

Track-Oriented Multiple Hypothesis Tracking Based on Tabu Search and Gibbs Sampling

Shaoming He, Hyo-Sang Shin, Antonios Tsourdos

Abstract

In order to circumvent the curse of dimensionality in multiple hypothesis tracking (MHT) data association, this paper proposes two efficient implementation algorithms using Tabu search and Gibbs sampling. As the first step, we formulate the problem of generating the best global hypothesis in MHT as the problem of finding a maximum weighted independent set of a weighted undirected graph. Then, the metaheuristic Tabu search with two basic movements is designed to find the global optimal solution of the problem formulated. To improve the computational efficiency, this paper also develops a sampling based algorithm based on Gibbs sampling. The problem formulated for the Tabu search based algorithm is reformulated as a max product problem to enable implementation of Gibbs sampling. The detailed algorithm is then designed and the convergence is also theoretically analysed. The performance of the two algorithms proposed are verified through numerical simulations and compared with that of a mainstream MDA implementation algorithm. The simulation results confirm that the proposed algorithms significantly improve the computational efficiency while maintaining or even enhancing the tracking performance.

Index Terms

Multiple Target Tracking, Multiple Hypothesis Tracking, Best Global Hypothesis Generation, Tabu Search, Gibbs Sampling

I. INTRODUCTION

Multiple target tracking (MTT) is a vital problem that arises in many surveillance systems, employing sensors to collect information of interest from the environment. Application examples of MTT include, but are not limited to, radar-based tracking of moving targets [1], sonar-based tracking of submarines [2], video-based tracking of pedestrians [3]–[8], extended target tracking [9] and biomedical cells [10]. The primary aim of MTT is to simultaneously estimate target trajectories and maintain the confirmed tracks for multiple targets of interest. Therefore, the two essential problems of MTT are: (1) state estimation/prediction; and (2) data association.

It is obvious that data association becomes more challenging in MTT than in single target tracking (STT). MTT usually suffers from high possibility of having multiple false alarms and clutters, i.e., measurement origin uncertainties, besides target-generated observations. Moreover, MTT encounters with target birth and death. All these facts make the data association complicated and hence a key issue in MTT [11], [12].

To this end, this paper tackles with MTT, especially focusing on the data association problem. The most fundamental and conventional data association for MTT is the nearest neighbour (NN) filter [13]. NN aims to find the observation nearest to each target, among all observations in the target validation region. This approach is simple and requires low computation demands. However, as this approach prunes all other feasible measurements from the association, its performance could significantly suffer in dense clutter environments. The global NN (GNN) filter is an extension of the NN filter. In the GNN filter, all possible associations within the gates are taken into account. To find the best association among all possible association, GNN applies an optimisation algorithm such as Jonker-Volgenant-Castanon or auction algorithm [14]. Nonetheless, since GNN also considers only the most likely measurement to each target, it has no measure to counteract the measurement uncertainty and only works well for the environments with low false alarm and widely spaced targets.

Considering the issues with the NN and GNN filters, it is desirable to conduct data association in a probabilistic manner [15]. The most well-known one-scan probabilistic data association for the MTT problem is joint probabilistic data association (JPDA) filter [16], [17]. JPDA is known to be a sub-optimal approximation (in the sense of minimum mean square error) for optimal Bayesian estimation. In JPDA, the target's state is approximated by a single Gaussian model with the moment matched to its corresponding multiple Gaussian mixtures. The key idea of JPDA is to enumerate all possible associations to calculate the marginal association probability. The issue is that calculating the marginal association probability is known to be #P-complete. To reduce the computational burden, some approximations of JPDA such as cheap JPDA [18], suboptimal JPDA [19] and near-optimal JPDA [20] were developed in the past few years.

Recent dramatic advance in computational capabilities has led to another widely-used data association approach, namely multiple hypothesis tracking (MHT). Unlike JPDA, MHT utilised several consecutive scans to solve the data association problem in a delayed logic and is known as a maximum a posteriori (MAP) estimator. Generally, MHT can be categorised into two main types: (1) hypothesis-oriented; and (2) track-oriented. The hypothesis-oriented MHT, originally proposed by

Shaoming He, Hyo-Sang Shin and Antonios Tsourdos are with the School of Aerospace, Transport and Manufacturing, Cranfield University, Cranfield MK43 0AL, UK (email: {shaoming.he,h.shin,a.tsourdos}@cranfield.ac.uk)

Reid [21], forms and expands a large number of hypotheses scan to scan for data association. On the other hand, the track-oriented MHT [22] maintains track trees for incompatible tracks and discards all the unreliable tree branches (track hypotheses) formed on last scans based on the best global hypothesis. In both categories, how to generate global hypotheses is a key part for the implementation. The naive brute force enumeration of all hypotheses is impractical for the real-time application as the number of hypotheses exponentially increases. In case of the hypothesis-oriented MHT, Cox [23] proposed an efficient method by considering only m -best hypotheses using Murty's ranking algorithm [24]. As to track-oriented MHT, multiple dimensional assignment (MDA) [25]–[27] is the mainstream for the implementation. The MDA problem with scan depth of more than two is known to be NP-hard. Solutions of this problem are generally obtained by an approximate algorithm such as Langrangian relaxation approach [14], [25], m -best assignment [14], and linear programming [28]. Note that the linear programming solution requires problem relaxation by allowing non-integer value for the association event and therefore the integral property of the solution is not ensured. This brings up the question of how to deal with the fractional assignments. In [28], the Langrangian relaxation solution was shown to be better than the m -best assignment and linear programming implementations. Recently, the MDA formulation of MHT was proven to be equivalent to the one finding the maximum weight independent set (MWIS) of a weighted undirected graph [29]. The MWIS formulation of MHT is easier not only to understand, but also to implement. However, the MWIS problem is still an NP-hard problem. To make it computationally tractable, one needs to seek to heuristic approaches or countable approximations. In [6], the authors utilised the MWIS formulation of MHT in the task of multi pedestrian tracking, where the MWIS problem was solved by a line search method. However, the quality of line search is highly dependent on the initial guess and hence easy to be trapped in a local optimum. Another MWIS implementation algorithm of MHT can be found in [30], where the exact solution [31] was sought at the price of increased computational demands. In order to relax the computational burden while providing reasonable results, some other approaches such as the evolutionary algorithm [32] or heuristic approach [33] are also implemented to the MWIS problem.

Thanks to its MAP property, this paper adapts MHT as the baseline data association approach. Especially, the track-oriented MHT is considered since it is proven to be more efficient than its hypothesis-oriented counterpart for practical implementation [34], [35]. The main focus of this paper is to address the issue with computational efficiency in the track-oriented MHT, without sacrificing the tracking performance.

The contribution of this paper is the development of two different implementation algorithms for the track-oriented MHT to mitigate the computational efficiency issue. The first approach takes advantages of the MWIS formulation and develops a metaheuristic Tabu search algorithm to find the best global hypothesis for the MWIS problem formulated. Hence, the first algorithm is named as Tabu-MHT algorithm in this paper. The Tabu search based MHT implementation algorithm is also a new approach for the MWIS problem. The Tabu search is chosen because the MWIS problem is a combinatorial optimization problem and the heuristic approach is considered to be well-suited to this type of problem [36]–[38]. At each Tabu iteration, two basic movement, namely add and swap, are used to find a better neighbourhood solutions of the initial candidates. A trigger policy is also proposed in the Tabu search to enhance the capability of escaping from the local optimum, i.e. the capability of finding the global optimum.

The second approach proposed exploits countable approximation based on a stochastic sampling method. In this approach, we reformulate the max summation MWIS problem to a new max product decoding (MPD) problem. The key idea of the newly proposed formulation is to consider each solution, that is association, as a random variable with a distribution. The product in the MPD problem is represented by this distribution and corresponds to the weight/track score in the MHT problem. This paper suggests to adopt the stochastic Gibbs sampling to approximate the distribution, so called Gibbs-MHT algorithm. It is then proven that the Markov chain generated by the proposed Gibbs sampling asymptotically converges to the target invariant distribution.

For the verification, the performance of the two proposed approaches is investigated via numerical simulations and compared with the mainstream Langrangian relaxation-based MDA algorithm. The two main performance indices considered are the tracking performance and running time. The simulation results with different scenarios show that the two proposed approaches outperform the mainstream MDA-MHT algorithm, especially for large scale problems. This confirms that the two approaches developed improve the computational efficiency of the MHT without sacrificing the tracking performance.

The rest of the paper is organised as follows. Sec. II presents some preliminaries of MHT, especially the track-oriented one. Sec. III provides different formulations of MHT. In Sec. IV, the proposed Tabu search-based implementation algorithm is detailed, followed by Gibbs sampling-aided implementation algorithm represented in Sec. V. Finally, some simulation results and conclusions are offered.

II. PRELIMINARIES

This section presents some concepts and implementation issues regarding track-oriented MHT.

