

Article

Hierarchical Network-Based Tracklets Data Association for Multiple Extended Target Tracking with Intermittent Measurements

Kaiyi Jiang, Yiguo Li, Tianli Ma *  and Lin Li

School of Electronics and Information Engineering, Xi'an Technological University, Xi'an 710021, China; jiangkaiyi@st.xatu.edu.cn (K.J.); liyiguo@st.xatu.edu.cn (Y.L.); lilin0624@xatu.edu.cn (L.L.)

* Correspondence: matianli111@xatu.edu.cn

Abstract: The key issue of multiple extended target tracking is to differentiate the origins of the measurements. The association of measurements with the possible origins within the target's extent is difficult, especially for occlusions or detection blind zones, which cause intermittent measurements. To solve this problem, a hierarchical network-based tracklet data association algorithm (ET-HT) is proposed. At the low association level, a min-cost network flow model based on the divided measurement sets is built to extract the possible tracklets. At the high association level, these tracklets are further associated with the final trajectories. The association is formulated as an integral programming problem for finding the maximum a posterior probability in the network flow model based on the tracklets. Moreover, the state of the extended target is calculated using the in-coordinate interval Kalman smoother. Simulation and experimental results show the superiority of the proposed ET-HT algorithm over JPDA- and RFS-based methods when measurements are intermittently unavailable.

Keywords: multiple extended targets; data association; tracklets; min-cost network flow; intermittent measurements



Citation: Jiang, K.; Li, Y.; Ma, T.; Li, L. Hierarchical Network-Based Tracklets Data Association for Multiple Extended Target Tracking with Intermittent Measurements. *Sensors* **2023**, *23*, 6372. <https://doi.org/10.3390/s23146372>

Academic Editor: Gemine Vivone

Received: 22 May 2023

Revised: 23 June 2023

Accepted: 4 July 2023

Published: 13 July 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Conventional multiple target tracking (MTT) algorithms assume that objects can be represented as points and allow only a single measurement per sensor scan. However, modern high-resolution sensors have revealed the existence of targets that can generate multiple measurements per scan [1]. This challenges the suitability of the conventional point-target assumption. In such scenarios, multiple extended target tracking (METT) provides a more appropriate approach as it specifically addresses the tracking of targets that can produce multiple measurements per scan. The MTT with more measurements per target is called multiple extended target tracking [2,3].

Data association plays a vital role in multi-target tracking by differentiating between false alarms and actual targets. Incorrect data association can significantly impact the performance of tracking multiple targets. For point targets, several methods for data association have been proposed, which can be found in [4]. On the other hand, data association for extended targets involves the challenging task of matching measurements to specific targets, and this complexity increases with the number of targets and measurements. To address this, multiple data association techniques have been developed specifically for multiple extended target tracking scenarios. Vivone introduced a method that incorporates a detector and joint probability data association (JPDA) tracker specifically designed for METT [5]. Additionally, the multi-detection JPDA (MD-JPDA) algorithm was developed to address many-to-one associations in high-resolution radar sensors [6]. However, MD-JPDA suffers from high complexity due to exhaustive combinations of measurements. To mitigate this issue, MD-JIPDA, an extension of MD-JPDA, integrates the existence probability, reducing complexity [7,8]. Another approach by Yang focuses on calculating

marginal association probabilities for METT without relying on exhaustive hypotheses and partitions [9]. Furthermore, several algorithms based on the random finite set (RFS) framework have been proposed. These include the probability hypothesis density (PHD) filter [10,11], cardinalized PHD filter [12], generalized labeled multi-bernoulli (GLMB) filter [13], and Poisson multi-bernoulli mixture (PMBM) [14]. These RFS-based methods offer both optimal and suboptimal state estimates for multiple extended targets. However, it is important to note that tuning parameters for RFS-based methods may impact system reliability [15]. In addition, the graphical model formulation is applied to deal with the METT problem. Su presents the belief propagation algorithm to obtain estimators based on a simplified measurement set [16]. In order to remove the deviation in the extended state estimation, a METT algorithm based on Loopy Belief Propagation is presented in [17]. Previously developed METT algorithms rely on continuous measurement availability [18].

Intermittent measurements are a common occurrence in practical applications due to occlusions, detection blind zones (DBZs), and low frame rates caused by radar operation. To handle subsequent miss-detections, Mahdi proposed the Interacting Multiple Model PHD tracker [19], which does not output track labels. Another method, the multiple-model multiple-hypothesis PHD (MM-MH-PHD) filter, adopts a multiple-model approach to estimate motion states in blind zones [20]. In the case of tracking maneuvering targets in blind areas while considering DBZ masking, the MM-GLMB filter is utilized in [21], incorporating a minimum detection speed. These works struggled to handle the METT problem. Hence, Yang performed goal association by combining the PHD filter with the network flow, using a network flow algorithm to solve the minimum cost maximum flow problem, which is able to handle temporary disturbances with good robustness. In Ref. [22], this combination provides a more robust multi-target tracking algorithm that can handle temporary interference and generate longer and more consistent target trajectories.

The algorithms mentioned above require the known intermittent probability when dealing with the tracking process of intermittent measurements. Once the intermittent probability changes, it will affect the tracking results and cause a decrease in tracking accuracy.

In this paper, a hierarchical network framework of the METT with intermittent measurements is proposed. The core of the algorithm lies in the association between the reliable tracklets. A layered network is exploited with respect to the low-level min-cost flow constructed by the clustered measurements set and the high-level min-cost flow constructed by these tracklets. The reliable tracklets can be obtained by using the A* algorithm on the low network. Then, the minimum cost flow algorithm is employed to obtain the trajectories from the directed acyclic graph. The experimental results show that the hierarchical network-based tracklets algorithm (ET-HT) is effective. Our main contributions are as follows: (1) A hierarchical network data association framework for multiple extended target tracking with intermittent measurements is proposed. (2) The min-cost network-based tracklets are built, and the A* search algorithm is used to solve the integer programming problem. (3) The multiple pedestrians tracking scenario is adopted to test the performance of the proposed algorithm.

The structure of this paper is as follows: Section 2 provides the problem statement. Section 3 describes the proposed ET-HT algorithm including the formulation of the point-based and tracklet-based network, and the global optimal solution. In Section 4, the simulation and experimental results are presented to prove the effectiveness of the ET-HT algorithm. Section 5 is the discussion. Section 6 is the conclusion.

2. Problem Formulations

The METT problem aims to estimate states and parameters while dealing with measurements from multiple extended targets and the presence of clutter. At each time step k , the states of the multiple extended targets are denoted as $\mathbf{S}_k = \{\mathbf{s}_{1,k}, \mathbf{s}_{2,k}, \dots, \mathbf{s}_{N_T,k}\}$, N_T representing the number of extended targets. Each extended target state is defined as $\mathbf{s}_{i,k} = [\mathbf{x}_{i,k}, \mathbf{X}_{i,k}]^T$. The subset of the state $\mathbf{x}_{i,k}$ includes the centroid's location $\mathbf{m}_{i,k}$ and the velocity $\dot{\mathbf{m}}_{i,k}$ of the centroid of the target, $\mathbf{x}_{i,k} = [\mathbf{m}_{i,k}, \dot{\mathbf{m}}_{i,k}]^T$. $\mathbf{X}_{i,k} \in \mathbb{R}^d$ is the extension

state that describes the size and shape of the i -th target. At the time step k , the unordered set of measurements containing the clutter and target $\mathbf{Z}_k$ is received by the sensor. $\mathbf{Z}_k = \{\mathbf{z}_{1,k}, \mathbf{z}_{2,k}, \dots, \mathbf{z}_{M_k,k}\}$, where $M_k = \sum_i^{N_T} M_{i,k} + M_{C,k}$ represents the number of measurements. $M_{i,k}$ is defined as the total number of measurements that originates from the i -th extended target. The value of clutter $M_{C,k}$ is Poisson-distributed with mean λ_C .

