Decentralized and Asynchronous Multi-Agent Active Search and Tracking when Targets Outnumber Agents

Arundhati Banerjee ARUNDHAT@CS.CMU.EDU
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213, USA

Jeff Schneider SCHNEIDE@CS.CMU.EDU
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213, USA

Abstract

Multi-agent multi-target tracking has a wide range of applications, including wildlife patrolling, security surveillance or environment monitoring. Such algorithms often assume that agents are pre-assigned to monitor disjoint partitions of the environment, reducing the burden of exploration. This limits applicability when there are fewer agents than targets, since agents are unable to continuously follow the targets in their fields of view. Multi-agent tracking algorithms additionally assume a central controller and synchronous inter-agent communication. Instead, we focus on the setting of decentralized multi-agent, multi-target, simultaneous active search-*and*-tracking with asynchronous inter-agent communication. Our proposed algorithm DecSTER uses a sequential monte carlo implementation of the probability hypothesis density filter for posterior inference combined with Thompson sampling for decentralized multi-agent decision making. We compare different action selection policies, focusing on scenarios where targets outnumber agents. In simulation, DecSTER outperforms baselines in terms of the Optimal Sub-Pattern Assignment (OSPA) metric for different numbers of targets and varying teamsizes.

Keywords: Probability Hypothesis Density, Thompson sampling, Decentralized Multi-Agent

1. Introduction

Searching for targets, detecting objects of interest (OOIs), localizing and following them are tasks integral to several robotics applications. When targets are in motion, agents (robots) face a non-stationary environment. Therefore, agents tracking an unknown number of moving targets should trade-off between exploring the possibly unobserved parts of the environment and exploiting its own posterior estimates to localize the previously detected targets at the current timestep. Unfortunately, prior work in multi-target tracking (MTT) has often assumed that the environment is known and exploration is not of primary concern (Robin and Lacroix, 2016). Moreover, with multiple agents, existing MTT methods simplify the explore-exploit dilemma by separating search and tracking into sequential tasks where each agent is assigned to track a target as soon as it is found (Papaioannou et al., 2020). Another approach is to assign sub-teams for executing these tasks separately

(Chen and Dames, 2022). Further, the majority of these multi-agent MTT (MAMTT) algorithms require either a central controller to coordinate joint tracking actions, or they depend on synchronized inter-agent communication for distributed inference and decision making. Unfortunately, such conditions may not be feasible in practice.

Contributions. In this work, we aim to develop a more practical approach for MAMTT. We assume that agents are outnumbered by targets, so that the multi-agent team is unable to continuously cover all targets in their fields of view. We propose a decentralized and asynchronous multi-agent algorithm, called DecSTER (*Dec*entralized Multi-agent Active *S*earch and *T*racking without continuous cov*er*age) for simultaneous multi-target active search and tracking without continuously following the targets. In simulation, we compare a number of common decision making objectives from the tracking literature after adapting them to our simultaneous active search-and-tracking setting. Our results show that DecSTER outperforms competitive baselines with different teamsizes and different target distributions in terms of the OSPA tracking performance.

2. Related Works

Target detection and tracking are both widely studied problems, typically considered as distinct tasks in various applications like search and rescue (Murphy, 2004), security surveillance (Doitsidis et al., 2012), computer games (Oskam et al., 2009), etc. We refer to Robin and Lacroix (2016) for a detailed survey of the many different approaches and taxonomy used in robotics and related fields for such scenarios.

In multi-target settings, the target state is modeled with approaches like Multiple Hypothesis Tracker (MHT) (Blackman, 2004), Joint Probabilistic Data Association (JPDA) (Fortmann et al., 1983) and Probability Hypothesis Density (PHD) filter (Mahler, 2003), all of which differ in how they perform data association (Stone et al., 2013). The PHD filter is particularly suited when unique identities for each target are not required, for example, in search and rescue tasks, where agents should detect and localize *all* survivors. In this work, we build on the Sequential Monte Carlo (SMC) PHD filter in Ristic et al. (2010).