A. Basic Concepts

The set of measurements received at scan k is defined as

$$Z_k = \{z_0^k, z_1^k, z_2^k, \dots, z_{M_k}^k\} \quad (1)$$

where M_k is the number of measurements received at scan k , z_i^k ($i \neq 0$) the i th measurement received at scan k , z_0^k the dummy measurement for the convenient representation of miss detection and false alarm.

A track or track hypothesis is defined as a sequence of measurements given by

$$T_i^k = \{z_{i_1}^1, z_{i_2}^2, \dots, z_{i_k}^k\} \quad (2)$$

where $z_{i_j}^j$ with $i_j \in \{0, 1, \dots, M_j\}$ represents the i_j th measurement that is hypothesised to be allocated to the i th track at scan $j \in \{1, \dots, k\}$.

The track-oriented MHT aims to find the most probable global hypothesis and use the matched state estimation from this particular hypothesis for the track maintenance. A global hypothesis is a set of tracks that are compatible, that is not in conflict: any two tracks in a global hypothesis do not share any common non-dummy observations at any scan and the two tracks cannot originate from one common node.

To find the best global hypothesis, MHT keeps a track tree in several consecutive scans for the track maintenance. A track tree in MHT encapsulates multiple tracks, also called track hypotheses, as each track carried from previous scan induces potentially many new tracks. Moreover, at each new scan, a new track tree is constructed for each measurement to consider the possibility the new measurement being a new target.

Each track hypothesis in MHT is associated with a track score, which is typically defined as a log likelihood ratio (LLR) $L(T_i^k)$ between the true target hypothesis probability and false alarm hypothesis probability. The true target hypothesis assumes that the sequence of measurement comes from the same target whereas the false alarm hypothesis assumes the sequence of observations originates from clutter. As there is no prior knowledge regarding the clutter measurements, we make the general assumptions in the MTT problem that the number of clutters or false alarms is locally Poisson distributed [39]. Let λ_F denote the expected number of clutters per unit volume of the validation gate, known as spatial density of clutters. The number of new targets appearing in the space is also assumed to be Poisson distributed. The expected number of new targets per unit volume, known as spatial density of new targets, is denoted as λ_N .

Following [14], the LLR score of track T_i^k can be computed recursively as

$$L(T_i^k) = L(T_i^{k-1}) + \Delta L(T_i^k) \quad (3)$$

where the incremental term $\Delta L(T_i^k)$ is defined as

$$\Delta L(T_i^k) = \begin{cases} \ln(1 - P_D P_G), & i_k = 0 \\ \ln \left[\frac{p(z_{i_k}^k | T_i^k) P_D}{\lambda_F} \right], & i_k \neq 0 \end{cases} \quad (4)$$

where P_D denotes the probability of detection, P_G the gating probability, $p(z_{i_k}^k | T_i^k)$ the track conditioned measurement likelihood. When a new track hypothesis is created, the track score of the new hypothesis is initialised as

$$L(T_i^1) = \ln \left(\frac{\lambda_N}{\lambda_F} \right) \quad (5)$$

The status of a track is usually characterised by the classical sequential probability ratio test (SPRT) based on its current score versus the upper and lower thresholds (T_{low}, T_{up}). Generally, a track is confirmed once its score exceeds the upper threshold T_{up} while a track is immediately deleted if its score is below the lower threshold T_{low} . If the track score falls between the upper and lower thresholds, the corresponding track is still tentative and is required to be further tested.

Following the standard SPRT, the upper and lower thresholds are defined as

$$T_{low} = \ln \left(\frac{\beta}{1 - \alpha} \right), \quad T_{up} = \ln \left(\frac{1 - \beta}{\alpha} \right) \quad (6)$$

where α is the false track confirmation probability, and β the true track deletion probability. As stated in [14], the allowable false track confirmation probability α can be defined from the system requirements on false track initiation. For example, if the system generates N_{FA} false alarms per scan and only N_{FC} false track confirmations are allowed within N_S scans, then, one can set $\alpha = N_{FA} / (N_S N_{FC})$. Note that the choice of β is less important since the true track deletion probability β has less effect on the track confirmation threshold. Typical choice of β is $\beta \leq 0.1$.

B. Implementation Issues

The key feature of MHT is the delayed-decision for data association, i.e. the track-oriented MHT postpones the confirmation of data association until decision ambiguity is cleared. It is known that maintaining as many track hypotheses as possible is advantageous to facilitate an irrevocable hard decision, delaying the association decision until enough data information is available.

The issue with keeping as many hypotheses as possible is that the number of track hypotheses, even for each track tree, exponentially increases. As the number of tracks of the track trees increases in an exponential way, it becomes intractable to

keep all tracks of each tree. To handle the issue with the combinatorial explosion of track hypotheses, it is essential to develop a computationally efficient algorithm in the MHT implementation.

In practice, gating is used to reduce the computational burden of the MHT data association [14]. When a new scan of data is received, the widely accepted ellipsoidal gating procedure is usually performed to remove the most unlikely measurements of each track. Only those measurements that are within the validation region are considered for the update of tracks. A typical ellipsoidal gate is given by [14]

$$\left(z_i^k - \hat{z}_j^{k|k-1}\right)^T S_k^{-1} \left(z_i^k - \hat{z}_j^{k|k-1}\right) \leq \gamma \quad (7)$$

where $\hat{z}_j^{k|k-1}$ denotes the predicted measurement of the j th track and S is the covariance of the innovation term $z_i^k - \hat{z}_j^{k|k-1}$. γ is the maximum Mahalanobis distance that controls the volume of the gate.

Clustering is another popular way to break the large association problem into small sub-problems. The clustering process partitions the full set of track hypotheses into a number of disjoint groups and thus the measurement-to-target association can be done in a parallel way to improve the computational efficiency. Each cluster is a collection of incompatible trees. An efficient clustering procedure can be found in [22].

Note that it is widely accepted that gating and clustering can relax the computational burden only up to a certain degree [14]. Therefore, it is still required to develop an implementation algorithm to further reduce the computational load.

III. PROBLEM FORMULATION

As discussed in the previous section, the aim of the track-oriented MHT is to find the best global hypothesis. The problem of finding the best global hypothesis can be formulated as an optimisation problem. The mainstream solution of this problem is called MDA formulation. Another commonly used formulation is called MWIS formulation. Since it is known that the MWIS formulation is more intuitive and easy to implement, this paper adapts the MWIS formulation as the baseline formulation for the proposed track-oriented MHT algorithms. This subsection first introduces the mainstream MDA formulation and then shows how to convert the MDA problem into an equivalent MWIS problem.

A. MDA Formulation

Let $s_i = L(T_i^k)$, N denote the number of tracks at current scan k , and M denote the summation of the number of resolved tracks at the root node (i.e., the confirmed tracks at the root scan) and the number of non-dummy measurements during the last $p-1$ scans, where p is the scan depth. Then, the best global hypothesis can be generated by solving a binary optimisation problem, called MDA, as

$$\begin{aligned} & \arg \max_a s \cdot a \\ & \text{s.t. } Aa = b, \quad b_j = 1, \quad j = 1, 2, \dots, M \\ & \quad \quad a_i \in \{0, 1\}, \quad i = 1, 2, \dots, N \end{aligned} \quad (8)$$

where $s = [s_1, s_2, \dots, s_N]$ is a vector of track scores. Here, $a_i \in \{0, 1\}$ is a binary variable indicating inclusion of the i th track in the solution, 1 for being included and 0 for being excluded. Matrix A is a $M \times N$ binary matrix that is used to satisfy the condition: (1) any two tracks in a global hypothesis are compatible; and (2) one measurement can only be assigned to one track. The optimisation problem formulated in (7) is known to be NP-hard when the scan depth is greater than 2 [25]–[27]. Approximate solutions of (8) can be obtained by Langrangian relaxation, m -best assignment or linear programming [28].

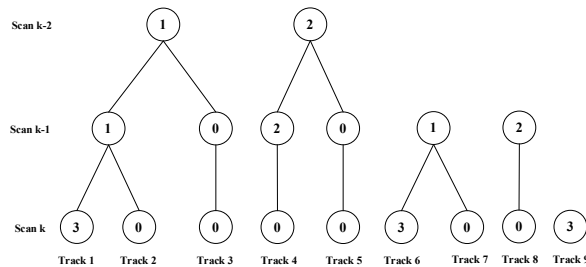


Fig. 1: An example of track tree for MDA formulation.

An example of track tree for the MDA formulation with a three scan depth is shown in Fig. 1, where the numbers represent the measurement IDs that are associated with tracks. In this example, we assume that we have two resolved/confirmed tracks at scan $k-2$. At scan $k-1$ and k , the received measurement sets are $\{z_1, z_2\}$ and $\{z_3\}$, respectively. Based on the MDA formulation, the matrix A in (8) is given by

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \quad (9)$$

where the first two rows are used to satisfy the constraint that any two tracks in a global hypothesis are compatible and the rest three rows are utilised to guarantee that one measurement can only be assigned to one track.

Ideally, all the candidate track hypotheses need to be considered in the solution space. The issue is that the number of track hypotheses exponentially increases. This implies considering all the tracks may become intractable even with heuristic or approximated approaches. One most commonly accepted approach to handle this issue is track-level pruning, which removes tracks with low track hypothesis probability. Once pruned, only remaining track hypotheses can be considered in the solution space.