In multiple extended target tracking, the measurements may intermittently be unavailable at any time. This is due to occlusions, detection blind zones, or even low frame rates caused by the work mode of the radar. To cater for the missing measurement, a binary variable representing the existence of the extended target is introduced, inspired by [23]. $\mathbf{r}_k = \{r_{1,k}, r_{2,k}, \dots, r_{N_T,k}\}$, with $r_{i,k} \in \{0, 1\}$, where the values of 0 and 1 correspond to the loss of the data and the measurement set of the i -th extended target is successfully received, respectively. The time-homogeneous binary Markov process has a transition probability matrix given by

$$\pi^+ = (\{r_{i,k+1} = n | r_{i,k} = m\})_{m,n \in \mathbb{S}} = \begin{bmatrix} 1 - q_{s,i} & q_{s,i} \\ p_{f,i} & 1 - p_{f,i} \end{bmatrix} \quad (1)$$

where $\mathbb{S} = \{0, 1\}$ is defined as the state space of the Markov process. The parameters $p_{f,i}$ and $q_{s,i}$ denote the failure rate and recovery rate, $0 < p_{f,i}, q_{s,i} < 1$, such that the Markov process $\{\mathbf{r}_k\}_{k \geq 0}$ is ergodic. Obviously, a smaller value of p_f and a larger value of q_s indicate a more reliable measurement received. Thus, the obtained measurement set of the i -th extended target at the time step k can be described as

$$\mathbf{Z}_{i,k} = r_{i,k} \{\mathbf{z}_{i,k}\}^{M_{i,k}} \quad (2)$$

The objective of METT is to accurately and robustly track multiple targets, even in challenging scenarios with occlusions, target interactions, intermittent measurements, and other complexities [24].

3. Hierarchical Network-Based Data Association for Multiple Extended Target Tracking

In this paper, emphasis is placed on the modeling of the data association for multiple extended targets (METs), utilizing a hierarchical dense neighborhood search approach. To achieve this, the density-based spatial clustering algorithm is employed to divide measurements at a specific time into clusters. Subsequently, a low-level association network is constructed based on the clustering results, enabling the calculation of tracklets. The association of clusters is formulated as a maximum a posteriori (MAP) problem, taking into account target initiation, termination, and false trajectories arising from clutter. At a higher level, the association network estimates trajectories by utilizing the tracklets derived from the low-level network. The paper presents the flowchart of the ET-HT filter in Figure 1, illustrating the stepwise process.

3.1. Pre-Processing

Since the set of measurements includes extended targets and clutter, it is necessary to extract the measured data of every extended object in one time step. Considering the property of the measurements of the extended target in that detections are spatially distributed around, the density-based spatial clustering algorithm is used to partition the set of measurements into multiple partitions. For the set of measurements $\mathbf{Z}_k$ which is received by the sensor at the time k , the set of subpartitions is

$$\Pi_k = (\pi_{1,k}, \pi_{2,k}, \dots, \pi_{N_\pi,k}) \quad (3)$$

where $\pi_{i,k}$ represents a single subpartition, and N_π is the number of all subpartitions. Note that the traditional partition algorithms include distance partition, subpartition, and K-means clustering. These methods require the specification of the amount of clusters.

The density-based spatial clustering can determine the number of clusters automatically in this paper, based on the intrinsic structure of the set of measurements. The details of the implementation can be found in [25].

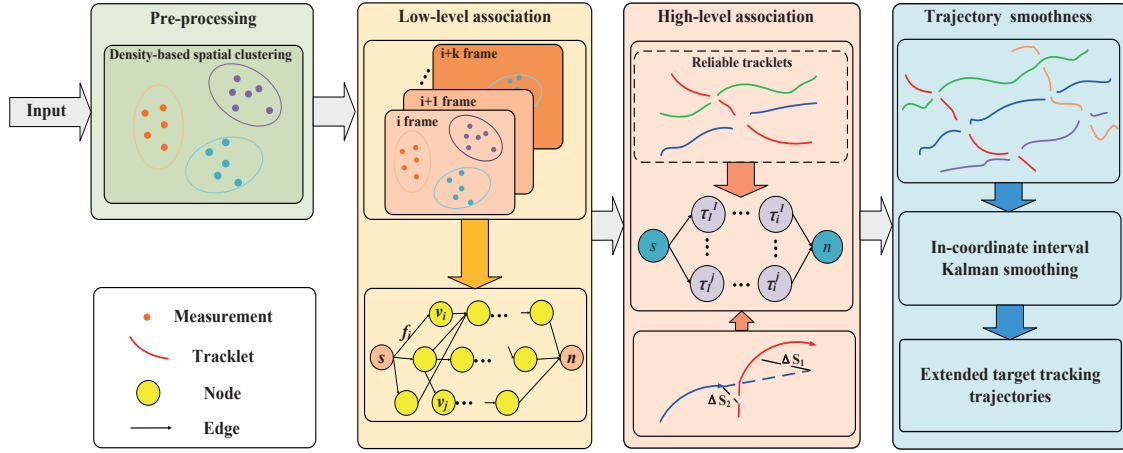


Figure 1. The flowchart of the proposed algorithm.

3.2. Hierarchical Association

a. Low-level association network

Due to the intermittent observations in the extended target tracking system, the extended target trajectories may be divided into several unconnected tracklets. Let $T_i = (\tau_i^1, \tau_i^2, \dots, \tau_i^n)$ represent the i -th extended target trajectory, where τ_i^j is defined as the j -th tracklet of the i -th target, and the number of trajectories is unknown. $\mathcal{T} = (T_1, T_2, \dots, T_m)$ is the set of the tracklets. In this section, the set of subpartitions is denoted as $\Pi_{1:k} = (\Pi_1, \Pi_2, \dots, \Pi_k)$. The low-level association defines $\Pi_{1:k}$ as the input, and uses the network flow method to generate the tracklets. The key here is to calculate a MAP estimate for $\mathcal{T}$ with a set of a cluster of measurements $\Pi_{1:k}$.

$$\mathcal{T}^* = \arg \max_{\mathcal{T}} \prod_{\mathcal{T}} P(\mathcal{T} | \Pi_{1:k}) \quad (4)$$

In this chapter, the $\mathcal{G}(V, E)$ is defined as a directed acyclic graph, where V represents the set of nodes and E represents the set of edges, with s denoting the nodes of the graph and n being the set of edges of the graph. This gives the set of graph nodes V and E , with v_i defined as a subpartition and each edge e_j in $\mathcal{G}(V, E)$ representing the motion between the subpartitions. In solving the data association process, the concept of network flow is used to represent f_{ij} as a directed flow variable from node v_i to node v_j , where f_{si} and f_{jn} denote the starting flow variable and the ending flow variable, respectively. Each flow in the graph is subject to the following constraints. Firstly, the sum of the flows arriving at node v_i is equal to the sum of the flows leaving this node at the same moment. For any tracklet τ_i^j :

$$f_{si} + \sum_{j:ji \in E} f_{ji} = \sum_{j:ij \in E} f_{ij} + f_{jn} \quad (5)$$

Secondly, the cost flow network must ensure that nothing other than a single extended object can be represented at one time. The upper bound of the sum of outgoing flows from node v_i is set to 1. For any node:

$$\forall v_i, v_j \sum f_{ij} < 1 \quad (6)$$

Considering that targets can appear or disappear anywhere in the cost-flow network, a source node and a sink node are introduced in [26], which are connected to the respective

nodes. The source and sink nodes are also subject to a constraint that enforces all the flows starting in s to end in n .