Prior work in MAMTT algorithms primarily considers centralized or distributed settings, the latter still necessitating synchronized communication among subgroups of agents at each time step. Coupled with a PHD filter, some of the common action selection methods previously proposed for tracking include mutual information and expected count based objectives (Dames et al., 2017), Renyi divergence maximization (Papaioannou et al., 2020) and Lloyd's algorithm for Voronoi-cell based control (Dames, 2020; Chen and Dames, 2020). In contrast with these deterministic objectives, Xin and Dames (2022) demonstrates the superior performance of stochastic optimization methods like Particle Swarm Optimization (PSO) and Simulated Annealing (SA) for better coverage and localization in such settings. Unfortunately, none of these prior approaches are applied in the decentralized and asynchronous multi-agent setup when agents are unable to support continuous target coverage.

3. Problem Setup

Consider a team of J UAVs tasked with search and tracking of an unknown number of moving targets $\{1, \ldots, k\}$ in a 2-dimensional (2D) region $\mathcal{G}$ of length n_l and width n_w

(Fig. 3a). The agents sense noisy location 2D coordinates of possible targets in their current field of view (FOV) following decentralized decision making, without pre-coordination. Each agent's FOV includes a contiguous rectangular region of the search space, and agents may choose to observe a wider (smaller) area at a greater (lower) vertical height but with more (less) observation noise. We therefore consider a hierarchical region sensing action space for each agent. The targets can move in any direction in the search space at different speeds. Agents can communicate asynchronously with their teammates (Fig. 3b). Over time T, agents observe different parts of the search space to detect and track all targets.

Target and Measurement Representations. The state of each target is denoted by $\mathbf{x} = \left[l_x, l_y, v_x, v_y\right]^{\mathrm{T}}$, where 2D coordinates $(l_x, l_y) \in [0, n_w] \times [0, n_l]$ and velocities $v_x, v_y \in [-v_{\max}, v_{\max}]$. Since both the cardinality and true locations of targets are unknown, we follow the Random Finite Set (RFS) representation for the multi-target state space $\mathcal{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_{|\mathcal{X}|}\}$, where $|\mathcal{X}|$ follows a Poisson distribution and the set elements are sampled i.i.d from a uniform distribution (Mahler, 2014). Following prior work, we use the Particle Hypothesis Density (PHD) filter (Mahler, 2003) to maintain a belief over the RFS $\mathcal{X}$.

The PHD $\nu(x)$ is the first statistical moment of a distribution over RFSs. Here it is a density over the target state space so that for any region E, the expected cardinality of the target RFS in that region is $\int_E \nu(x)dx$. The PHD filter tracks the evolving target density over the search space using target motion models and agent observations.

Sensing model. An agent with pose $\mathbf{q} = \left[q_x, q_y\right]^{\mathrm{T}}$ executes a sensing action $\mathbf{a_q}$, receiving a measurement set $\mathcal{Z} = \{\mathbf{z}_1, \ldots, \mathbf{z}_m\}$. Any target $\mathbf{x}$ within the agent's FOV may generate a measurement $\mathbf{z}$, with a probability of detection $p_d(\mathbf{x}|\mathbf{q})$. In this work, we assume a constant $p_d(\cdot)$ when the target $\mathbf{x}$ is within the FOV at $\mathbf{q}$, and 0 otherwise. The agent follows a linear sensing model with additive i.i.d white noise: $\mathbf{z} = h(\mathbf{x}) + \boldsymbol{\omega}$, where $h(\mathbf{x}) = \left[l_x, l_y\right]^{\mathrm{T}}$ and $\boldsymbol{\omega} \sim \mathcal{N}(\mathbf{0}, \sigma_h^2\mathbf{I})$. Additionally, $\mathcal{Z}$ includes i.i.d false positives with clutter rate $\lambda_{\mathbf{q}}$. The measurements $\mathcal{Z}$ are modeled as a (Poisson) RFS, as are clutter $\kappa(z)$ (false positives) and target births $b(x)$. We refer to Mahler (2014); Ristic et al. (2010) for a detailed understanding of the PHD filter prediction and update steps using the target motion model, the agent's sensor parameters and gathered measurements following a sensing action (Appendix A). As Mahler (2003) explains, using the first order moment to approximate the multi-target belief and deriving recursive PHD update equations to approximate the evolving posterior is justifiable when both sensor covariances and false alarm densities are small, so that (the distribution of) observations from true targets are centered around target states with negligible spread and there is lower noise due to false alarms.