Track-level pruning can easily be performed using the approximated track hypothesis probability (ATHP) [40]. Based on the law of total probability, the track hypothesis probability $p(T_i^k)$ can be calculated as

$$p(T_i^k) = \sum_{H_i^k \ni T_i^k} p(H_i^k) \quad (10)$$

where $p(H_i^k)$ denotes the probability of the global hypothesis H_i^k , and is defined as

$$p(H_i^k) = \frac{\exp(L(H_i^k))}{1 + \sum_{H_i^k \in \mathcal{H}} \exp(L(H_i^k))} \quad (11)$$

where $L(H_i^k) = \sum_{T_i^k \in H_i^k} L(T_i^k)$ is the LLR score of the global hypothesis H_i^k , \mathcal{H} being the set containing all global hypotheses.

Without track-level pruning, it could be intractable to enumerate all the global hypotheses. Applying the target-level pruning, poor track hypotheses can be removed and consequently only countable global hypotheses can be considered in enumeration.

B. MWIS Formulation

The best global hypothesis generation problem can be also formulated as an MWIS problem of a weighted undirected graph. An independent set of an undirected graph is defined as a subset with no pairwise adjacent elements of the vertex set. The MWIS is the heaviest subset of non-adjacent nodes/vertices of a graph. A weighted undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{W})$ at each scan for this problem is constructed by assigning a track T_i^k to a graph vertex v_i and each vertex with a weight w_i corresponds to its track score $L(T_i^k)$. If two tracks cannot co-exist within a node or are incompatible due to shared observations, an edge is constructed to connect them. Let $|\mathcal{V}|$ denote the cardinality (the number of elements) of vertex set \mathcal{V} . Then, the problem of best global hypothesis generation can be transformed to finding the MWIS as

$$\begin{aligned} & \arg \max_x \sum_i^{|\mathcal{V}|} w_i x_i \\ & \text{s.t. } x_i + x_j \leq 1, \quad \forall (v_i, v_j) \in \mathcal{E} \end{aligned} \quad (12)$$

where $x_i \in \{0, 1\}$ is a binary variable to indicate that whether or not the i th track is included in the solution. To improve the computational efficiency, this paper assumes that track-level pruning is performed. Therefore, the track hypotheses remained after track pruning are considered in the solution space.

An example of track trees within 3 scan depth is illustrated in Fig. 2, in which the numbers represent the measurement IDs that are associated with tracks. For this simple example, tracks 1, 2 and 3 originate from a common parent node 1; tracks 4 and 5 originate from a common node 3; tracks 2, 3 and 5 share a common non-dummy measurement 2. Based these observations, one can construct an undirected graph as shown Fig. 3, where the numbers are track IDs.

Note that the MWIS problem is also known to be NP-hard. Therefore, it is imperative to find some heuristic or countable approximation solutions.

IV. TABU-MHT ALGORITHM

This section proposes a Tabu search based approach to find the best global hypothesis. Tabu search is one of the state-of-the-art general metaheuristic algorithms to find the approximate optimal solutions for a wide variety of hard combinatorial optimisation problems [41]. Tabu search usually consists of two phases: initial solution generation phase and local search phase. Compared with other heuristic algorithms, Tabu search enhances the performance of local search by introducing prohibition rules, which prevent the search from visiting the previously-searched candidate solutions.

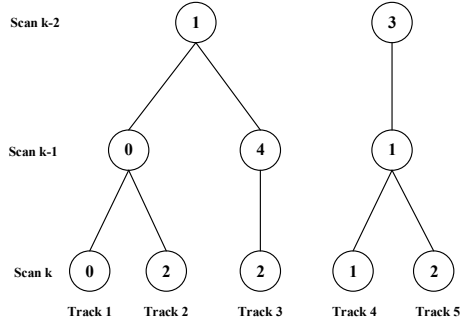


Fig. 2: An example of track tree.

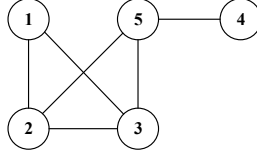


Fig. 3: Undirected graph of the considered example.

For any feasible solution $\mathcal{F} = \{i \in \mathcal{V} : i \neq j \forall j \in \mathcal{V} \text{ s.t. } \{i, j\} \notin \mathcal{E}\}$, the quality of the solution, denoted as $W(\mathcal{F}) = \sum_i^{|\mathcal{V}|} w_i x_i$, is quantified by the total weight (summation) of all vertices. Given two solutions \mathcal{F}_1 and \mathcal{F}_2 , \mathcal{F}_1 is better than \mathcal{F}_2 if and only if $W(\mathcal{F}_1) > W(\mathcal{F}_2)$. The object function of Tabu search is thus given by $W(\mathcal{F})$. The aim here is to find \mathcal{F} maximising $W(\mathcal{F})$ using the two phases, initial phase and local search phase.

A. Initial Solution Generation

Let Θ denote the search space that composed of all possible solutions. The goal of the initial phase is to randomly generate a set of possible candidate solutions. To achieve this, we first sort all tracks in a descending manner based on their track scores. The initial phase starts from an initial solution \mathcal{F} by randomly selecting a unique track vertex from the first N_1 sorted tracks as these N_1 tracks have relatively higher scores. Then, iteratively, randomly pick up a compatible track vertex $v = \{v : v \in \mathcal{V} \setminus \mathcal{F}, (v, l) \notin \mathcal{E}, \forall l \in \mathcal{F}\}$ and add it into \mathcal{F} until there is no feasible vertex can be added based on the independence constraint. In general, the initial phase only aims to find a candidate solution, i.e. a global hypothesis, and ignores the quality of the solution, but ensures a rich diversity of the initial solutions for local search in a different region. This is achieved by using a search length strategy that triggers a new initial solution if the quality of the previous solution never improve for a number of consecutive iterations.

B. Local Search Phase

Local search is used to provide an attempt to find better immediate neighbourhood solutions for the initial solutions. A neighbourhood of a candidate solution confines the search space and is typically defined by a move operator *move*. This operator transforms a given solution \mathcal{F} to a better neighbourhood solution, that is $\bar{\mathcal{F}} \triangleq \mathcal{F} \oplus \text{move}$. The proposed algorithm considers two natural operators: *ADD*(i) and *SWAP*(i, j). These two operators are defined based on the two sets, \mathcal{M}_A and \mathcal{M}_S . The set \mathcal{M}_A is composed of the vertices that are all excluded from the candidate solution \mathcal{F} and not adjacent to all vertices of \mathcal{F} . The set \mathcal{M}_S consists of the vertices that are all excluded from the set \mathcal{F} and adjacent to only one vertex in \mathcal{F} . The operator *ADD*(i) is defined as adding a vertex in \mathcal{M}_A to \mathcal{F} . The neighbourhood solution resulted from this operator is $\bar{\mathcal{F}}_A = \{\bar{\mathcal{F}}_A : \mathcal{F} \oplus \text{ADD}(i), i \in \mathcal{M}_A\}$. Clearly, the *ADD*(i) move operator always improves the quality of a solution. The operator *SWAP*(i, j) is defined as replacing an existed vertex $j \in \mathcal{F}$ with its adjacent vertex $i \in \mathcal{M}_S$ to generate a neighbourhood solution: $\bar{\mathcal{F}}_S = \{\bar{\mathcal{F}}_S : \mathcal{F} \oplus \text{SWAP}(i, j), i \in \mathcal{M}_S\}$. This operator can either improve or deteriorate the quality of a solution. From this analysis, it is clear that an overall better solution can be selected by a union of these two neighbourhood solutions, i.e., a hybrid movement. Note that both \mathcal{M}_A and \mathcal{M}_S could be an empty set if there is no such sets exist.

Tabu search introduces the so-called Tabu list to discourage the search from coming back to previously-examined solutions. Invoking the fact that the operator *ADD*(i) always improves the quality of the current solution, the prohibition rule in the proposed algorithm is designed as: a vertex that joins the solution \mathcal{F} is free to leave \mathcal{F} while a vertex that is removed from the solution is not allowed to be added again into \mathcal{F} for next T_S iterations (Tabu tenure) except for the case it provides a better solution.

Consider the undirected graph shown in Fig. 4, where the weights are the same as their IDs. Suppose the candidate solution is given by $\mathcal{F} = \{1, 4, 5, 7\}$. For $ADD(i)$, it is easy to verify that is \mathcal{M}_A an empty set based on the independence constraint. Similarly, one can obtain $\mathcal{M}_S = \{6, 8, 9\}$ for $SWAP(i, j)$ and thus replacing $\{5\}$ with $\{6\}$ and replacing $\{7\}$ with $\{8\}$ or $\{9\}$ results in a better solution. Since the weight of node 9 is larger than that of node 6 and 8, the local search automatically choose node 9 at the current iteration, leading to an improved candidate solution $\mathcal{F} = \{1, 4, 5, 9\}$. In the next iteration, one can also obtain $\mathcal{M}_A = \{8\}$ for $ADD(i)$ and $\mathcal{M}_S = \{6, 7\}$ for $SWAP(i, j)$. Since that the node 7 is excluded in the previous update step and thus will be automatically blocked by the local search phase through a Tabu list as discussed above. Consequently, the most favourable solution in terms of its quality obtained by these two operators at the this iteration is $\mathcal{F} = \{1, 4, 6, 8, 9\}$.