$$\sum_i f_{si} = \sum_i f_{in} \quad (7)$$

Through the network optimization process, Equation (4) is transformed into an integer programming (IP) problem, and the logarithm of the objective function is given by

$$\begin{aligned} \mathcal{T}^* &= \arg \max_{\mathcal{T}} \prod_{\mathcal{T}} P(\mathcal{T} | \Pi_{1:k}) \\ &= \arg \max_{\mathcal{T}} \sum_{\mathcal{T}} c_{si} f_{si} + c_{ij} f_{ij} + c_{jn} f_{jn} \\ &= \arg \max_{\mathcal{T}} \sum_i c_{si} f_{si} + \sum_{ij} c_{ij} f_{ij} + \sum_j c_{jn} f_{jn} \end{aligned} \quad (8)$$

where c_{si} is the flow cost from the source node to the subpartition π_i , c_{ij} is the flow cost from the subpartition π_i to the subpartition π_j , and c_{jn} is the flow cost from the subpartition π_j to the sink node. Hence, the IP problem can be described as

$$\begin{aligned} \min \quad & \sum_i c_{si} f_{si} + \sum_{ij} c_{ij} f_{ij} + \sum_j c_{jn} f_{jn} \\ \text{s.t.} \quad & \forall \pi_i, \pi_j \quad \sum f_{ij} < 1 \\ & \forall \pi_i, \pi_j \quad f_{ij} \geq 0 \end{aligned} \quad (9)$$

The cost can be defined as follows:

$$\begin{cases} c_{si} = -\log P_s(\pi_i) \\ c_{ij} = -\log P_l(\pi_j | \pi_i) \\ c_{jn} = -\log P_n(\pi_j) \end{cases} \quad (10)$$

In this paper, the data association problem is transformed into a MAP estimation for $\mathcal{T}$. The A* search algorithm is used to solve the problem, and the optimum trajectory $\mathcal{T}^*$ is obtained.

b. High-level association network

In this section, the tracklet-based network flow framework is represented as $\mathcal{G}^*(V^*, E^*)$ to obtain the trajectories of the multiple extended targets. It can be also formulated as a MAP problem.

$$\begin{aligned} \Psi^* &= \arg \max_{\Psi'} P(\Psi' | \mathcal{T}^*) \\ &= \arg \max_{\Psi'} P(\mathcal{T}^* | \Psi') P(\Psi') \\ &= \arg \max_{\Psi'} \prod_{\mathcal{T}_i^* \in \mathcal{T}^*} P(\mathcal{T}_i^* | \Psi') \prod_{\Psi_j \in \Psi'} P(\Psi_j) \end{aligned} \quad (11)$$

where Ψ_i denotes the merged trajectory, $\Psi_i = \{\mathcal{T}_{i,0}^*, \mathcal{T}_{i,1}^*, \dots, \mathcal{T}_{i,l}^*\}$. l denotes the number of tracklets in Ψ_i , and $\Psi' = \{\Psi_i\}$ is the merged trajectories set.

Note that the set of linear constraints is similar to those of Equations (5)–(7). The only difference is in the tracklet-based network flow model, where each node represents a tracklet extracted in the low-level association stage, which is a set of continuous measurements in a batch of time frames.

Assuming that the likelihoods of the input tracklets are conditionally independent given the merged trajectories set, and each merged trajectory is independent, the cost flow network of the tracklets is given by

$$P(\Psi_j) = P_{\Psi}(\mathcal{T}_{i,0}^*) \dots P_l(\mathcal{T}_{i,l-1}^* | \mathcal{T}_{i,l}^*) P_t(\mathcal{T}_{i,l}^*) \quad (12)$$

where $P(\Psi_j)$ represents a Markov Chain. $P_s(\mathcal{T}_{i,0}^*)$, $P_l(\mathcal{T}_{i,l-1}^*|\mathcal{T}_{i,l}^*)$ and $P_t(\mathcal{T}_{i,l}^*)$ represent the initial probability, transition probability, and terminal probability, respectively.

The link cost between two tracklets is defined as

$$P_l(\mathcal{T}_i^*|\mathcal{T}_j^*) = p_m(\mathcal{T}_i^*|\mathcal{T}_j^*)p_t(\mathcal{T}_i^*|\mathcal{T}_j^*) \quad (13)$$

where $p_m(\mathcal{T}_i^*|\mathcal{T}_j^*)$ and $p_t(\mathcal{T}_i^*|\mathcal{T}_j^*)$ represent the motion cost and time cost, respectively. The motion-associated probability is defined as

$$p_m(\mathcal{T}_i^*|\mathcal{T}_j^*) = -\ln(\mathcal{N}(\Delta p_1, \Sigma)\mathcal{N}(\Delta p_2, \Sigma)) \quad (14)$$

$$\Delta p_1 = p_i^{tail} + v_i^{tail} \Delta t - p_j^{head} \quad (15)$$

$$\Delta p_2 = p_j^{head} - v_j^{head} \Delta t - p_i^{tail} \quad (16)$$

where Δt denotes the frame gap between the last detection set of the tracklet $\mathcal{T}_j^*$ and the first detection set of the tracklet $\mathcal{T}_i^*$. p_i^{head} and p_i^{tail} are the center position of the first and last detection set of the tracklet $\mathcal{T}_i^*$, respectively. v_i^{head} and v_i^{tail} indicate the estimated speed of the tracklet $\mathcal{T}_i^*$ at the head and tail, respectively. The motion affinity between two tracklets is shown in Figure 2.

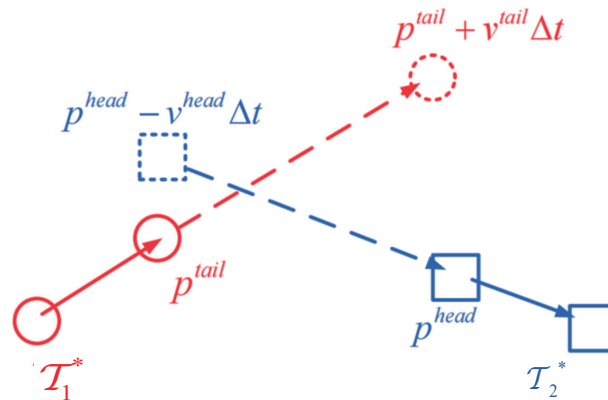


Figure 2. The motion affinity between two tracklets.

In this case, it is assumed that the error of the predicted location and the central location of the detection set Δp_1 and Δp_2 follows a Gaussian distribution. The smaller the error between the predicted location and the actual position of the target to be connected, the greater the motion similarity between the corresponding track slices will be. The temporal associated cost is defined as

$$p_t(\mathcal{T}_i^*|\mathcal{T}_j^*) = \begin{cases} 1, & \Delta t \in [1, \varsigma] \\ 0, & \text{otherwise} \end{cases} \quad (17)$$

where ς is the upper bound of the frame gap. The initialization probabilities and termination probabilities for each tracklet are set to be

$$P_\Psi(\mathcal{T}_i^*) = P_t(\mathcal{T}_j^*) = 1 \quad (18)$$

Similar to the low-level association process, the network flow model based on tracklets is established to solve the motion cost and time cost between the tracklets, and the optimal trajectories are obtained using the A* search algorithm.