Since targets are in continuous motion, our agents must be able to deal with the uncertainty arising from observation noise as well as due to asynchronous communication of time-dependant observations in their posterior PHD updates. In order to enable time-ordered assimilation of received observations by all agents, we assume that any agent j communicates the tuple $(t, \mathbf{a}_t^{\mathbf{q}_j}, \mathcal{Z}_t^j)$ where $\mathbf{a}_t^{\mathbf{q}_j}$ and $\mathcal{Z}_t^j$ are respectively the agent's sensing action and measurement set at time t.

4. Our approach

We now describe our algorithm DecSTER for multi-agent active search and tracking without continuous coverage. Agent j at time t has a history of available actions and observations

Algorithm 1 TS-PHD-I

Input: PHD $\nu = \{(w_1, \mathbf{x}_1), \ldots, (w_\rho, \mathbf{x}_\rho)\}$.
Sample $\tilde{\rho}$ particles $\{\mathbf{x}_i\}_{i=1}^{\tilde{\rho}}$ from ν, proportional to the weights $\{w_1, \ldots, w_\rho\}$.
Cluster the $\tilde{\rho}$ particles using k-means with $\tilde{n} = \sum_{i=1}^{\tilde{\rho}} w_i$ centroids to get the TS.

Algorithm 2 TS-PHD-II

Input: PHD ν=$\{(w_1, \mathbf{x}_1), \ldots, (w_\rho, \mathbf{x}_\rho)\}$. $\hat{n}_\mathcal{G}$=$\sum_{i=1}^{\rho} w_i$. $\hat{\mathcal{X}}$=$\{\hat{\mathbf{x}}_1, \ldots, \hat{\mathbf{x}}_{\hat{n}_\mathcal{G}}\}$ from ν.
Sample $\tilde{n} \sim \text{Poisson}(\hat{n}_\mathcal{G})$. Sample uniformly random target locations $\tilde{\mathcal{X}}_R$.
Sample $\tilde{n}$ target locations ($\tilde{\mathcal{X}}$) from $\hat{\mathcal{X}} \cup \tilde{\mathcal{X}}_R$ as the TS.

$\mathcal{D}_t^j = \{(t', \mathbf{a}_{t'}^{\mathbf{q}'_j}, \mathcal{Z}_{t'}^{j'})\}_{t'<t, j' \in \{1,\ldots,J\}}$. Using $\mathcal{D}_t^j$, it computes the PHD ν_t^j over the target RFS. In our SMC-PHD implementation, $\nu_t^j = \{(w_{t,1}^j, \mathbf{x}_{t,1}^j), \ldots, (w_{t,\rho}^j, \mathbf{x}_{t,\rho}^j)\}$ where $\mathbf{x}_{t,1}^j, \ldots, \mathbf{x}_{t,\rho}^j$ are the ρ particles with weights $w_{t,1}^j, \ldots, w_{t,\rho}^j$. The SMC-PHD filter propagation steps follow from Ristic et al. (2010). In our decentralized setup, each agent maintains its own posterior PHD ν_t^j. Next, we will describe the decision making step executed by agent j at time t.

Thompson sampling for decision making. Prior work in multi-agent active search with static targets has demonstrated the effectiveness of Thompson sampling (TS) as a decentralized decision making algorithm (Ghods et al., 2021; Bakshi et al., 2023). Here we propose a TS strategy so agents can trade-off exploratory sensing actions that might discover undetected targets, with exploitative sensing actions that help localize and track the previously detected dynamic targets in our simultaneous search-and-tracking setting.

To the best of our knowledge, prior work has not studied the problem of TS in a continuous (not discretized) search space with a PHD posterior. This is challenging because the PHD is not a distribution and does not include second order uncertainty information, whereas TS is typically applied in the Bayesian setting with the samples drawn from a posterior distribution for which both first and second order moment estimates are available (Russo et al., 2018). Prior work in Zhou et al. (2022) has proposed particle Thompson sampling (PTS) and regenerative PTS (RPTS) algorithms for particle filters where particles are sampled proportional to their weights. Therefore, we adopt a similar principle in our first proposed TS strategy for the SMC-PHD posterior, called TS-PHD-I (Algorithm 1). But this method tends to sample more particles from the regions in the PHD where the agent already estimates targets might be present. The samples drawn are thus more likely to be biased against regions of the target state space where the agent might be less certain about its observations owing to false positives or missed detections. Furthermore, this method does a poor job of modeling the uncertainty about the number of true targets.