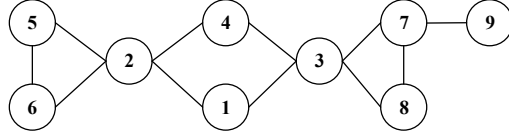


Fig. 4: Undirected graph of the considered example.

The proposed algorithm examines the possible neighbourhood solutions based on their qualities at every iteration and selects the most preferable one. The prohibition rule provides the possibility of local diversification by preventing the previously-visited solutions from being re-searched for next several iterations. A new restart from a different randomized initial solution is triggered if the quality of the solution obtained by the local search never improves for a maximum allowable consecutive search depth. Although the heuristic approach cannot have theoretical guarantee of the optimal solution, the proposed randomized multi-restart policy can help the search to escape from the local basin and hence can help to prevent from being trapped in a local optimum solution.

Track level pruning using the ATHP concept is performed to enhance the computational efficiency. For pruning, we first define a set \mathcal{B}_{global} that is composed of feasible global hypothesis candidates and update it at every iteration. Using global hypotheses in the set \mathcal{B}_{global} obtained, the proposed algorithm then calculates the ATHP for each track, whose equation is given in (10), and prunes the ones smaller than the threshold value designed.

The algorithm of the proposed Tabu search-based best global hypothesis generation is summarised in Algorithm 1.

V. GIBBS-MHT ALGORITHM

The paper also proposes a computationally tractable approximation to address the original MWIS problem described in (12). Note that the approximation proposed is based on Gibbs sampling.

Gibbs sampling is a stochastic approach for Bayesian inference to approximate the posterior multivariate probability distribution in polynomial time [42], [43], especially for large scale problems. Gibbs sampler is a special case of the Metropolis-Hastings algorithm with an acceptance probability of one. The key idea of Gibbs sampling is that, given a joint distribution, it is simpler to recursively sample random samples from a conditional distribution than directly from the posterior joint distribution.

Our proposition is to apply Gibbs sampling to sample candidate solutions of x in (12) and evaluate their corresponding quality. Hence, the second MHT algorithm developed uses Gibbs sampling to generate an initial solution and recursively sample candidate solutions using the conditional distribution. As candidate solutions and their corresponding quality are recursively generated, sorting the solutions and finding the approximated solution, providing the best quality, are straightforward.

To apply Gibbs sampling, it is required to reformulate the original MWIS problem into a problem in which the key idea of Gibbs sampling can be applied. Therefore, we first reformulate the original max-sum problem given in (8) to a new max-product problem based on the property of track scores. From (3)-(5), we have

$$s_i = L(T_i^k) = \ln c_i^k \quad (13)$$

where c_i^k denotes likelihood ratio between the true target hypothesis probability and false alarm hypothesis probability. This implies that the max-summation problem (12) can be transformed to the max-product problem as

$$\begin{aligned} & \arg \max_X c(X), \\ \text{s.t. } & c(X) \triangleq \left(\prod_{i \in \mathcal{V}} \varphi_i(x_i) \right) \left(\prod_{(v_i, v_j) \in \mathcal{E}} \varphi_c(x_i, x_j) \right) \\ & X = (x_1, \dots, x_{|\mathcal{V}|}) \end{aligned} \quad (14)$$

with

$$\varphi_i(x_i) = \begin{cases} e^{s_i}, & x_i = 1 \\ 1, & x_i = 0 \end{cases} \quad (15)$$

Algorithm 1. Best global hypothesis generation by Tabu search

Input: A weighted graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{W})$ generated by all tracks at scan k , maximum allowable search depth L , maximum allowable iterations n_{max}

Output: The best global hypothesis defined by set \mathcal{F}^* , global hypothesis candidate set \mathcal{B}_{global}

```

1) begin
2)  $iter \leftarrow 0$  {set the initial iteration counter as zero}
3)  $\mathcal{F}^* \leftarrow \emptyset$  {set the initial solution as an empty set}
4)  $\mathcal{B}_{global} \leftarrow \emptyset$  (set the global hypothesis candidate set as an empty set)
5) while  $iter < n_{max}$  do
6)  $\mathcal{F}^* \leftarrow$  random initial possible solution generated from  $\mathcal{G}$ 
7) initialise the Tabu list
8)  $\mathcal{F}_{local} \leftarrow \mathcal{F}$  { $\mathcal{F}_{local}$  is the local optimal solution in each iteration during the loop}
9)  $L_{local} \leftarrow 0$   $L_{local}$  is the consecutive for local search
10) while  $L_{local} < L$  do
11)  $\bar{\mathcal{F}}_A, \bar{\mathcal{F}}_S \leftarrow$  construct neighbourhood solutions based on  $ADD(i)$  and  $SWAP(i, j)$ 
12)  $\mathcal{F} \leftarrow$  select the overall best solution  $\bar{\mathcal{F}}_A \cup \bar{\mathcal{F}}_S$  according to the quality function
13)  $L_{local} \leftarrow L_{local} + 1$ 
14)  $iter \leftarrow iter + 1$ 
15) update the Tabu list
16) if  $W(\mathcal{F}) > W(\mathcal{F}_{local})$  do
17)  $L_{local} \leftarrow 0$ 
18)  $\mathcal{F}_{local} \leftarrow \mathcal{F}$ 
19) end if
20) end while
21) if  $\mathcal{F}_{local} \notin \mathcal{B}_{global}$  do
22)  $\mathcal{B}_{global} \leftarrow \{\mathcal{B}_{global}, \mathcal{F}_{local}\}$  {update the global hypothesis candidate set}
23) end if
24) if  $W(\mathcal{F}_{local}) > W(\mathcal{F}^*)$  do
25)  $\mathcal{F}^* \leftarrow \mathcal{F}_{local}$  {update the best global hypothesis}
26) end if
27) end while
28) return  $\mathcal{F}^*, \mathcal{B}_{global}$ 

```

$$\varphi_c(x_i, x_j) = \begin{cases} 0, & x_i = x_j = 1 \\ 1, & \text{otherwise} \end{cases} \quad (16)$$

where $\varphi_i(x_i)$ is the node potential, guaranteeing a non-negative un-normalised probability to each possible value of the random variable $x_i \in \{0, 1\}$; $\varphi_c(x_i, x_j)$ is the edge potential, providing a non-negative un-normalised weight to all the combinations that (x_i, x_j) can take. The cost function $c(X)$ can be regarded as an un-normalised joint distribution of a pairwise undirected graphical model. The solution of (14) is a decoding problem, in which the goal is to find the most probable joint distribution.

The proposed max-product formulation enables utilisation of Gibbs sampling. The key idea proposed in this paper is to consider each solution as a random variable that satisfies a distribution $\pi(X)$. For the realisation, we first construct a finite-state Markov chain whose state space is the set Θ of all feasible solutions (i.e. independent sets of graph \mathcal{G}). Then, the Gibbs sampler is utilised to generate random variables x_i based on the proposed distribution proposal $\pi(X)$ so as to find the best candidate set. To ensure that the independent set solution with higher weights are more likely sampled, it is natural to construct the sampling proposal π proportional to its corresponding weight/track scores as

$$\pi(X) \propto \left(\prod_{i=1}^{|\mathcal{V}|} \varphi_i(x_i) \right) \left(\prod_{(v_i, v_j) \in \mathcal{E}} \varphi_c(x_i, x_j) \right) \quad (17)$$

Enumerating all possible independent sets is impossible for real applications, which means direct sampling from (17) is difficult. Hence, the concept of Gibbs sampling is adopted here, i.e. samples are generated by sweeping through all the posterior conditionals, rather than from the posterior distribution itself. Then, the one-step transition kernel from one possible solution $X = (x_1, \dots, x_{|\mathcal{V}|})$ to another possible solution $\bar{X} = (\bar{x}_1, \dots, \bar{x}_{|\mathcal{V}|})$ is obtained as

$$\pi(\bar{X} | X) = \prod_{m=1}^{|\mathcal{V}|} \pi_m(\bar{x}_m | \bar{x}_1, \dots, \bar{x}_{m-1}, x_{m+1}, \dots, x_{|\mathcal{V}|}) \quad (18)$$

More specifically, given solution x , the solution \bar{x} can be obtained by recursive sampling according to the following individual

conditional distributions

$$\begin{aligned}
\bar{x}_1 &\sim \pi_1(\bar{x}_1 | x_2, \dots, x_{|\mathcal{V}|}) \\
&\vdots \\
\bar{x}_m &\sim \pi_m(\bar{x}_m | \bar{x}_1, \dots, \bar{x}_{m-1}, x_{m+1}, \dots, x_{|\mathcal{V}|}) \\
&\vdots \\
\bar{x}_{|\mathcal{V}|} &\sim \pi_{|\mathcal{V}|}(\bar{x}_{|\mathcal{V}|} | \bar{x}_1, \dots, \bar{x}_{|\mathcal{V}|-1})
\end{aligned} \tag{19}$$

After constructing the Markov chain, it is necessary to prove that the generated Markov chain asymptotically converges to its invariant distribution and the convergence property of the proposed Gibbs sampling. These aspects are investigated in Theorem 1.