3.3. Trajectory Smoothness

In this section, the in-coordinate interval Kalman smoothing algorithm is used to calculate the extended target's state[27]. The algorithm consists of forward filtering and backward smoothing. Assume that the interval between two frames is defined as $k(i) = k(0) + i \times T$, ($i = 0, 1, 2, \dots, M_l$), where T is the sampling time. At the time $k(i)$ ($i = 0, 1, 2, \dots, M_l - 1$), the Kalman filter only predicts the state of the extended target until a new measurement arrives at time $k(M_l)$ [28]. The time update step is as follows:

$$\mathbf{s}_{k(i)|k(i-1)} = \mathbf{F}_{k(i-1)}\mathbf{s}_{k(i-1)} \quad (19)$$

$$\mathbf{P}_{k(i)|k(i-1)} = \mathbf{F}_{k(i-1)}\mathbf{P}_{k(i-1)}\mathbf{F}_{k(i-1)}^T + \mathbf{w}_{k(i)} \quad (20)$$

where $\mathbf{s}_{k(i)}$ denotes the extended target state at the time $k(i)$, and $\mathbf{F}_{k(i-1)}$ is the transition matrix. Once the measurements arrive at time $k(M_l)$, the forward filtering step is

$$\mathbf{s}_{k(M_l)|k(M_l-1)} = \mathbf{F}_{k(M_l-1)}\mathbf{s}_{k(M_l-1)} \quad (21)$$

$$\mathbf{P}_{k(M_l)|k(M_l-1)} = \mathbf{F}_{k(M_l-1)}\mathbf{P}_{k(M_l-1)}\mathbf{F}_{k(M_l-1)}^T + \mathbf{w}_{k(M_l)} \quad (22)$$

$$\mathbf{K}_{k(M_l)} = \mathbf{P}_{k(M_l)|k(M_l-1)}\mathbf{H}_{k(M_l)}^T \left[\mathbf{H}_{k(M_l)}\mathbf{P}_{k(M_l)|k(M_l-1)}\mathbf{H}_{k(M_l)}^T + \mathbf{e}_{k(M_l)} \right]^{-1} \quad (23)$$

$$\mathbf{s}_{k(M_l)} = \mathbf{s}_{k(M_l)|k(M_l-1)} + \mathbf{K}_{k(M_l)}[\bar{\mathbf{z}}_{k(M_l)} - \mathbf{H}_{k(M_l)}\mathbf{s}_{k(M_l)|k(M_l-1)}] \quad (24)$$

$$\mathbf{P}_{k(M_l)} = \left[\mathbf{I} - \mathbf{K}_{k(M_l)}\mathbf{H}_{k(M_l)} \right] \mathbf{P}_{k(M_l)|k(M_l-1)} \quad (25)$$

where $\bar{\mathbf{z}}_{k(M_l)}$ represents the central location of the measurements of the extracted extended target at time $k(M_l)$. Here, the state and covariance matrix at each time are calculated via the backward recursive method starting from the last time k_{max} , given by

$$\mathbf{s}_{k|k} = \mathbf{F}_k\mathbf{s}_{k|k+1} \quad (26)$$

$$\mathbf{P}_{k|k} = \mathbf{F}_k\mathbf{P}_{k|k+1}\mathbf{F}_k^T \quad (27)$$

$$\mathbf{J}_k = \mathbf{P}_{k|k+1}\mathbf{F}_{k+1}^T\mathbf{P}_{k|k+1}^{-1} \quad (28)$$

$$\mathbf{s}_{k|k_{max}} = \mathbf{s}_{k|k} + \mathbf{J}_k(\mathbf{s}_{k+1|k_{max}} - \mathbf{s}_{k|k+1}) \quad (29)$$

$$\mathbf{P}_{k|k_{max}} = \mathbf{P}_{k|k} + \mathbf{J}_k(\mathbf{P}_{k+1|k_{max}} - \mathbf{P}_{k|k+1})\mathbf{J}_k^T \quad (30)$$

where $\mathbf{s}_{k|k_{max}}$ and $\mathbf{P}_{k|k_{max}}$ represent the smoothness state estimation and covariance matrix at time k , respectively. $\mathbf{J}_k$ is the smoothness gain matrix.

4. Numerical Simulation and Experiments

In this section, the feasibility of the proposed algorithm is verified by numerical simulation and experimental verification.

4.1. Case 1: Numerical Simulation

In the simulation, the validity of the ET-HT filter is tested. Consider a 2D surveillance region which is set as $[3000, 10,000] \text{ m} \times [0, 6000] \text{ m}$, with a clutter intensity of $c_k = 5 \times 10^{-17}$. The time step is set to 30, and the sampling period is defined as $T = 1 \text{ s}$. There are two

extended targets, and their initial positions are set to [6000, 4500] m and [3500, 3000] m, respectively. Their start velocities are set to [0, −150] m/s and [200, −20] m/s, respectively. The start time and terminal time in this system are [1, 5] s and [5, 30] s, respectively. The kinematic state is given by

$$\mathbf{s}_{i,k} = \mathbf{F}_{i,k}\mathbf{s}_{i,k-1} + \mathbf{w}_{i,k} \quad (31)$$

where $\mathbf{s}_{i,k} = [x_{i,k}, y_{i,k}, \dot{x}_{i,k}, \dot{y}_{i,k}]$ is the state variable. $x_{i,k}$ and $y_{i,k}$ represent the location of the targets. $\dot{x}_{i,k}$ and $\dot{y}_{i,k}$ represent the velocity of the targets. $\mathbf{F}_{i,k}$ is the kinematic state transition matrix of the i -th target. $\mathbf{w}_{i,k}$ represents the Gaussian process noise of the i -th target with zero mean and covariance $\mathbf{Q}_{i,k}$, $\mathbf{Q}_{i,k} = \text{diag}[100, 100, 0, 0]$.

The measurement model is defined as

$$\mathbf{z}_{i,k} = \mathbf{H}_k\mathbf{s}_{i,k} + \mathbf{e}_{i,k} \quad (32)$$

where $\mathbf{H}_k$ is the observation matrix, $\mathbf{z}_{i,k}$ is defined as the measurements generated by the i -th target at time k , and $\mathbf{e}_{i,k}$ denotes the Gaussian measurement noise of the i -th target with zero mean and covariance $\mathbf{R}_{i,k}$, $\mathbf{R}_{i,k} = \text{diag}[100, 100]$.

In the simulation, the clutter Poisson rate λ_c is set to 600, and then the clutter density $\lambda_c c_k$ is 3×10^{-15} . The failure and recovery rate are set to $p_f = 0.2$ and $q_s = 0.8$, respectively. The simulated target scenario with intermittent measurements is shown in Figure 3.

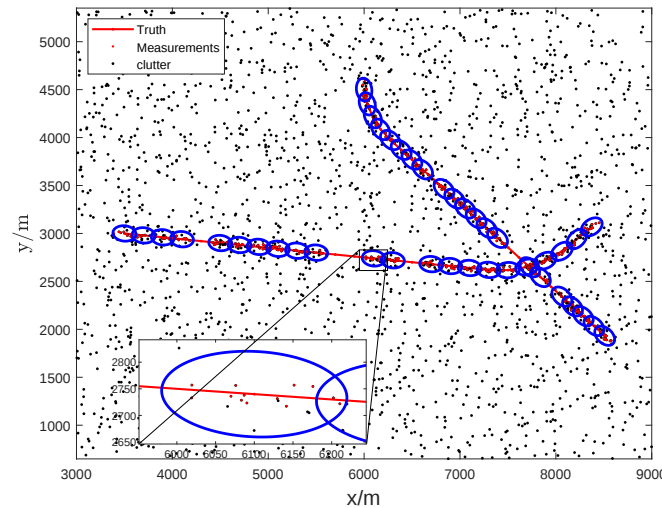


Figure 3. Target trajectories and measurements, the blue ellipses represent the shape of the extended targets.

In the experiment, the absolute mean number of targets estimation error (NTE) [29] and the optimal subpattern assignment (OSPA) distance [30] are taken into account as metrics to evaluate the performance of the proposed ET-HT filter against the ET-PHD, ET-JPDA, ET-PMBM, and ET-NFPHD filters.