To address the drawbacks of Algorithm 1, we propose a second approach TS-PHD-II in Algorithm 2. Recall that the expected cardinality of the target RFS $\mathcal{X}$ over a region $E \subseteq \mathcal{G}$ is given by $\hat{n} = \int_E \nu(\mathbf{x}) d\mathbf{x}$. In case of the SMC-PHD representation, $\hat{n} = \sum_i w_i, \forall \mathbf{x}_i \in E$ (Ristic et al., 2010) i.e. the sum of particle weights of the SMC-PHD in the region E is the expected cardinality of $\mathcal{X}$ in that region. Further, Mahler (2003) shows that the PHD is the best Poisson approximation of the multitarget posterior in terms of KL divergence. We therefore draw a sample $\tilde{n}$ of the cardinality of the target RFS from a Poisson distribution with mean $\hat{n} = \sum_i w_i$ (i.e. $\tilde{n} \sim \text{Poisson}(\hat{n})$). Then we sample $\tilde{n}$ locations of the possible targets $\tilde{\mathcal{X}} = \{\tilde{\mathbf{x}}_1, \dots, \tilde{\mathbf{x}}_{\tilde{n}}\}$ by drawing from a mixture of already estimated target locations in the PHD and some locations drawn uniformly at random over the search space. These $\tilde{n}$ particles $\tilde{\mathcal{X}} = \{\tilde{\mathbf{x}}_1, \dots, \tilde{\mathbf{x}}_{\tilde{n}}\}$ form our TS.

Objective. The Optimal Sub-Pattern Assignment (OSPA) metric is typically used in the MTT literature for evaluating the tracking performance of an algorithm and is defined as the error between two sets, taking into account both the cardinality error and the distance error between the set elements (Appendix D). Given a true target set $\mathcal{X}$ and an estimated set $\mathcal{Y}$ of possible target locations, our goal is to minimize $\text{OSPA}(\mathcal{X}, \mathcal{Y})$. Since the ground truth $\mathcal{X}$ is unknown, each agent j instead draws a TS $\tilde{\mathcal{X}}_t^j$ from the predicted PHD $\bar{\nu}_{t+1}^j$. Assuming observations are generated by $\tilde{\mathcal{X}}_t^j$ for any action $\mathbf{a}$ and $\mathcal{Y}_t^j$ is the estimated target set following the PHD filter update, agent j then selects:

$$\mathbf{a}_t^j = \arg\min_{\mathbf{a}} \mathbb{E}_{\mathcal{Y}_t^j | \tilde{\mathcal{X}}_t^j, \mathbf{a}} [\text{OSPA}(\tilde{\mathcal{X}}_t^j, \mathcal{Y}_t^j)] \tag{1}$$

In our decentralized and asynchronous multi-agent setting, each agent j individually selects $\mathbf{a}_t^j$ with its own sampled $\tilde{\mathcal{X}}_t^j$. Hence the stochasticity in the sampling procedure enables agents to make decentralized explore-exploit decisions for simultaneous search-and-tracking in their action space.

Remark 1. Prior work in search-and-tracking (Papaioannou et al., 2020; Chen and Dames, 2022) tends to separate the search and tracking phases of the task, and maintains either a visit count or dynamic occupancy grid to compute the action selection objective during the exploration phase. Such methods scale poorly with the size of the environment since agents need to maintain a discretization over the search space (Van Nguyen et al., 2022). In contrast, our SMC inference for multi-target belief is parallelizable over particles in the posterior PHD, while our TS-based decision making is scalable with increasing teamsize J.

5. Results

We now present our experimental results. For a description of the setup and parameters, please refer to Appendix D. In the following experiments, we measure performance in terms of the average OSPA for the entire team of agents. The plots show mean across 10 random trials with the shaded regions indicating standard error.

TS-PHD-I vs. TS-PHD-II. Fig. 2 shows that decision making with TS-PHD-II (DecSTER-II) outperforms that with TS-PHD-I (DecSTER-I) by achieving a lower OSPA for the same number of measurements per agent. J denotes the teamsize. TS-PHD-II samples both the cardinality and locations of the target RFS from the PHD, so the samples for different agents are sufficiently diverse to capture the uncertainty regarding the true multi-target ground truth. In contrast, the samples from TS-PHD-I are generally clustered around the agent's current estimate of target locations. Fig. 2 further demonstrates the scalability of TS in the decentralized multi-agent active search-and-tracking setting. When teamsize increases n times, agents achieve similar OSPA with n times fewer measurements per agent (Ghods et al., 2021).