Theorem 1. *Given any initial feasible joint event, the distribution of Gibbs samples (18) asymptotically converges to the target distribution (17) with an exponential rate as*

$$|\pi^n(\bar{X} | X) - \pi(\bar{X})| \leq (1 - 2\beta)^{\lfloor n/2 \rfloor} \tag{20}$$

where $\pi^n(\bar{X} | X)$ denotes the n th power of transition kernel $\pi(\bar{X} | X)$, $\beta = \min_{X, \bar{X} \in \Theta} \pi^2(\bar{X} | X) \in (0, 0.5]$ the least likely two-step transition probability.

Proof: In general, the convergence of finite-state Markov chain is guaranteed by its irreducibility and regularity. The irreducibility of a Markov chain is quantified in terms of the possibility that one state can be transferred to another state within a finite number of steps. The regularity of a Markov chain can be checked by the positivity of the entries of some finite power of its transition matrix.

Let $\vartheta = \{\vartheta_i | \vartheta_i = \bar{x}_i, i \in \{1, \dots, m\}, \vartheta_j = x_j, j \in \{m+1, \dots, |\mathcal{V}|\}\}$. From (14) and (18), we have

$$\begin{aligned}
&\pi_m(\bar{x}_m | \bar{x}_1, \dots, \bar{x}_{m-1}, x_{m+1}, \dots, x_{|\mathcal{V}|}) \\
&= \frac{\pi(\bar{x}_1, \dots, \bar{x}_m, x_{m+1}, \dots, x_{|\mathcal{V}|})}{\pi(\bar{x}_1, \dots, \bar{x}_{m-1}, x_{m+1}, \dots, x_{|\mathcal{V}|})} \\
&\propto \pi(\bar{x}_1, \dots, \bar{x}_m, x_{m+1}, \dots, x_{|\mathcal{V}|}) \\
&\propto \left(\prod_{i=1}^{|\mathcal{V}|} \varphi_i(\vartheta_i) \right) \left(\prod_{(v_i, v_j) \in \mathcal{E}} \varphi_c(\vartheta_i, \vartheta_j) \right) \\
&= \left(\varphi_m(\vartheta_m) \prod_{(v_m, v_j) \in \mathcal{E}} \varphi_c(\vartheta_m, \vartheta_j) \right) \\
&\quad \times \left(\prod_{i \in \mathcal{V}, i \neq m} \varphi_i(\vartheta_i) \prod_{(v_i, v_j) \in \mathcal{E}, i \neq m} \varphi_c(\vartheta_i, \vartheta_j) \right) \\
&\propto \varphi_m(\vartheta_m) \prod_{(v_m, v_j) \in \mathcal{E}} \varphi_c(\vartheta_m, \vartheta_j)
\end{aligned} \tag{21}$$

‘Proportion to’ in inequality (21) highlights the dependence of individual conditional distribution on ϑ_m , while all other parts are formed as the normalisation constant.

Let 0_n denote the n dimensional zero vector. Then, it follows from (21) that

$$\begin{aligned}
\pi(0_{|\mathcal{V}|} | X) &\propto \prod_{m=1}^{|\mathcal{V}|} \varphi_m(0) > 0 \\
\pi(\bar{X} | 0_{|\mathcal{V}|}) &\propto \prod_{m=1}^{|\mathcal{V}|} \varphi_m(x_m) > 0
\end{aligned} \tag{22}$$

Then, the two-step transition kernel from any $X \in \Theta$ to any $\bar{X} \in \Theta$ satisfies

$$\begin{aligned}
\pi^2(\bar{X} | X) &= \sum_{\zeta \in \Theta} \pi(\bar{X} | \zeta) \pi(\zeta | X) \\
&> \pi(\bar{X} | 0_{|\mathcal{V}|}) \pi(0_{|\mathcal{V}|} | X) > 0
\end{aligned} \tag{23}$$

which means that the Markov chain $\{X^{(t)}\}_{t=1}^{\infty}$ generated by the Gibbs sampler is irreducible and recurrent, and therefore the Markov chain will asymptotically converge to its invariant distribution, e.g. the posterior of the joint event, by the ergodic theorem [44].

Since $\pi^2(\bar{X}|X) > 0$, applying Lemma 2, presented in Appendix A, to $\pi^2(\bar{X}|X)$ gives

$$\begin{aligned} \max_X \pi^{2n}(\bar{X}|X) - \min_X \pi^{2n}(\bar{X}|X) &\leq (1 - 2\beta)^n \\ \lim_{n \rightarrow \infty} \max_X \pi^{2n}(\bar{X}|X) &= \lim_{n \rightarrow \infty} \min_X \pi^{2n}(\bar{X}|X) \geq \beta > 0 \end{aligned} \quad (24)$$

Since Lemma 1, presented in Appendix A, states that $\max_X \pi^n(\bar{X}|X)$ is non-increasing and $\min_X \pi^n(\bar{X}|X)$ is non-decreasing in n , (24) can be reformulated as

$$\begin{aligned} \max_X \pi^n(\bar{X}|X) - \min_X \pi^n(\bar{X}|X) &\leq (1 - 2\beta)^{\lfloor n/2 \rfloor} \\ \lim_{n \rightarrow \infty} \max_X \pi^n(\bar{X}|X) &= \lim_{n \rightarrow \infty} \min_X \pi^n(\bar{X}|X) > 0 \end{aligned} \quad (25)$$

Due to the asymptotical convergence property of the proposed Markov chain, we have

$$\pi(\bar{X}) = \lim_{n \rightarrow \infty} \max_X \pi^n(\bar{X}|X) = \lim_{n \rightarrow \infty} \min_X \pi^n(\bar{X}|X) > 0 \quad (26)$$

Note that $\pi(\bar{X})$ lies between the minimum and maximum $\pi^n(\bar{X}|X)$ for any given state X , (20) can be directly ensured. QED. ■

Theorem 1 shows that, given any initial state $X^{(1)} \in \Theta$, the generated Markov chain $\{X^{(t)}\}_{t=1}^{\infty}$ exponentially converges to its stationary invariant distribution. Due to the convergence property, one can easily select a feasible solution as an initial state for Gibbs sampler. For example, one can simply choose the initial solution by random selection of compatible tracks, as presented in Sec. III.A.

Note that since Gibbs sampler is initialised with random values, samples generated at early iterations, known as the burn-in phase, usually cannot represent the target distribution and need to be discarded. With the aid of Gibbs sampler, one can also easily obtain the approximated marginal track hypothesis probability $p(x_i = 1)$ as

$$p(x_i = 1) = \frac{n_i}{n_{\max} - n_{\text{burn-in}}} \quad (27)$$

where n_{\max} denotes maximum allowable Gibbs samples, $n_{\text{burn-in}}$ the burn-in samples, n_i the number of samples that $x_i = 1$ occurs after the pre-convergent burn-in phase. The proposed Gibbs sampling-aided implementation algorithm also performs track-level pruning to improve its computational efficiency. Since the approximated hypothesis probability $p(x_i = 1)$ can be readily obtained from (27), track level pruning can be easily performed by thresholding $p(x_i = 1)$.

Algorithm 2 summarises the proposed Gibbs sampling-based implementation algorithm.

VI. NUMERICAL SIMULATIONS

This section validates the proposed algorithms through numerical simulations with comparisons to the mainstream MDA-MHT implementation algorithm.

A. Simulation Setup

The following six different challenging scenarios are conducted. Scenario 1: twelve constant moving targets with different birth and death time. Scenario 2: 20 close moving targets with three 3g coordinated turns. Scenario 3: 40 close moving targets with three 3g coordinated turns. Scenario 4: 20 random moving targets. Scenario 5: 40 random moving targets. Scenario 6: 60 random moving targets. The ground truths of these six scenarios are presented in Fig. 5. For all these scenarios, the measurements are generated with detection probability $P_D = 0.95$ and the clutter is assumed to be uniformly distributed in the surveillance region with its number at one scan being Poisson with spatial density $\lambda_F = 10^{-7}$.

