believe that it is the summation over joint events rather than the form of the joint likelihood that provides the key benefit of using the JPDAF. As will be discussed, the restriction makes it possible to realistically implement JPDAF for large numbers of targets, providing further strong motivation for the restriction.

3. EFFICIENT MULTI-TARGET PARTICLE FILTERING

The problem with the approach described in the previous section is that an enumeration of the valid association matrices in (16) results in a combinatorial explosion; the number of valid association matrices grows exponentially with the number of targets and measurements. The aim in this section is to present an approach to calculating exactly the numerator (and so denominator) in (16) in sub-exponential time.

3.1. Net Formation

As described to this point, mutual exclusion enumerates all possible assignment matrices that satisfy the constraints of (9) and (10).

The key idea of the approach described here is for the nodes for the jth target to enumerate the valid associations possible for the $N_T - j$ remaining targets and not the associations of the j targets considered up to the jth target. The reason for the huge computational saving that results is that several parent nodes for one target can result in the same child node.

So, an additional step is introduced that, when creating a new node for the jth target checks to see if there are any existing nodes for the jth target that can result in the same set of valid associations possible for the $N_T - j$ remaining targets. If this node exists already then it is added to the list of children of the parent node and the parent added to the child's list of parents. The parent-child relationship is then labelled with the association hypothesis, the so-called identity, that gave rise to the connection; hence there can be multiple connections between the same pair of parent and child nodes as a result of different choices of measurement hypotheses that give rise to the same set of valid possible associations for the remaining targets. Only if no node exists with the same set of valid associations possible for the $N_T - j$ remaining targets is a new node created. The result is that the *tree* is replaced by a net of interconnected nodes, referred to hereafter as a *net*.

The identity for each node is an indication of how measurement assignments made for tracks already considered affect assignments for tracks remaining to be considered. It is defined as the intersection of the current track's accumulated measurements with the union of the identity of the parent node under consideration and the measurement index to be implemented by the child node in this instance. For a current track, the accumulated measurements represent those measurement assignments for remaining (i.e. as yet unprocessed) tracks that can be affected by a measurement assignment for the tracks up to and including the current track.

The leaf nodes (the nodes for the N_Tth target) in the tree that enumerates association matrices are then equivalent to different descents through the net. The fact that descents through the net explicitly share common structure means that the net provides a succinct description of the problem. Indeed, the use of this net opens the door to summing the exponentially increasing number of terms in (16) in sub-exponential time as will be discussed in the sequel.

3.2. Example

Continuing with the example in Section 2.2.1, the resulting hypotheses net is displayed in Figure 2.

To conserve space, we only detail the node creation for the second track. The approach that we present exploits redundancy in the problem; by referring to Table 1 in Section 2.2.1 one can see that the measurement hypothesis labelled 1 only appears in the list of possible measurements for track 1. This means that in figure 1 the structure of the sub-trees resulting from the enumeration of nodes 0 and 1 in the previous track are identical. Instead of enumerating two sets of identical sub-trees, the proposed approach exploits the that fact that one can enumerate a single node representing both of the valid associations possible for the current track. The net therefore behaves like a 'road-map' of the tree.

To enable one to efficiently interpret the resulting net, the node measurement hypotheses that result in a node (from a parent to a child) are labelled together with the identity of each node.

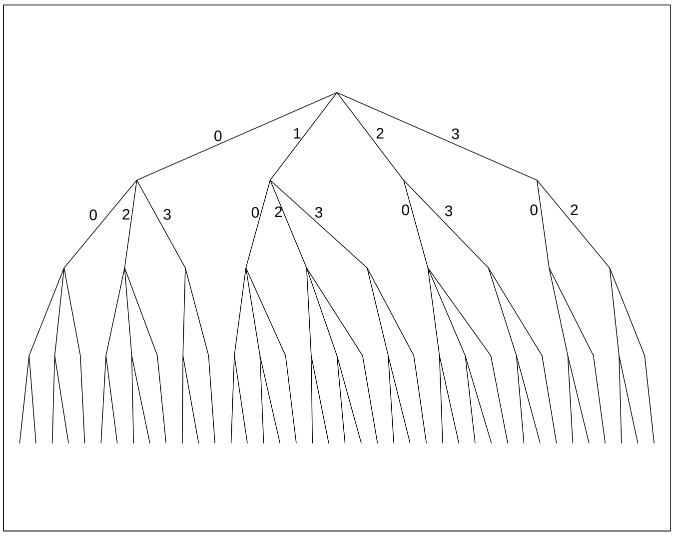

Figure 1: The tree of association hypotheses for the example presented in Section 2.2.1.