The absolute mean number of targets estimation error is defined as

$$NTE(\mathbf{X}_k, \mathbf{Y}_k) = E\{|\mathbf{X}_k| - |\mathbf{Y}_k|\} \quad (33)$$

where $\mathbf{X}_k$ and $\mathbf{Y}_k$ are two finite subsets, and $|\mathbf{X}_k|$ and $|\mathbf{Y}_k|$ represent the potential of the two subsets, respectively. The OSPA distance between $\mathbf{X}_k$ and $\mathbf{Y}_k$ is defined as the distance between the position and the potential of the two sets.

$$OSPA(p, c, \mathbf{X}_k, \mathbf{Y}_k) = \left(\frac{1}{n} \left(\min_{\pi \in \Pi_n} \sum_{i=1}^n d^{(c)}(x_i, y_{\pi(i)})^p + c^p(m-n) \right) \right)^{1/p}, n \leq m \quad (34)$$

where c is the penalty cost for the cardinality mismatch, and p is the order of the OSPA metric, $1 < p < \infty$, $c > 0$. In the simulation, they are set to 10 and 2, respectively.

The OSPA distances and NTEs of the five filters are depicted in Figure 4. Comparatively, the ET-HT filter exhibits a lower OSPA distance and NTEs when compared to the other filters. In contrast, the performance of the ET-JPDA filter is influenced by clutter and necessitates prior knowledge regarding the number of targets. As the estimation error increases, the OSPA distance and NTE of the ET-JPDA filter also escalate. On the other hand, the ET-PHD filter leverages the first-order approximate moment of the multi-target density to convey valuable information about target potential. When the motion trajectories intersect, the targets are partitioned into one cell, which leads to an increase in the OSPA distance. Meanwhile, the ET-PHD and ET-PMBM filters require precise models of both targets and clutter. However, the lack of prior information and the interference of clutter result in a large discrepancy between the estimated and actual number of targets, which incurs increases in the NTE. The ET-NFPHD filter utilizes the PHD filter to process sensor measurements and extract the PHD component at different time steps. These PHD components are integrated into the network flow graph as nodes. By solving the minimum cost maximum flow problem among these nodes, the optimal data association path can be determined. As a result, the ET-NFPHD filter demonstrates a reduced OSPA distance and NTE compared to the ET-PHD filter. However, due to the direct modeling of the PHD filter's update data, the algorithm suffers from accumulated update errors, resulting in inferior tracking performance compared to the ET-HT filter proposed in this study.

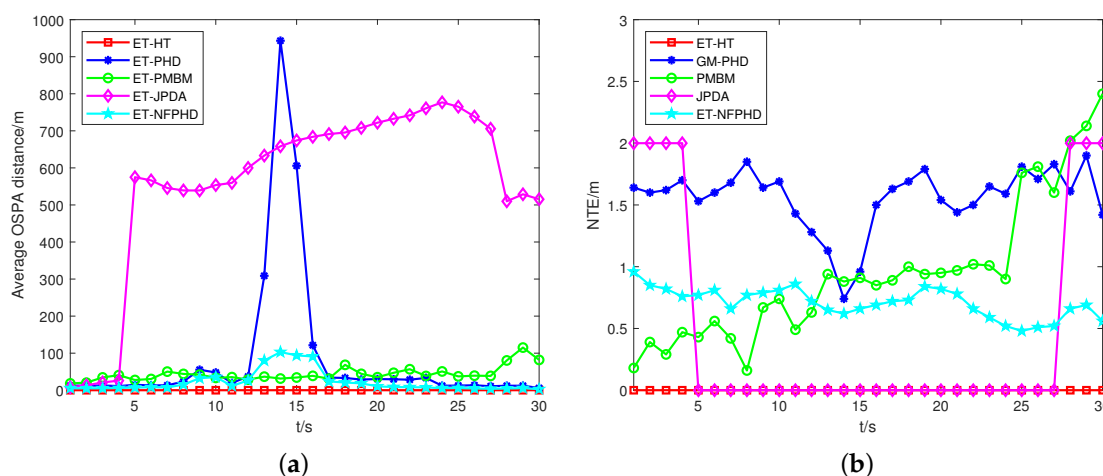


Figure 4. The average OSPA distances and NTEs of a single example run, with $q_s = 0.8$, $p_f = 0.2$, $\lambda_C = 300$. (a) OSPA, (b) NTE.

In order to illustrate the influence of clutter on the experiment, we designed simulation experiments with different clutter numbers.

Figure 5 presents the average OSPA distances and NTEs of the five filters at varying clutter rates. The ET-HT filter consistently demonstrates smaller average OSPA distance and NTEs compared to the ET-JPDA, ET-PMBM, ET-PHD and ET-NFPHD filters. To delve further into the influence of intermittent probability, the investigation extends to a different failure rate p_f and recovery rate q_s . The OSPA distances and NTEs of the five filters are illustrated in Figure 6a,b, respectively, considering diverse recovery rates in 100 Monte Carlo runs. Likewise, Figure 7a,b depicts the OSPA distances and NTEs of the filters for varying failure rates with 100 Monte Carlo runs.

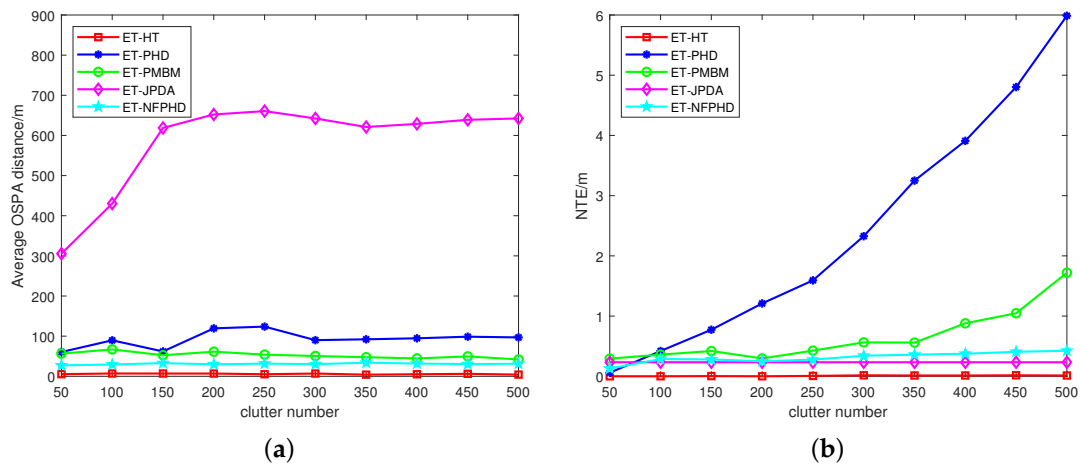


Figure 5. The average OSPA distances and NTEs for 100 Monte Carlo runs, with $q_s = 0.8$, $p_f = 0.2$. (a) OSPA, (b) NTE.

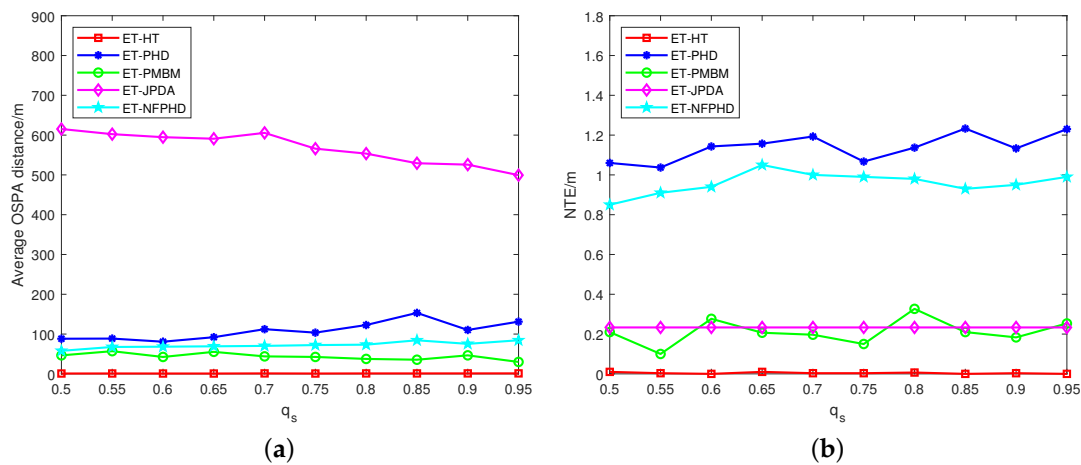


Figure 6. The average OSPA distances and NTEs of different recovery rates for 100 Monte Carlo runs, with $p_f = 0.2$, $\lambda_C = 600$. (a) OSPA, (b) NTE.