The state variable of each target contains planar position and velocity and we apply simple constant velocity model with linear position measurement model for target state prediction and correction. More specifically, the state transition matrix and process noise matrix are given as

$$F \triangleq \mathbb{I}_{2 \times 2} \otimes \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}, \quad Q \triangleq \mathbb{I}_{2 \times 2} \otimes \sigma_v^2 \begin{bmatrix} T^3/3 & T^2/2 \\ T^2/2 & T \end{bmatrix} \quad (28)$$

where $\mathbb{I}_{2 \times 2}$ denotes the 2×2 identity matrix, \otimes the Kronecker product, $T = 1s$ the sampling period, and $\sigma_v = 22m/s^2$ is standard deviation of the process noise. The measurement matrix and measurement noise matrix are as follows

$$H \triangleq \mathbb{I}_{2 \times 2} \otimes \begin{bmatrix} 1 & 0 \end{bmatrix}, \quad R \triangleq \sigma_r^2 \mathbb{I}_{2 \times 2} \quad (29)$$

Algorithm 2. Best global hypothesis generation by Gibbs sampling

Input: A weighted graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{W})$ generated by all tracks at scan k , maximum allowable Gibbs samples n_{gibbs} , burn-in samples

$n_{burn-in}$

Output: The best global hypothesis defined by X_{best} , marginal distribution $p(x_i), \forall i \in [|\mathcal{V}|]$

```

1) begin
2)  $iter \leftarrow 0$  {set the initial iteration counter as zero}
3)  $x^{(1)} \leftarrow$  random initial possible solution generated from  $\mathcal{G}$ 
4)  $c_{best} \leftarrow -\text{inf}$  {set the initial best cost as -inf}
5) for  $i = 1 : |\mathcal{V}|$  do
6)    $n_i \leftarrow 0$  {set the initial event counter as zero}
7) end for
8) while  $iter < n_{max}$  do
9)   for  $i = 1 : |\mathcal{V}|$  do
10)    generating samples according to (19)
11)   end for
12)    $X^{(iter)} = (x_1^{(iter)}, \dots, x_{|\mathcal{V}|}^{(iter)})$  {one Gibbs sample}
13)   if  $iter > n_{burn-in}$  do
14)      $c(X^{(iter)}) \leftarrow$  calculate the current cost function
15)     if  $c(X^{(iter)}) > c_{best}$  do
16)        $c_{best} \leftarrow c(X^{(iter)})$ 
17)        $X_{best} \leftarrow X^{(iter)}$ 
18)     end if
19)     for  $i = 1 : |\mathcal{V}|$  do
20)       if  $x_i^{(iter)} = 1$  do
21)          $n_i \leftarrow n_i + 1$ 
22)       end if
23)     end for
24)   end if
25)    $iter \leftarrow iter + 1$ 
26) end while
27) for  $i = 1 : |\mathcal{V}|$  do
28)    $p(x_i = 1) \leftarrow n_i / (n_{max} - n_{burn-in})$  {Approximated
marginal track probability}
29) end for
30) return  $X_{best}, p(x_i)$ 

```

where $\sigma_r = 100m$ is the standard deviation of the measurement noise. For the above model, the standard Kalman filter for each target is applied as

$$\begin{aligned}
\hat{x}_{k|k-1} &= F\hat{x}_{k-1|k-1} \\
P_{k|k-1} &= FP_{k-1|k-1}F^T + Q \\
K &= P_{k|k-1}H^T(HP_{k|k-1}H^T + R)^{-1} \\
\hat{x}_{k|k} &= \hat{x}_{k|k-1} + K(z_k - H\hat{x}_{k|k-1}) \\
P_{k|k} &= P_{k|k-1} - KHP_{k|k-1}
\end{aligned} \tag{30}$$

For every new track, the state estimation error covariance is initialised using the one-point method [45] with maximum velocity bound $600m/s$ and the new target density is set as $\lambda_N = 10^{-11}$. Gating is performed with a threshold such that the gating probability is $P_G = 0.999$. For SPRT, the false track confirmation probability and the true track deletion probability are, respectively set as $\alpha = 10^{-6}$ and $\beta = 10^{-3}$. Since the detection probability and clutter spatial density may vary in practical implementations, we choose these two parameters as $P_D = 0.75$ and $\lambda_F = 10^{-8}$, which are not matched with the ones used in measurement generation. The scan depth of all MHTs is set as $p = 3$. In order to remove the hypothesis with low probability, the threshold of track level pruning is selected as $p_{th} = 10^{-6}$. Any hypothesis with probability that is below this threshold is removed. The design parameters for Tabu search-based approach are chosen as $L = 10$, $n_{max} = 100$, and parameters for Gibbs sampling-based implementation algorithm are set as $n_{max} = 2000$, $n_{burn-in} = 1000$. All experiments are performed on Matlab 2016b platform using an Intel Core i5-6500 CPU.

To demonstrate the effectiveness, the performance of the proposed algorithms are compared with that of the mainstream MDA formulation of MHT with Langrangian relaxation approach. As shown in [28], Langrangian relaxation implementation algorithm outperforms both m -best assignment and linear programming solutions. Therefore, it is reasonable to compare the proposed methods with the Langrangian approach.

For performance evaluation, the following commonly accepted metrics [46] are considered.

(1) Miscorrelation ratio of true tracks R_{MC} . The ratio of the number of false measurement-to-track associations to total true tracks.

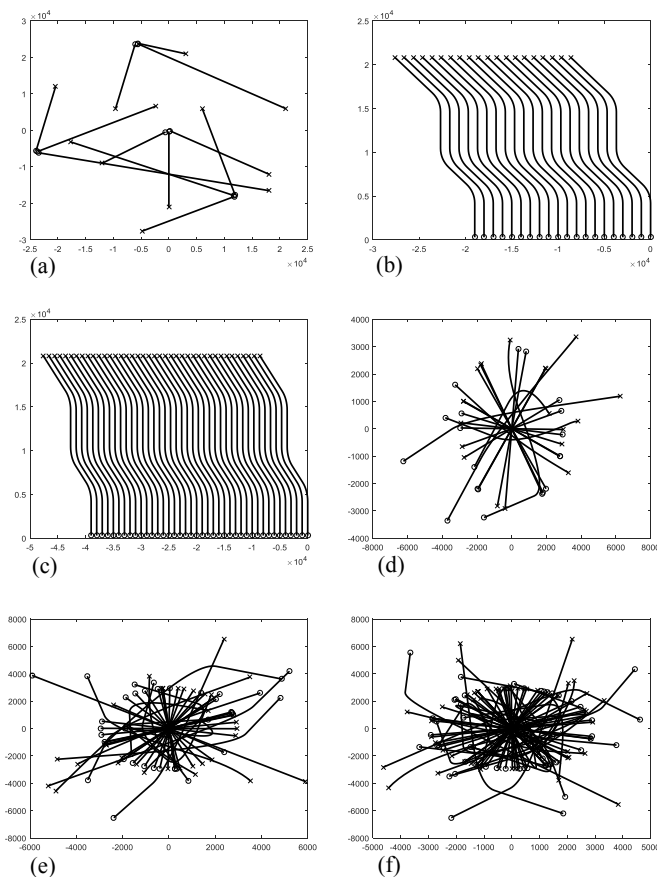


Fig. 5: Ground truth with start point o and death point x: (a) scenario 1; (b) scenario 2; (c) scenario 3; (d) scenario 4; (e) scenario 5; and (f) scenario 6.

$$d_p^c(X, Y) \triangleq \begin{cases} \left[\frac{1}{n} \left(\min_{\pi \in \Pi_n} \sum_{i=1}^m d^c(x_i, y_{\pi(i)})^p + c^p (n - m) \right) \right]^{1/p}, & m \leq n \\ d_p^c(Y, X), & m > n \end{cases} \quad (31)$$

(2) Correct correlation ratio of true tracks R_{CC} . The ratio of the number of correct measurement-to-track associations to total true tracks.

(3) Average number of true target IDs associate with a true track N_{ID} . For perfect matching in the ideal case, all measurements that allocated to a true track should match to the true target, i.e. $N_{ID} = 1$. Similarly, $N_{ID} > 1$ holds for imperfect matching.

(4) Average true track life T_{avg} . Average true track life measured in scan numbers.

(5) Running time t_e . The execution time for a sample run.

(6) Optimal sub-pattern assignment (OSPA) distance. Overall evaluation of cardinality and position estimation error for MTT. Let X and Y be the position estimation set and true target position set, respectively. The cardinality of these two sets are m and n , respectively. Denote Π_n as the set of all permutations on $\{1, 2, \dots, n\}$ for any positive integer n . $d^c(x_i, y_{\pi(i)}) = \min(d(x_i, y_{\pi(i)}), c)$ with $d(x_i, y_{\pi(i)})$ is the cut-off Euclidean distance between two vectors with $d(x_i, y_{\pi(i)})$ being the Euclidean distance. Then, for $c > 0$ and $1 \leq p < \infty$, the OSPA distance $d_p^c(X, Y)$ is defined as (31) [47]. The order parameter p in (31) determines the sensitivity of $d_p^c(X, Y)$ in penalizing estimation outliers, while the cut-off parameter c determines the relative weighting of the penalties allocated to cardinality and localization errors. In all simulations, these two parameters are set as $p = 2$, $c = 1000$.

B. Simulation Results

Table 1 summarises all the performance metrics obtained by these three different approaches of fifty Monte-Carlo runs, where the values are calculated as the mean value of all runs. The best one of each metric is highlighted in bold font. One sample run of OSPA distances are presented in Fig. 6.