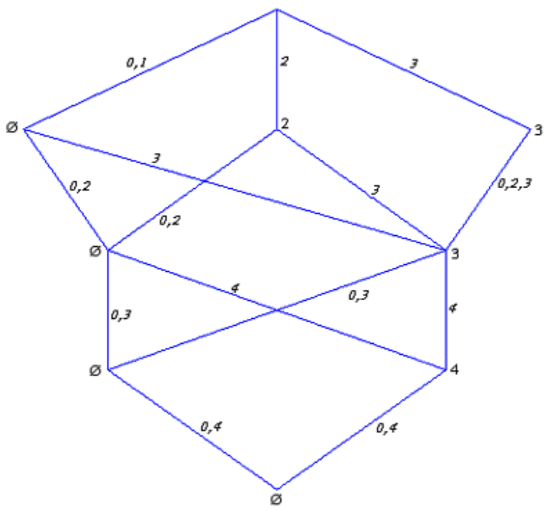

Figure 2: The net of association hypotheses for the example presented in Section 2.2.1.

3.3. Probability Calculations

It still remains to calculate the quantities in the numerator of (16) from this net. The sum is over all descents of the net that include a given element in the assignment matrix being unity. Since all descents must pass through one of the nodes in the net for the jth target, the sum over the valid hypotheses can be replaced with a sum over the nodes:

$$\sum_{\Omega\in V,\omega_{ji}=1} p\left(\Omega\right) = \sum_{n_j=1}^{N_{n_j}} p_T\left(i|n_j\right) \tag{19}$$

where there are N_{n_j} nodes for the jth target and where $p_T\left(i|n_j\right)$ is the sum of the joint likelihoods for descents that include $\omega_{ji}=1$ that go through the n_jth node (for the jth target).

If the restriction of (18) is imposed then $p_T\left(i|n_j\right)$ can be efficiently calculated as a product of three terms: a term relating to the descents to those of the nodes' parents that have the ith measurement associated to the parent-child relationship, a term relating to the connection of the node to its parent (ie. $p\left(\omega_{ji}\right)$) and a term relating to the descents possible for the remaining $N_T - j$ targets:

$$p_T\left(i|n_j\right) = p_U\left(n_j\right) p\left(\omega_{ji}\right) p_D^{\star}\left(i,n_j\right) \tag{20}$$

where

$$p_D^{\star}\left(i,n_j\right) = \sum_{n'_{j-1}\in P(n_j),R\left(n'_{j-1},n_j\right)=i} p_D\left(n'_{j-1}\right) \tag{21}$$

and where $P\left(n_j\right)$ is the list of parents of node n_j and $R\left(n'_{j-1},n_j\right)$ is the index of the measurement associated with the parent-child relationship of the n_jth node (for the jth target) and the n'_{j-1}th node (for the $(j-1)$th target).

To calculate $p_D\left(n_j\right)$ and $p_U\left(n_j\right)$, a similar approach can be used to the two steps constituting the forward-backward algorithm used in Hidden Markov Models, HMMs[19]:

$$p_D\left(n_j\right) = \sum_{n'_{j-1}\in P(n_j)} p\left(\omega_{jR\left(n'_{j-1},n_j\right)}\right) p_D\left(n'_{j-1}\right) \tag{22}$$

$$p_U\left(n_j\right) = \sum_{n'_{j+1}\in C(n_j)} p\left(\omega_{(j+1)R\left(n_j,n'_{j+1}\right)}\right) p_U\left(n'_{j+1}\right) \tag{23}$$

where it should be noted that the same node can appear more than once in the node's lists of parents, $P\left(n_j\right)$, and children, $C\left(n_j\right)$.

So, to calculate the probabilities, the downwards and upwards probabilities to each node are calculated. These probabilities are combined with the probability of each association hypothesis to give the sum of probabilities of all descents through the net that pass though each node (and are related to the given hypothesis). Hence it is possible to calculate the quantities of interest from the net.

3.4. Ordering of Targets

The approach described to this point will drastically reduce the computational expense associated with the calculation of the numerator of (16) given an ordering of the targets. However, to minimise the computational expense the targets need to be ordered such that the number of nodes in the net is minimised.