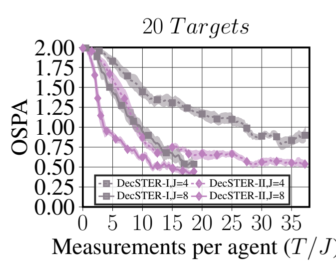


Figure 2

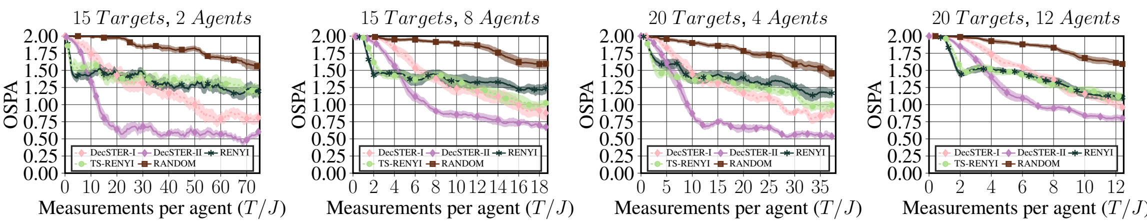


Figure 1: **Baseline comparisons**. For different numbers of targets and with fewer agents than targets, DecSTER with TS-PHD-II (denoted DecSTER-II) outperforms random sensing (RANDOM) and information greedy baselines (RENYI, TS-RENYI) by achieving a lower OSPA for the same number of measurements per agent.

Baselines comparisons. We compare our algorithms, DecSTER-I using TS-PHD-I and DecSTER-II using TS-PHD-II, with a random sensing baseline (RANDOM) and information-greedy baselines RENYI and TS-RENYI (Appendix C). Fig. 1 shows that DecSTER outperforms all the baselines for different number of targets k and team sizes J. Since RENYI agents are information-greedy, the lack of stochasticity in their decision making objective leads different agents to select the same action in the decentralized asynchronous multi-agent setting. Moreover, their objective computation depends only on the particles in the predicted PHD filter and does not account for previously undetected targets. This highlights the drawback of using Renyi divergence as an optimization objective for explore-exploit decisions in the search-and-tracking setting, in contrast with its success in the tracking only setting where exploration is not a concern (Papaioannou et al., 2020). As a result, we propose the TS-RENYI baseline in order to encourage exploration with samples drawn from TS-PHD-II. We observe that TS-RENYI still does not perform noticeably better than RENYI. This is because the weights of the particles in the SMC-PHD filter relate to the expected cardinality of the target set, therefore Renyi divergence does not account for any measure of the distance error between the current target estimates $\hat{\mathcal{X}}_t^j$ in the PHD filter (or the drawn Thompson sample $\tilde{\mathcal{X}}_t^j$) and the expected one-step lookahead estimate $\hat{\mathcal{X}}_{t+1}^j$. In contrast, the OSPA objective accounts for both localization error as well as cardinality error in the estimated target set. Thus we observe that our algorithm DecSTER-I is competitive with or outperforms random sensing and information-greedy baselines, and DecSTER-II consistently achieves the lowest OSPA among all with the same number of measurements per agent.

6. Conclusion

We introduce DecSTER, a novel decentralized and asynchronous algorithm for multi-agent multi-target active search-and-tracking that relaxes the restrictive assumption of requiring continuous target coverage. In simulation, DecSTER outperforms competitive baselines that greedily optimize for information gain or expected target detections. A key contribution is adapting TS to effectively drive exploration and exploitation using the SMC-PHD filter. Future work includes theoretical analysis of the proposed TS methods and learning improved models of environment uncertainty for non-stationary multi-target tracking.

References

Nikhil Angad Bakshi, Tejus Gupta, Ramina Ghods, and Jeff Schneider. Guts: Generalized uncertainty-aware thompson sampling for multi-agent active search. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pages 7735–7741, 2023. doi: 10.1109/ICRA48891.2023.10160597.