TABLE I: Comparison results of Monte-Carlo simulations.

	Tracker	R_{MC}	R_{CC}	N_{ID}	T_{avg}	t_e	OSPA
Scenario 1	MDA-MHT	0.0671	0.9654	1.2265	44.2603	23.8042s	191.4025
	Tabu-MHT	0.0835	0.9636	1.3087	53.9755	20.2255s	179.9225
	Gibbs-MHT	0.0809	0.9660	1.1864	50.0071	16.5422s	181.6081
Scenario 2	MDA-MHT	0.0784	0.9538	1.0568	53.6069	7.7664s	153.5685
	Tabu-MHT	0.0882	0.9549	1.0604	63.4154	5.9839s	149.8836
	Gibbs-MHT	0.0854	0.9557	1.0470	53.6069	5.0561s	152.0427
Scenario 3	MDA-MHT	0.0758	0.9570	1.0414	51.2003	15.9088s	159.4479
	Tabu-MHT	0.0911	0.9540	1.0609	62.0841	13.1394s	154.6242
	Gibbs-MHT	0.0854	0.9591	1.0383	57.6434	11.3569s	155.4897
Scenario 4	MDA-MHT	0.2279	0.4102	2.3988	15.0391	11.3561s	597.8526
	Tabu-MHT	0.2386	0.4382	1.6182	24.9880	8.5660s	357.7429
	Gibbs-MHT	0.2479	0.4502	1.6652	26.4104	5.1800s	317.1836
Scenario 5	MDA-MHT	0.2142	0.3013	2.7500	11.6479	237.5427s	687.0442
	Tabu-MHT	0.2242	0.3313	2.8577	22.0750	37.1126s	362.0764
	Gibbs-MHT	0.2557	0.3160	2.7442	26.5581	19.6717s	323.1499
Scenario 6	MDA-MHT	0.2138	0.3042	2.9515	12.3592	1403.6657s	786.1761
	Tabu-MHT	0.2829	0.3091	2.8969	23.5577	124.7696s	393.8384
	Gibbs-MHT	0.2205	0.3375	2.7596	25.6316	43.7115s	362.9895

For scenarios 1 - 3, the performance of these three different MHT implementation algorithms is comparable. Although MDA-MHT has lowest miscorrelation rate of true tracks, it is evident that Tabu-MHT provides the best overall performances in terms of OSPA distance and has longer true track life, which is more desirable for long-time consecutive tracking. As a comparison, Gibbs-MHT is the most efficient one among these three approaches and shows better performance in ID switch as well as correct correlation ratio. The peak values in Fig. 6 (a) are caused by track initialisation of new birth targets. As for the latter three random moving cases, although MDA-MHT still has lowest miscorrelation rate of true tracks, its OSPA distance is almost twice than the proposed two algorithms, leading to the fact that the proposed two MHT implementation algorithms could have more robust estimation performance. Gibbs-MHT shows best performance in estimation accuracy and is also the most efficient one. Notably, for the cases with a large number of targets (scenario 6 for example), the execution time of Gibbs-MHT is only 43.7115s, whereas Tabu-MHT takes 124.7696s and MDA-MHT requires 1403.6657s.

Based on these results, it can be concluded that the proposed Gibbs-MHT and Tabu-MHT algorithms yield better overall performance, compared with the MDA-MHT algorithm. Note that it is clear that the number of the track hypotheses exponentially increases as the size and complexity of the problem grow. Therefore, the proposed algorithms are expected to provide better performance as the problem becomes more complex. The simulation results obtained confirm this expectation: as the complexity and the size of the scenario increases, the performance improvement becomes more significant. Furthermore, it is obvious from the simulation results that the proposed Gibbs-MHT drastically reduces the computational complexity without sacrificing the tracking performance.

VII. CONCLUSIONS

This paper developed two new algorithms for MHT implementation with the MWIS formulation: Tabu-MHT and Gibbs-MHT algorithms. The main purpose of these developments is to improve the computational efficiency while maintaining/enhancing the tracking performance. For the Tabu search based algorithm, random initial solutions are first generated to guarantee the diversity of the final solution. Then, two basic operators are used in each local search iteration to find better neighbourhood solutions. In the second proposed MHT algorithm, the MWIS problem is first reformulated to a max-product problem to successfully apply Gibbs sampling. Then, the Gibbs sampling based algorithm is designed to find the best approximated global hypothesis. Theoretical analysis shows that the proposed Markov chain generated by Gibbs sampler converges to its invariant distribution with exponential rate. Simulation results confirm that the proposed MHT implementation algorithms are highly efficient and provide better overall performance than the mainstream MDA-MHT in most scenarios. Especially, as can be deduced, it is clear from the simulation results that the performance improvement of the proposed algorithms becomes significant in large-scale problems.

APPENDIX A LEMMAS 1 AND 2

This appendix collects two key lemmas in [44] that are used in the proof of Theorem 1.

Lemma 1. *Let π be the transition matrix of a finite-state Markov chain and let π^n be the n th order transition probabilities. Then, for any state and $n \geq 1$, we have*

$$\begin{aligned} \max_{\theta} \pi^{n+1}(\chi|\theta) &\leq \max_{\zeta} \pi^n(\chi|\zeta) \\ \min_{\theta} \pi^{n+1}(\chi|\theta) &\geq \min_{\zeta} \pi^n(\chi|\zeta) \end{aligned} \tag{32}$$

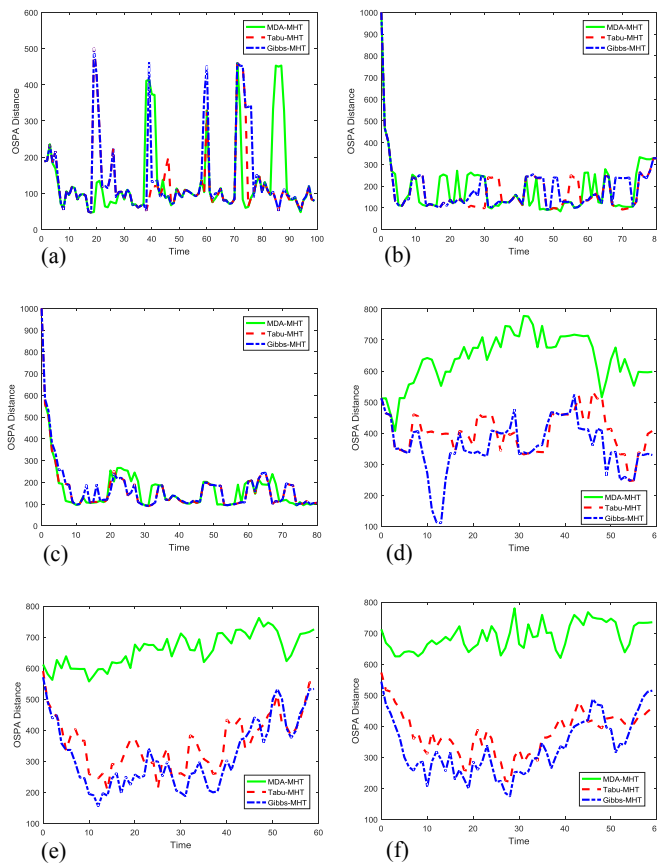


Fig. 6: OSPA distance with MDA-MHT (green), Tabu-MHT (red) and Gibbs-MHT (blue): (a) scenario 1; (b) scenario 2; (c) scenario 3; (d) scenario 4; (e) scenario 5; and (f) scenario 6.

Lemma 2. Let $\alpha = \min_{\chi, \theta} \pi(\chi | \theta)$ and the transition matrix π of a finite-state Markov chain satisfy $\pi > 0$. Then, for any state and $n \geq 1$, we have

$$\begin{aligned}
 & \max_{\theta} \pi^{n+1}(\chi | \theta) - \min_{\theta} \pi^{n+1}(\chi | \theta) \\
 & \leq \left[\max_{\zeta} \pi^n(\chi | \zeta) - \min_{\zeta} \pi^n(\chi | \zeta) \right] (1 - 2\alpha) \\
 & \max_{\zeta} \pi^n(\chi | \zeta) - \min_{\zeta} \pi^n(\chi | \zeta) \leq (1 - 2\alpha)^n \\
 & \lim_{n \rightarrow \infty} \max_{\zeta} \pi^n(\chi | \zeta) = \lim_{n \rightarrow \infty} \min_{\zeta} \pi^n(\chi | \zeta) \geq \alpha > 0
 \end{aligned} \tag{33}$$