One simple step is to adopt the same approach as described by Werthmann[23]; the targets are divided into clusters. The clusters are defined such that the cluster's set of measurements has no intersection with the other clusters' sets of measurements. These sets of measurements for each cluster are the union of all the gated

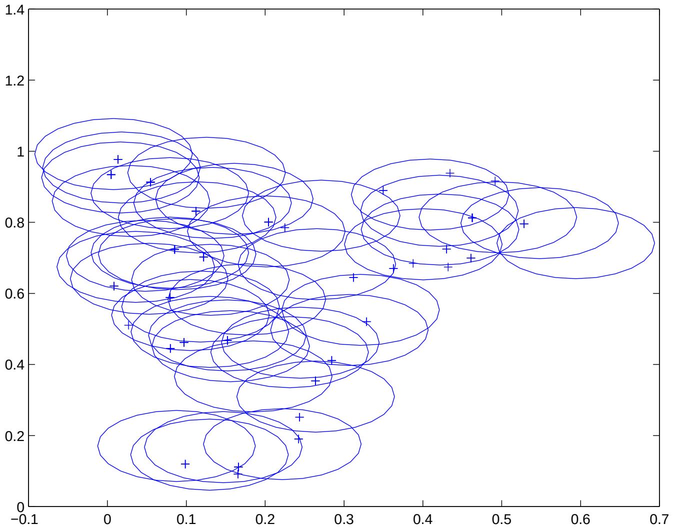

Figure 3: Illustration of gating arrangement for example involving a large number of targets.

measurements for all the targets within the cluster. The sums in (16) then only need consider the cluster containing the given target since all the joint likelihoods (for all the targets) factorise with one term for each of the clusters.

This clustering can be thought of as an ordering of the targets according to the cluster to which each target belongs. The extension to ordering the targets within the clusters is non-trivial to solve optimally, so a simple heuristic approach, that has been found to be effective, is advocated here.

The idea is to start with N_T lists of targets where the jth list initially contains just the jth target. At each iteration of the ordering algorithm, the union of the gated measurements for all the targets in each list is calculated. Then the pair of lists with the largest number of measurements common to both lists' unions is found. This pair of lists is then combined. The combination process considers potentially reversing the order of each of the lists and appending the lists in either order: a appended with b; a reversed and appended with b; b appended with a; b appended with a reversed. The combination is chosen that minimises the intersection between the newly adjoining targets' sets of gated measurements. This new combined list then replaces the two constituent lists. This process is repeated N_T times until there is one list of targets.

4. RESULTS

To demonstrate the gains offered by the approach, an exemplar data association problem is considered. For the example considered, the number of targets in the cluster is large enough to make it infeasible to use the standard hypotheses enumeration technique.

To generate the problems, a cluster of 80 targets and 80 measurements were sampled uniformly across an area such that $0 < x < 1$ and $0 < y < 1$. The gates for each target were circular regions centered on each target such that a measurement was considered within the gate if the squared distance of the measurement from the target was less than $d = 0.01$. Based on the data generated, the largest cluster that we subsequently analyse, consists of 33 measurements and 41 targets and is shown in Figure 3.

The total number of association matrices that satisfy (10) is greater than 10^{22}. While the number of valid association matrices will be smaller, it is unlikely to be a very different order of magnitude. Indeed, the combinatorial explosion results in the tree based approach being unable to be implemented in a useful timescale.

The net resulting from the proposed approach is shown in Figure 4 and, at its widest, is 64 nodes wide. As a result the terms in (16) can be calculated in a few seconds on a desktop PC.

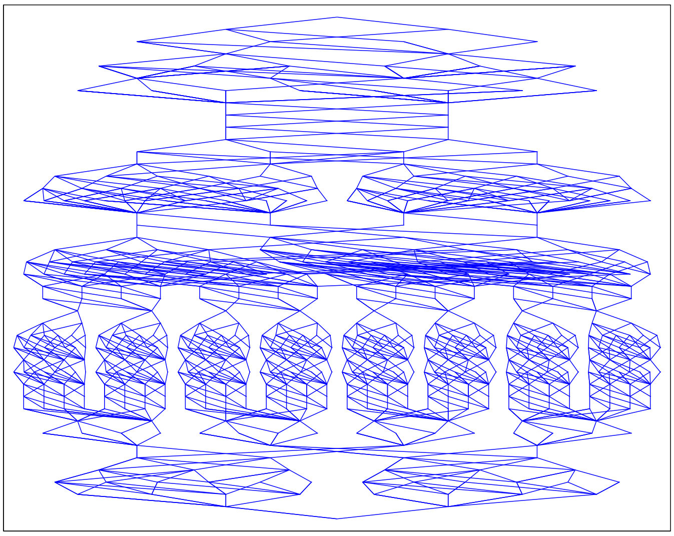

Figure 4: Net Multi-target tracking example using computationally efficient approach

5. CONCLUSIONS

A fast method has been described that can be used to exactly calculate the sum over an exponentially growing number of joint events in substantially less than exponential time. It has been shown that it is now possible to use mutual exclusion to track a large number of targets.