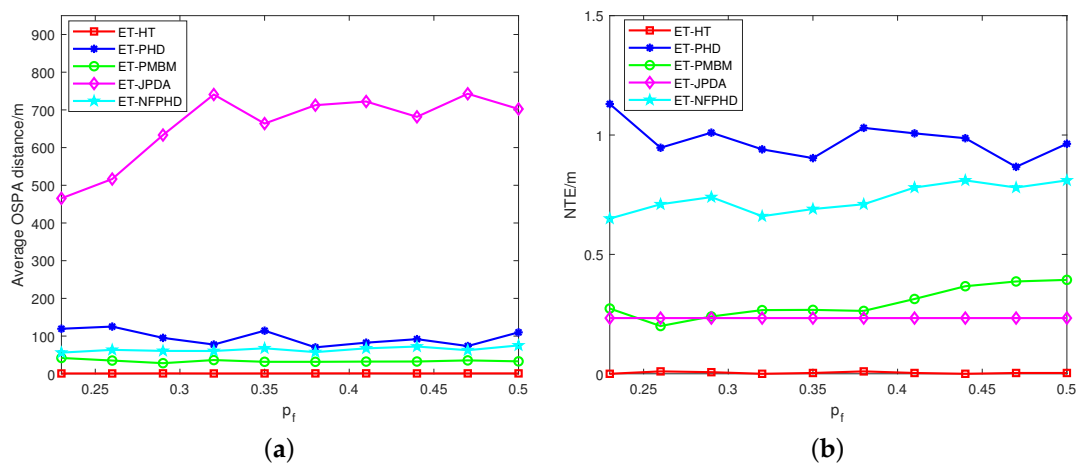


Figure 7. The average OSPA distances and NTEs of different failure rates for 100 Monte Carlo runs, with $q_s = 0.8$, $\lambda_C = 600$. (a) OSPA, (b) NTE.

The results from Figures 6 and 7 clearly demonstrate that the ET-HT filter exhibits superior tracking performance compared to the ET-PHD, ET-PMBM, ET-JPDA, and ET-NFPD filters irrespective of the transformation of the measurement recovery and loss rates. This is because the ET-HT filter associates the tracklets into long tracks through both low-level and high-level associations, thereby reducing the impact of intermittent measurements.

On the other hand, the experimental results of the ET-JPDA filter show that the OSPA distance of the ET-JPDA filter tends to decrease as the measurement recovery rate q_s increases, as shown in Figure 6a. However, if the measurement loss rate p_f increases, as shown in Figure 7a, it may result in a mismatch between the tracking gate and the measurement, thereby increasing the OSPA distance. Similarly, the results obtained using the ET-PHD filter shown in Figures 6b and 7b demonstrate that with an increasing recovery rate, some pieces of clutter are incorrectly identified as targets, causing errors in target estimation and an increase in the NTE. Conversely, a higher loss rate p_f leads to greater data loss, resulting in a decrease in the effective data volume at each time step. In the ET-PHD filters, significant data loss implies that fewer measurement data are associated with targets, leading to a decrease in the estimated number of targets and NTE. The presence of clutter impacts the partitioning of measurements, causing the OSPA distance and NTE of the ET-PMBM filter to exceed those of the ET-HT filter. The depicted results in Figures 6 and 7 demonstrate the consistent superiority of the ET-NFPD filter over the ET-PHD filter, irrespective of variations in intermittent probability. These findings serve as evidence supporting the notion that the integration of the flow network method effectively enhances the accuracy of the PHD filter.

4.2. Time Complexity

The time complexity of the ET-HT filter consists of two main components, namely the adaptive spectral clustering algorithm and the worst-case scenario of the search subgraph. The time complexity of the former is $O(n \times (N_{tc} - n_{tc})) + O(n_{tc}^3) + O(K_{tc}N_{tc}I_{tc}) \cdot O(n_{tc} \times (N_{tc} - n_{tc}))$, where K_{tc} is the number of nodes in the undirected graph, n_{tc} is the sample points, N_{tc} is the number of trajectories, and I_{tc} is the iteration number of the K-means algorithm. The worst time complexity of the searching sub-graph is $O(n_{\max}^2)$ [31].

The time complexity of the ET-JPDA filter is $O(N_{tc} \times n_{tc}^2)$, where N_{tc} is the initial number of targets, and n_{tc} is the sampling time. The time complexity of the ET-PHD filter is $O(N_{tc} \times n_{tc}^2) + O(n_{tc} \times M_{tc} \times N_{tc}^2 \times d_{tc}^2)$, where the parameter M_{tc} is generally greater than 1000, and d_{tc} is the dimension of the state vector. The time complexity of the ET-PMBM filter is $O(N_{tc}^3 \times H_{tc}^3)$, where H_{tc} represents the number of hypothetical combinations. The time complexity of the ET-NFPD filter is $O(N_{tc} \times n_{tc}^2) + O(K_{tc}N_{tc}I_{tc})$.

The proposed algorithm and existing algorithms are coded with MatlabR2019b, and the experiments are run on a computer with an Intel Core i74770k CPU at 3.5 GHz. The running times of the five filters for 100 Monte Carlo runs are shown in Table 1.

Table 1. Average running times of five filters.

Filter	ET-HT	ET-PHD	ET-PMBM	ET-JPDA	ET-NFPD
Time (t/s)	10.1369	2232.9	96.3917	1.1254	411.9493

Table 1 demonstrates that the ET-HT filter exhibits a reduced running time in comparison to both the ET-PHD filter and the ET-PMBM filter. This indicates that the ET-HT filter possesses a lower time complexity, leading to significant enhancements in computational efficiency. Although the running time of the ET-HT filter exceeds that of the ET-JPDA filter, its accuracy in estimating the state and number of targets is significantly superior. Hence, it is crucial to prioritize accuracy in tracking rather than solely emphasizing the attainment of low time complexity.

4.3. Case 2: Real Data Verification

To evaluate the applicability of the proposed filter, case 2 incorporates real scenarios. These selected scenes encompass more randomness and a higher number of moving targets, reflecting real situations. The positional data of each pedestrian were obtained with a pedestrian detector with discriminatively trained part-based models [32]. Firstly, a classifier that uses a decision rule learned from the data was constructed to determine whether the image window contains pedestrians. Next, the image was divided into multiple sub windows and each sub window was described using features that were invariant to lighting effects. Then, each sub window was provided to the classifier, which determined whether the sub-window contained pedestrians, and their position coordinates were extracted.

The input for the ET-HT filter consisted of a collection of measurement points gathered from all frames during the motion sequence. In situations where pedestrians were obscured by a streetlight, the measurement points could not be extracted, thereby simulating intermittent observation scenarios.

The ET-HT filter framework was applied to a scenario from the PETS 2009S0CC dataset. In this dataset, we focused on a specific scene featuring six pedestrians walking. Their paths intersected at specific frames (295, 305, 345, 365, and 367). Additionally, in frames 319, 339, and 351, the pedestrians were obstructed by a street light, as illustrated in Figure 8a. Figure 8b shows the extracted measurements. Figure 9 shows the tracking results of the ET-HT filter. Following a similar approach as in case 1, the PETS2009S0CC dataset was utilized to evaluate the performance of the ET-HT filter, ET-JPDA filter, ET-PHD filter, ET-PMBM, and ET-NFPHD filters of a single example run. The OSPA distances and NTEs were computed for each filter with the time step $t = 10$ s. The obtained results are illustrated in Figure 10.