Samuel S Blackman. Multiple hypothesis tracking for multiple target tracking. *IEEE Aerospace and Electronic Systems Magazine*, 19(1):5–18, 2004.

Jun Chen and Philip Dames. Distributed multi-target search and tracking using a coordinated team of ground and aerial robots. In *Robotics science and systems*, 2020.

Jun Chen and Philip Dames. Active multi-target search using distributed thompson sampling. 2022.

Philip Dames, Pratap Tokekar, and Vijay Kumar. Detecting, localizing, and tracking an unknown number of moving targets using a team of mobile robots. *The International Journal of Robotics Research*, 36(13-14):1540–1553, 2017.

Philip M Dames. Distributed multi-target search and tracking using the phd filter. *Autonomous robots*, 44(3-4):673–689, 2020.

Lefteris Doitsidis, Stephan Weiss, Alessandro Renzaglia, Markus W Achtelik, Elias Kosmatopoulos, Roland Siegwart, and Davide Scaramuzza. Optimal surveillance coverage for teams of micro aerial vehicles in gps-denied environments using onboard vision. *Autonomous Robots*, 33:173–188, 2012.

Thomas Fortmann, Yaakov Bar-Shalom, and Molly Scheffe. Sonar tracking of multiple targets using joint probabilistic data association. *IEEE journal of Oceanic Engineering*, 8(3):173–184, 1983.

Ramina Ghods, Arundhati Banerjee, and Jeff Schneider. Decentralized multi-agent active search for sparse signals. In *Uncertainty in Artificial Intelligence*, pages 696–706. PMLR, 2021.

Ronald PS Mahler. Multitarget bayes filtering via first-order multitarget moments. *IEEE Transactions on Aerospace and Electronic systems*, 39(4):1152–1178, 2003.

Ronald PS Mahler. *Advances in statistical multisource-multitarget information fusion*. Artech House, 2014.

Robin R Murphy. Human-robot interaction in rescue robotics. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 34(2):138–153, 2004.

Thomas Oskam, Robert W Sumner, Nils Thuerey, and Markus Gross. Visibility transition planning for dynamic camera control. In *Proceedings of the 2009 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, pages 55–65, 2009.

Savvas Papaioannou, Panayiotis Kolios, Theocharis Theocharides, Christos G Panayiotou, and Marios M Polycarpou. A cooperative multiagent probabilistic framework for search and track missions. *IEEE Transactions on Control of Network Systems*, 8(2):847–858, 2020.

Branko Ristic, Daniel Clark, and Ba-Ngu Vo. Improved smc implementation of the phd filter. In *2010 13th International Conference on Information Fusion*, pages 1–8. IEEE, 2010.

Branko Ristic, Ba-Ngu Vo, and Daniel Clark. A note on the reward function for phd filters with sensor control. *IEEE Transactions on Aerospace and Electronic Systems*, 47(2): 1521–1529, 2011.

Cyril Robin and Simon Lacroix. Multi-robot target detection and tracking: taxonomy and survey. *Autonomous Robots*, 40:729–760, 2016.

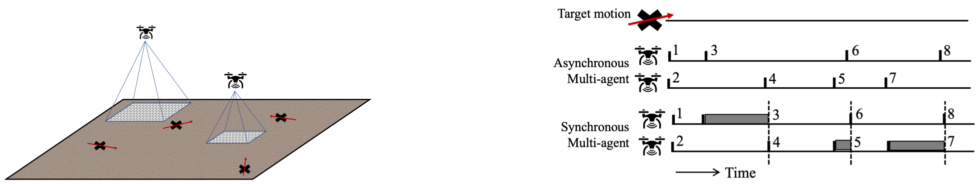
Daniel J Russo, Benjamin Van Roy, Abbas Kazerouni, Ian Osband, Zheng Wen, et al. A tutorial on thompson sampling. *Foundations and Trends® in Machine Learning*, 11(1): 1–96, 2018.

Lawrence D Stone, Roy L Streit, Thomas L Corwin, and Kristine L Bell. *Bayesian multiple target tracking*. Artech House, 2013.

Hoa Van Nguyen, Ba-Ngu Vo, Ba-Tuong Vo, Hamid Rezatofighi, and Damith C Ranasinghe. Multi-objective multi-agent planning for discovering and tracking multiple mobile objects. *arXiv preprint arXiv:2203.04551*, 2022.