Acknowledgements

The authors would like to thank the Royal Commission for the Exhibition of 1851 (www.royalcommission1851.org.uk). This research was supported by the UK MoD Corporate Research Programme.

REFERENCES

1. S Arulampalam, S R Maskell, N J Gordon, and T Clapp. A tutorial on particle filters for on-line non-linear/non-Gaussian Bayesian tracking. *IEEE Transactions on Signal Processing*, 50(2):174–188, 2001.
2. D Ballantyne, H Chan, and M Kouritzin. A branching particle-based nonlinear filter for multi-target tracking. In *The 4th International Conference on Information Fusion*, August 2001.
3. Y Bar-Shalom and W D Blair, editors. *Multitarget-Multisensor Tracking: Applications and Advances Volume III.* Artech House, Boston/London, 2000.
4. Y Bar-Shalom and X R Li. *Multitarget-Multisensor Tracking: Principles and Techniques.* Academic Press, 1995.
5. S Blackman and R Popoli. *Design and Analysis of Modern Tracking Systems.* Artech House, Boston/London, 1-58053-006-0, 1999.
6. A Doucet, J F G de Freitas, and N J Gordon, editors. *Sequential Monte Carlo Methods in Practice.* Springer-Verlag: New York, 2001.
7. T E Fortmann, Y Bar-Shalom, and M Scheffe. Multi-target tracking using joint probabilistic data association. *Proceedings of the 1980 IEEE Conference on Decision and Control, Albuquerque*, pages 807–812, 1980.
8. N J Gordon, D J Salmond, and A F M Smith. Novel approach to nonlinear/non-gaussian Bayesian state estimation. *IEE Proceedings - F*, 140(2):107–113, 1993.

9. C Hue, J-P Le-Cadre, and P Perez. Sequential monte carlo methods for multiple target tracking and data fusion. *IEEE Transactions on Signal Processing*, 50(2), February 2002.
10. R E Kalman. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, pages 35–46, 1960.
11. J MacCormick and A Blake. A probabilistic exclusion principle for tracking multiple objects. In *International Conference on Computer Vision, Kerkyra, Corfu*, 1999.
12. S R Maskell. Efficient hypothesis management. UK Patent Office Application Number 0315349.1. Filed on 1 July 2003.
13. S R Maskell. Efficient hypothesis management - with application to multi-target tracking. Technical report, DERA, June 2001. RESTRICTED - COMMERCIAL.
14. S R Maskell, N J Gordon, and M Rollason D J Salmond. Efficient multitarget tracking using particle filters. In *SPIE, Signal and Data Processing of Small Targets*, 2002.
15. M Orton and A Marrs. A Bayesian approach to multi-target tracking data fusion with out-of sequence measurements. Technical report, CUED, 2001.
16. L Y Pao. Multisensor multitarget mixture reduction algorithms for target tracking. *Journal of Guidance Control and Dynamics*, 17:1205–1211, 1994.
17. K R Pattipati, Y Bar-Shalom, T Kirubarajan, and S Deb. A generalized S-D assignment algorithm for multisensor-multitarget state estimation. In *INFORMS*, October 1998.
18. A B Poore, N Rijavec, T N Barker M, and Munger. Data association problems posed as multidimensional assignment problems: Algorithm development. In *SPIE, Signal and Data Processing of Small Targets*, pages 172–182, 1993.
19. L R Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–285, 1989.
20. B D Reid. An algorithm for tracking multiple targets. *IEEE Transactions on Automatic Control*, AC-24(6):843–854, 1979.
21. D Schulz, W Burgard, D Fox, and A B Cremers. Tracking multiple moving targets with a mobile robot using particle filters and statistical data association. In *IEEE International Conference on Robotics and Automation*, pages 1665–1670, 2001.
22. R van der Merwe, A Doucet, J F G de Freitas, and E Wan. The unscented particle filter. In T K Leen, T G Dietterich, and V Tresp, editors, *Advances in Neural Information Processing Systems 13*. MIT Press, 2000.
23. J R Werthmann. A step-by-step description of a computationally efficient version of multiple hypothesis tracking. In *SPIE, Signal and Data Processing of Small Targets*, 1992.