REFERENCES

- [1] Y. Bar-Shalom, F. Daum, and J. Huang, "The probabilistic data association filter," *IEEE Control Systems Magazine*, vol. 29, no. 6, pp. 82–100, 2009.
- [2] D. E. Clark and B. Judith, "Bayesian multiple target tracking in forward scan sonar images using the PHD filter," *IEE Proceedings-Radar, Sonar and Navigation*, vol. 152, no. 2, pp. 327–334, 2005.
- [3] S. H. Rezatofighi, A. Milan, Z. Zhang, Q. Shi, A. Dick, and I. Reid, "Joint probabilistic data association revisited," *Proceedings of the IEEE International Conference on Computer Vision*, pp. 3047–3055, 2016.
- [4] A. Dehghan, S. M. Assari, and M. Shah, "GMMCP tracker: Globally optimal Generalized Maximum Multi Clique problem for multiple object tracking," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 4091–4099, 2015.
- [5] H. Wang, S. K. Nguang, and S. Member, "Multi-Target Video Tracking Based on Improved Data Association and Mixed Kalman / H_{∞} Filtering," *IEEE Sensors Journal*, vol. 16, no. 21, pp. 7693–7704, 2016.
- [6] W. Brendel, M. Amer, and S. Todorovic, "Multiobject Tracking as Maximum-Weight Independent Set," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Providence, RI, 2011, pp. 1273–1280.
- [7] X. Chen, L. An, and B. Bhanu, "Multitarget Tracking in Nonoverlapping Cameras Using a Reference Set," *IEEE Sensors Journal*, vol. 15, no. 5, pp. 2692–2704, 2015.
- [8] B. Gulmezoglu, M. B. Guldogan, and S. Gezici, "Multiperson tracking with a network of ultrawideband radar sensors based on gaussian mixture PHD filters," *IEEE Sensors Journal*, vol. 15, no. 4, pp. 2227–2237, 2015.
- [9] P. Zong and M. Barbaray, "Improved Multi-Bernoulli Filter for Extended Stealth Targets Tracking Based on Sub-Random Matrices," *IEEE Sensors Journal*, vol. 16, no. 5, pp. 1428–1447, 2016.
- [10] N. N. Kachouie and P. W. Fieguth, "Extended-Hungarian-JPDA: Exact single-frame stem cell tracking," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 11, pp. 2011–2019, 2007.
- [11] A. Dallil, M. Oussalah, and A. Ouldali, "Sensor Fusion and Target Tracking Using Evidential Data Association," *IEEE Sensors Journal*, vol. 13, no. 1, pp. 285–293, 2013.

- [12] A. G. Vicente, I. Bravo Munoz, P. Jimenez Molina, and J. L. Lazaro Galilea, "Embedded vision modules for tracking and counting people," *IEEE Transactions on Instrumentation and Measurement*, vol. 58, no. 9, pp. 3004–3011, 2009.
- [13] R. Singer, R. Sea, and K. Housewright, "Derivation and evaluation of improved tracking filter for use in dense multitarget environments," *IEEE Transactions on Information Theory*, vol. 20, no. 4, pp. 423–432, 1974.
- [14] S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*. Norwood, MA, USA: Artech House, 1999.
- [15] Y. Bar-Shalom, T. Kirubarajan, and X. Lin, "Probabilistic data association techniques for target tracking with applications to sonar, radar and EO sensors," *IEEE Aerospace and Electronic Systems Magazine*, vol. 20, no. 8 II, pp. 37–54, 2005.
- [16] T. E. Fortmann, Y. Bar-Shalom, and M. Scheffe, "Sonar tracking of multiple targets using joint probabilistic data association," *IEEE Engineering of Oceanic Engineering*, vol. 8, no. 3, pp. 173–184, 1983.
- [17] Y. Chen, Q. Zhao, Z. An, P. Lv, and L. Zhao, "Distributed Multi-Target Tracking Based on the K-MTSCF Algorithm in Camera Networks," *IEEE Sensors Journal*, vol. 16, no. 13, pp. 5481–5490, 2016.
- [18] R. J. Fitzgerald, "Development of Practical PDA Logic for Multitarget Tracking by Microprocessor," *American Control Conference, 1986*, pp. 889–898, 1986.
- [19] J. A. Roecker and G. L. Phillis, "Suboptimal Joint Probabilistic Data Association," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 29, no. 2, pp. 510–517, 1993.
- [20] J. A. Roecker, "A Class of Near Optimal JPDA Algorithms," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 30, no. 2, pp. 504–510, 1994.
- [21] D. Reid, "An algorithm for tracking multiple targets," *IEEE Transactions on Automatic Control*, vol. 24, no. 6, pp. 843–854, 1979.
- [22] T. Kurien, "Issues in the design of practical multitarget tracking algorithms," in *Multitarget-Multisensor Tracking: Advanced Applications*, 1990, pp. 43–83.
- [23] I. J. Cox and S. L. Hingorani, "Efficient implementation of Reid's multiple hypothesis tracking algorithm and its evaluation for the purpose of visual tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 2, pp. 138–150, 1996.
- [24] K. G. Murthy, "An algorithm for ranking all the assignments in order of increasing costs," *Operations Research*, vol. 16, no. 3, pp. 682–687, 1968.
- [25] S. Deb, M. Yeddapanudi, K. Pattipati, and Y. Bar-Shalom, "A generalized s-d assignment algorithm for multisensor-multitarget state estimation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 33, no. 2 PART 1, pp. 523–538, 1997.
- [26] A. B. Poore and A. J. Robertson III, "A New Lagrangian Relaxation Based Algorithm for a Class of Multidimensional Assignment Problems," *Computational Optimization and Applications*, vol. 8, no. 2, pp. 129–150, 1997.
- [27] R. T. Collins, "Multitarget data association with higher-order motion models," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1744–1751, 2012.
- [28] J. Areta, Y. Bar-Shalom, M. Levedahl, and K. R. Pattipati, "Hierarchical track association and fusion for a networked surveillance system," *Journal of Advances in Information Fusion*, vol. 1, no. 2, 2006.
- [29] D. Papageorgiou and S. Michael, "The maximum weight independent set problem for data association in multiple hypothesis tracking," *Optimization and Cooperative Control Strategies*, pp. 235–255, 2009.
- [30] C. Kim, F. Li, A. Ciptadi, and J. M. Rehg, "Multiple hypothesis tracking revisited," *Proceedings of the IEEE International Conference on Computer Vision*, vol. 11-18-December-2015, pp. 4696–4704, 2016.
- [31] P. R. Östergård, "A new algorithm for the maximum-weight clique problem," *Nordic Journal of Computing*, vol. 8, no. 4, pp. 424–436, 2001.
- [32] S. M. A. Nayeem and M. Pal, "Genetic algorithmic approach to find the maximum weight independent set of a graph," *Journal of Applied Mathematics and Computing*, vol. 25, no. 1, pp. 217–229, 2007.
- [33] B. Nogueira, R. G. Pinheiro, and A. Subramanian, "A hybrid iterated local search heuristic for the maximum weight independent set problem," *Optimization Letters*, pp. 1–17, 2017.
- [34] S. Blackman, "Multiple hypothesis tracking for multiple target tracking," *IEEE Aerospace and Electronic Systems Magazine*, vol. 19, no. 1, pp. 5–18, 2004.
- [35] S. S. Blackman, R. J. Dempster, and R. W. Reed, "Demonstration of multiple-hypothesis tracking (MHT) practical real-time implementation feasibility," in *International Symposium on Optical Science and Technology*, 2001, pp. 470–475.
- [36] P. Pardalos, L. Pitsoulis, T. Mavridou, and M. Resende, "Parallel search for combinatorial optimization: Genetic algorithms, simulated annealing, tabu search and grasp," *Parallel Algorithms for Irregularly Structured Problems*, pp. 317–331, 1995.
- [37] R. Battiti and G. Tecchiolli, "The continuous reactive tabu search: blending combinatorial optimization and stochastic search for global optimization," *Annals of Operations Research*, vol. 63, no. 2, pp. 151–188, 1996.
- [38] E. H. Aarts and J. K. Lenstra, *Local search in combinatorial optimization*. Princeton University Press, 1997.
- [39] F. W. Smith and J. A. Malin, "Models for radar scatterer density in terrain images," *IEEE transactions on aerospace and electronic systems*, no. 5, pp. 642–647, 1986.
- [40] S. S. Blackman, *Multiple-target tracking with radar applications*. Dedham, MA: Artech House, Inc., 1986.
- [41] F. Glover, "Tabu search: A tutorial," *Interfaces*, vol. 20, no. 4, pp. 74–94, 1990.
- [42] S. Geman and D. Geman, "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-6, no. 6, pp. 721–741, 1984.
- [43] G. Casella and E. I. George, "Explaining the Gibbs sampler," *The American Statistician*, vol. 46, no. 3, pp. 167–174, 1992.
- [44] R. G. Gallager, *Stochastic processes: theory for applications*. Cambridge, UK: Cambridge University Press, 2013.
- [45] Y. Bar-Shalom, X. R. Li, and K. Thiagalingam, *Estimation with applications to tracking and navigation: theory algorithms and software*. New York: John Wiley & Sons, 2004.
- [46] H. Leung, H. U. Zhijian, and M. Blanchette, "Evaluation of multiple radar target trackers in stressful environments," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 35, no. 2, pp. 663–674, 1999.
- [47] D. Schuhmacher, B. T. Vo, and B. N. Vo, "A consistent metric for performance evaluation of multi-object filters," *IEEE Transactions on Signal Processing*, vol. 56, no. 8 I, pp. 3447–3457, 2008.