Pujie Xin and Philip Dames. Comparing stochastic optimization methods for multi-robot, multi-target tracking. In *International Symposium on Distributed and Autonomous Systems*, 2022.

Zeyu Zhou, Bruce Hajek, Nakjung Choi, and Anwar Walid. Regenerative particle thompson sampling. *arXiv preprint arXiv:2203.08082*, 2022.



(a) (b)

Figure 3: **Problem setup.** (a) Agents sense different regions of the search space at different vertical heights, receiving noisy 2D location coordinates of the possible targets in their field of view, along with false positive measurements. The targets shown as black crosses move in the search space with different velocities shown by the red arrows. (b) The line at the top indicates the target's continuous motion with time. In our asynchronous multi-agent setup, agents can collect observations without waiting for their teammates whereas in the synchronous setting, the solid boxes indicate the agents' idle wait times.

Appendix A. Brief description of the PHD filter

The (noisy) target dynamics from state $\boldsymbol{\xi}$ to $\mathbf{x}$ is captured by the target motion model $f(\mathbf{x}|\boldsymbol{\xi})$. The survival probability $p_s(\mathbf{x})$ denotes the target's chances of persisting over successive time steps. The PHD filter formulates the following steps to propagate the posterior density over target states.[1]

$$\text{Prediction: } \bar{\nu}_t(\mathbf{x}) = b(\mathbf{x}) + \int_E f(\mathbf{x}|\boldsymbol{\xi}) p_s(\boldsymbol{\xi}) \nu_{t-1}(\boldsymbol{\xi}) d\boldsymbol{\xi} \tag{2}$$

$$\text{Update: } \nu_t(\mathbf{x}) = (1 - p_d(\mathbf{x}|\mathbf{q}))\bar{\nu}_t(\mathbf{x}) + \sum_{\mathbf{z}\in\mathcal{Z}_t} \frac{\psi_{\mathbf{z},\mathbf{q}}(\mathbf{x})\bar{\nu}_t(\mathbf{x})}{\eta_{\mathbf{z}}(\bar{\nu}_t)} \tag{3}$$

$$\eta_{\mathbf{z}}(\nu) = \kappa(\mathbf{z}|\mathbf{q}) + \int_E \psi_{\mathbf{z},\mathbf{q}}(\mathbf{x})\nu(\mathbf{x})d\mathbf{x} \tag{4}$$

$$\psi_{\mathbf{z},\mathbf{q}}(\mathbf{x}) = g(\mathbf{z}|\mathbf{x},\mathbf{q})p_d(\mathbf{x}|\mathbf{q}) \tag{5}$$

Here, $\psi_{\mathbf{z},\mathbf{q}}(\mathbf{x})$ is the probability that the agent at $\mathbf{q}$ receives the measurement $\mathbf{z}$ from a target $\mathbf{x}$ and $g(\mathbf{z}|\mathbf{x},\mathbf{q})$ is the measurement likelihood model. The PHD filter can handle appearing and disappearing targets by defining an appropriate birth density $b(\mathbf{x})$ over the search space, but in our experiments the number of ground truth targets k is fixed.

Appendix B. Our proposed algorithm

Algorithm 3 outlines our proposed algorithm called DecSTER (*Dec*entralized Multi-Agent Active *S*earch-and-*T*racking without continuous cov*er*age) for simultaneous multi-target active search and tracking without continuously following the targets.

1. For a detailed understanding of the PHD filter, please refer to Mahler (2014).

Algorithm 3 DecSTER for agent j at time t

1: **Input:** PHD $\nu_t^j = \{(w_1^j, \mathbf{x}_1^j), \ldots, (w_\rho^j, \mathbf{x}_\rho^j)\}$
2: Compute predicted PHD $\bar{\nu}_{t+1}^j$ (Eq. (2)).
3: Draw TS $\tilde{\mathcal{X}}_t^j \sim \bar{\nu}_{t+1}^j$.
4: Assuming pseudo-measurements at $\tilde{\mathcal{X}}_t^j$, estimate expected target set $\mathcal{Y}_t^j$ and select action $\mathbf{a}_t^j$ (Eq. (1)).
5: Observe $\mathcal{Z}_t^j$. Update PHD ν_{t+1}^j (Eq. (3)).
6: Estimate target set $\hat{\mathcal{X}}_{t+1}^j$ from ν_{t+1}^j (Ristic et al., 2010).
7: Asynchronously communicate $(t, \mathbf{a}_t^j, \mathcal{Z}_t^j)$ with team.