As illustrated in Figure 9, the ET-HT filter achieves a better performance when dealing with real data with intermittent measurements and is able to accurately estimate the trajectories of the targets. The results in Figure 10a,b show that the tracking error of the ET-JPDA filter increases with the increase in clutter number, while the ET-PHD, ET-PMBM, and ET-NFPHD filters have large errors in dealing with real scenes. On the contrary, the ET-HT filter has better robustness and accuracy for tracking targets in real scenes.

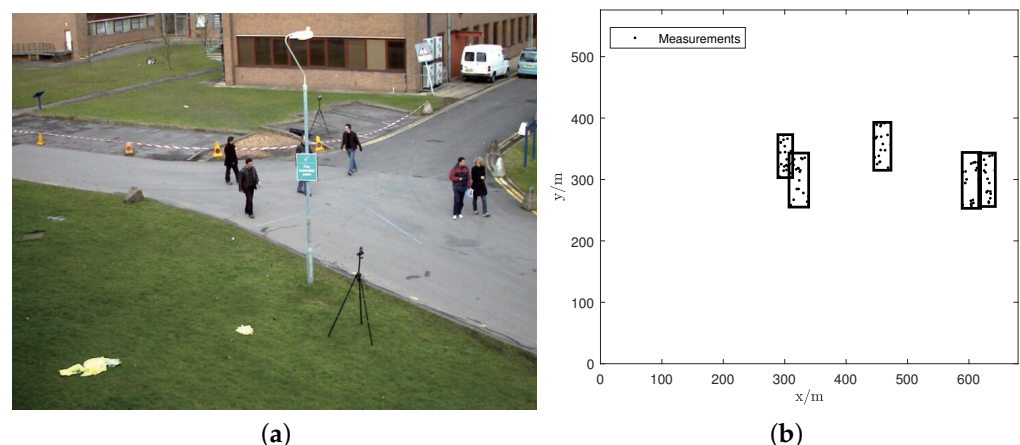


Figure 8. An example of the PETS2009S0CC dataset. The extracted measurement points are fed into the ET-HT filter. Subsequently, following processing with the density-based spatial clustering algorithm, these measurement points can be considered as trajectories depicting pedestrian motion. (a) An example of the PETS 2009S0CC dataset where a pedestrian is obstructed by a street light, and (b) the extracted measurements.

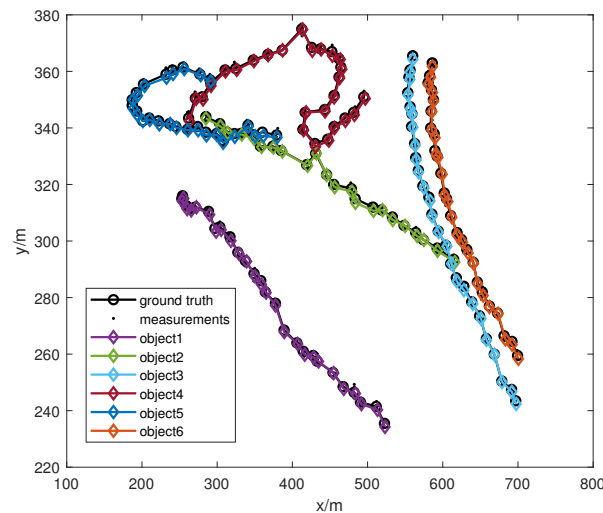


Figure 9. The tracking results for the ET-HT filter

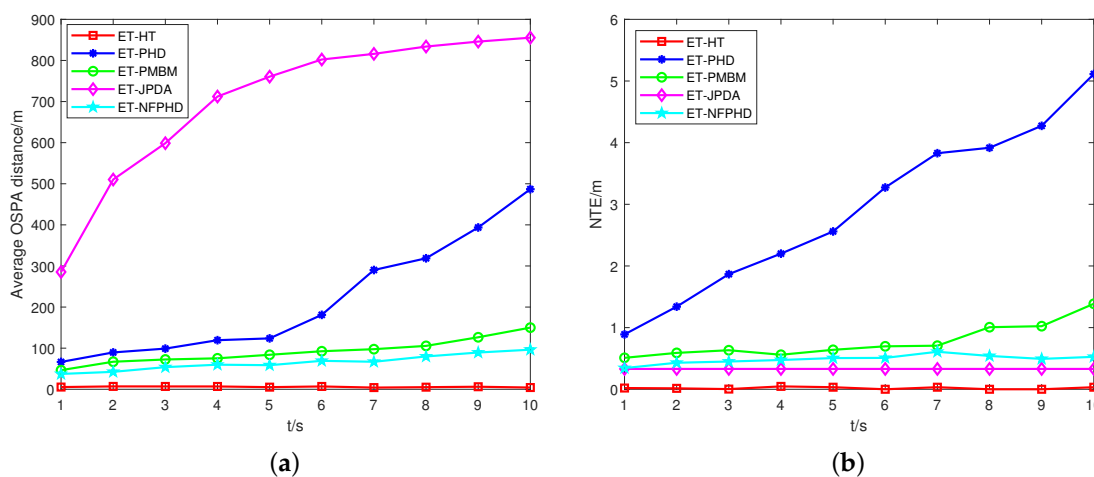


Figure 10. The average OSPA distances and NTEs of a single example run, with $t = 10$. (a) OSPA, (b) NTE.

The validation of the ET-HT filter on real data demonstrates its advantages in addressing target tracking problems with intermittent measurements. These findings have significant implications for practical applications that demand precise target tracking, such as intelligent traffic systems, video surveillance, and object recognition.

5. Discussion

The primary focus of this paper is to tackle the challenge of tracking multiple extended targets when dealing with intermittent measurements. The proposed filter in this study, ET-HT, is designed to extract all potential short tracks and employs a network flow model to compute the trajectories of these targets. Comparative experiments have been conducted to validate the algorithm's performance, demonstrating its superior robustness and accuracy in handling intermittent measurements. However, it is important to note that this paper exclusively addresses the tracking method for extended targets and does not explore the issue of tracking group targets with intermittent measurements. Additionally, the implications of target separation and merging phenomena on tracking performance in the presence of intermittent measurements have not been thoroughly investigated in

this paper. Further exploration of this aspect is warranted and represents a crucial area for future research.

6. Conclusions

In this paper, we present a hierarchical network-based tracklet data association framework for tracking multiple extended targets with intermittent measurements. At low association levels, all possible tracklets are extracted and utilized to construct a high-level network flow model. The trajectories of the multiple extended targets can then be obtained using the A* search algorithm. The effectiveness of the proposed ET-HT filter is demonstrated through simulations and experimental results, particularly in challenging scenarios involving clutter, newborn targets, and occlusion. Moving forward, our future research aims to expand the algorithm's capabilities by considering observation delay, integrating it into multi-object trackers, and exploring the potential of incorporating data from multiple sensors to enhance estimation accuracy and overall tracking performance.

Author Contributions: K.J. and T.M. conceived of and designed the algorithm; K.J. performed the experiments, analyzed the data, and drafted the paper; L.L. and Y.L. revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Natural Science Foundation of China, grant number 82101969, in part by the Regional Innovation Capability Guidance Program Project of Shaan xi Province, grant number 2022QFY01-16, in part by the Key Research and Development Project of Shaan xi Province, grant number 2022GY-242, and in part by the Science and Technology Plan Project of Xian Yang City, grant number 2021ZDYF-SF-0020.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data based on School of Electronics and Information Engineering is not universal, so it is not disclosed. Interested readers can contact us directly by email.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Mahler, R.P.S. *Advances in Statistical Multisource-Multitarget Information Fusion*; Artech House: Norwood, MA, USA, 2014; Volume 10, pp. 142–149.
2. Lan, J.; Li, X.R. Tracking of extended object or target group using random matrix: New model and approach. *IEEE Trans. Aerosp. Electron. Syst.* **2016**, *52*, 2973–2989. [\[CrossRef\]](#)
3. Stateczny, A.; Kazimierski, W. Multisensor Tracking of Marine Targets—Decentralized Fusion of Kalman and Neural Filters. *Int. J. Electron. Telecommun.* **2011**, *57*, 65–70. [\[CrossRef\]](#)
4. Morelande, M.R.; Kreucher, C.M.; Kastella, K. A Bayesian Approach to Multiple Target Detection and Tracking. *IEEE Trans. Signal Process.* **2007**, *55*, 1589–1604. [\[CrossRef\]](#)
5. Vivone, G.; Braca, P. Joint Probabilistic Data Association Tracker for Extended Target Tracking Applied to X-Band Marine Radar Data. *IEEE J. Ocean. Eng.* **2016**, *41*, 1007–1019. [\[CrossRef\]](#)
6. Habtemariam, B.; Tharmarasa, R.; Thayaparan, T. A Multiple-Detection Joint Probabilistic Data Association Filter. *IEEE J. Sel. Top. Signal Process.* **2013**, *7*, 461–471. [\[CrossRef\]](#)
7. Schuster, M.; Reuter, J.; Wanielik, G. Multi- detection joint integrated probabilistic data association using random matrices with applications to radar-based multi object tracking. *IEEE J. Adv. Inf. Fusion* **2013**, *12*, 175–188.
8. Xie, Y.; Huang, Y.; Song, T.L. Iterative joint integrated probabilistic data association filter for multiple-detection multiple-target tracking. *Digit. Signal Process.* **2018**, *72*, 232–243. [\[CrossRef\]](#)
9. Yang, S.; Wolf, L.M.; Baum, M. Marginal Association Probabilities for Multiple Extended Objects without Enumeration of Measurement Partitions. In Proceedings of the 2020 IEEE 23rd International Conference on Information Fusion, Rustenburg, South Africa, 6–9 July 2020; pp. 1–8.
10. Li, P.; Ge, H.W.; Yang, J.L.; Wang, W. Modified Gaussian inverse Wishart PHD filter for tracking multiple non-ellipsoidal extended targets. *Signal Process.* **2018**, *150*, 191–203. [\[CrossRef\]](#)
11. Peng, M.S.; Linares, R.; Bageshwar, V.L. Extended Target Tracking and Shape Estimation via Random Finite Sets. In Proceedings of the 2019 American Control Conference, Milwaukee, WI, USA, 10–12 July 2019; Volume 72, pp. 232–243.
12. Zhibin, Z.; Liping, S.; Xuan, C. Labeled box-particle CPHD filter for multiple extended targets tracking. *J. Syst. Eng. Electron.* **2019**, *30*, 57–67.

13. Liang, Z.; Liu, F.; Li, L. Improved generalized labeled multi-Bernoulli filter for non-ellipsoidal extended targets or group targets tracking based on random sub-matrices. *Digit. Signal Process.* **2020**, *99*, 102669–102682. [\[CrossRef\]](#)
14. Granstrom, K.; Svensson, L.; Reuter, S.; Xia, Y.; Fatemi, M. Likelihood-Based Data Association for Extended Object Tracking Using Sampling Methods. *IEEE Trans. Intell. Veh.* **2018**, *3*, 30–45. [\[CrossRef\]](#)
15. Zhang, H.; Gao, L.; Xu, M.; Wang, Y. An improved probability hypothesis density filter for multi-target tracking. *Optik* **2019**, *182*, 23–31. [\[CrossRef\]](#)
16. Su, Z.Z.; Ji, H.B.; Zhang, Y. Data association for extended target tracking by BP. *IET Radar Sonar Navig.* **2018**, *12*, 1484–1492. [\[CrossRef\]](#)
17. Su, Z.Z.; Ji, H.B.; Zhang, Y. Loopy belief propagation based data association for extended target tracking. *Chin. J. Aeronaut.* **2020**, *33*, 2212–2223. [\[CrossRef\]](#)
18. Liu, W.; Zhu, S.; Wen, C.; Yu, Y. Structure modeling and estimation of multiple resolvable group targets via graph theory and multi-Bernoulli filter. *Automatica* **2018**, *89*, 274–289. [\[CrossRef\]](#)
19. Yazdian, D.M.; Azimifar, Z. An improvement on GM-PHD filter for target tracking in presence of subsequent miss-detection. In Proceedings of the 2015 23rd Iranian Conference on Electrical Engineering, Tehran, Iran, 10–14 May 2015.
20. Ma, T.L.; Wang, X.M.; Li, T. Multiple-model multiple hypothesis probability hypothesis density filter with blind zone. *Int. J. Ind. Syst. Eng.* **2017**, *27*, 180–195. [\[CrossRef\]](#)
21. Wu, W.; Sun, H.; Cai, Y.; Jiang, S.; Xiong, J. Tracking multiple maneuvering targets hidden in the DBZ based on the MM-GLMB filter. *IEEE Trans. Signal Process.* **2020**, *68*, 2912–2924. [\[CrossRef\]](#)
22. Yang, S.; Teich, F.; Baum, M. Network flow labeling for extended target tracking PHD filters. *IEEE Trans. Ind. Inform.* **2019**, *15*, 4164–4171. [\[CrossRef\]](#)
23. You, K.Y.; Fu, M.Y.; Xie, L.H. Mean Square Stability for Kalman Filtering with Markovian Packet Losses. *Automatica* **2011**, *47*, 1247–1257. [\[CrossRef\]](#)
24. Liu, Z.; Ji, L.; Yang, F.; Qu, X.; Yang, Z.; Qin, D. Cubature information Gaussian mixture probability hypothesis density approach for multi extended target tracking. *IEEE Access* **2019**, *7*, 103678–103692. [\[CrossRef\]](#)
25. Guo, P.; Liu, Z.; Wang, J.J. HRRP multi-target recognition in a beam using prior-independent DBSCAN clustering algorithm. *IET Radar Sonar Navig.* **2019**, *13*, 1366–1372. [\[CrossRef\]](#)
26. Xi, Z.; Liu, H. Multiple Object Tracking Using the Shortest Path Faster Association Algorithm. *Sci. World J.* **2014**, *2014*, 481719–481730. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Zhao, L.; He, Z. An in-coordinate interval adaptive Kalman filtering algorithm for INS/GPS/SMNS. In Proceedings of the IEEE 10th International Conference on Industrial Informatics, Beijing, China, 25–27 July 2012; pp. 41–44.
28. Aravkin, A.; Burke, J.; Ljung, L.; Lozano, A.; Pillonetto, G. Generalized Kalman smoothing: Modeling and algorithms. *Automatica* **2017**, *86*, 63–86. [\[CrossRef\]](#)
29. Wang, Y.; Meng, H.; Liu, Y.; Wang, X. Collaborative penalized Gaussian mixture PHD tracker for close target tracking. *Signal Process.* **2014**, *102*, 1–15. [\[CrossRef\]](#)
30. Schuhmacher, D.; Vo, B.T.; Vo, B.N. A consistent metric for performance evaluation of multi-object filters. *IEEE Trans. Signal Process.* **2008**, *56*, 3447–3457. [\[CrossRef\]](#)
31. Ma, T.; Gao, S.; Chen, C.; Song, X. Multitarget Tracking Algorithm Based on Adaptive Network Graph Segmentation in the Presence of Measurement Origin Uncertainty. *Sensors* **2018**, *18*, 3791. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Felzenszwalb, P.F.; Girshick, R.B.; McAllester, D.; Ramanan, D. Object detection with discriminatively trained part-based models. *IEEE Trans. Pattern Anal. Mach. Intell.* **2009**, *32*, 1627–1645. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.