Appendix C. Baselines descriptions

We compare DecSTER-I and DecSTER-II with the following baselines. Note that all of them use the same PHD filter inference method, but differ in the action selection policy.

1) RANDOM. Each agent j selects its next sensing action uniformly at random. **2) RENYI.** At t, agent j computes the predicted PHD $\bar{\nu}_{t+1}^j$ and generates a (pseudo) measurement set $\bar{\mathcal{Z}}_t^j$ for any action $\mathbf{a} \in \mathcal{A}$ assuming the estimated target set $\hat{\mathcal{X}}_t^j$ from ν_t^j as ground truth. It then selects the action $\mathbf{a}_t^j$ that maximizes the Renyi divergence (with $\alpha = 0.5$) between $\bar{\nu}_{t+1}^j$ and its expected updated PHD ν'^j_{t+1} (Eq. (3)). With the SMC-PHD formulation, the Renyi divergence is (Ristic et al., 2011):

$$\sum_{i=1}^{\rho} \bar{w}_i + \frac{\alpha}{1-\alpha} \sum_{i=1}^{\rho} w'_i - \frac{1}{1-\alpha} \sum_{i=1}^{\rho} w'^{\alpha}_i \bar{w}_i^{1-\alpha} \tag{6}$$

where $\bar{w}_i$ and w'_i are the weights of the particle i in $\bar{\nu}_{t+1}^j$ and ν'^j_{t+1} respectively. **3) TS-RENYI.** We modify RENYI to use $\tilde{\mathcal{X}}_t^j \sim \bar{\nu}_{t+1}^j$ (with TS-PHD-II) for computing the (pseudo) measurement set $\bar{\mathcal{Z}}_t^j$ and the updated weights w'_i.

Appendix D. Additional results

We first describe our experimental setup. Consider a 2D search space with dimensions $n_l \times n_w = 16 \times 16$. There are k targets moving in this region, whose starting locations and velocities are chosen uniformly at random, such that $v_{\max} = 0.1$. A team of J agents are tasked with search-and-tracking of all the targets over $T = 150$ steps. The agents' action space $\mathcal{A}$ consists of hierarchical region sensing actions of width 1×1, 2×2, 4×4 and 8×8, $|\mathcal{A}| = 340$. Since actions with larger FOV receive noisier observations, we vary the false positive (clutter) rate as $\lambda \in \{0.005, 0.04, 1, 5\}$ for action widths 1, 2, 4 and 8 respectively. The survival probability in the PHD filter is set at $p_s = 1$ and the detection probability $p_d = 0.9$ for targets in the agent's FOV. We choose $\tilde{\rho} = 100$ (Algorithm 1).

The agents assume the target motion model $\mathbf{x}_{t+1} = \mathbf{F}\mathbf{x}_t + \boldsymbol{\epsilon}$, where $\mathbf{F} = \begin{bmatrix} 1 & 0 & \Delta T & 0 \\ 0 & 1 & 0 & \Delta T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$, $\Delta T = 1$ and $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q})$, $\mathbf{Q} = \begin{bmatrix} 0.03 & 0 & 0.05 & 0 \\ 0 & 0.03 & 0 & 0.05 \\ 0.05 & 0 & 0.1 & 0 \\ 0 & 0.05 & 0 & 0.1 \end{bmatrix}$. The sensing model is $\mathbf{z} = \mathbf{H}\mathbf{x} + \boldsymbol{\omega}$, where $\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$ and $\boldsymbol{\omega} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$, $\sigma = 0.1$. $\mathbf{I}$ is the identity matrix.

Given sets $\mathcal{X}$ and $\mathcal{Y}$, where $|\mathcal{X}| = m \leq |\mathcal{Y}| = n$ without loss of generality,

$$\text{OSPA}(\mathcal{X}, \mathcal{Y}) = \Big(\frac{1}{n} \min_{\pi \in \Pi_n} \sum_{i=1}^{m} d_c(x_i, y_{\pi(i)})^p + c^p(n-m)\Big)^{\frac{1}{p}}$$

where c is the cut-off distance, $d_c(x, y) = min(c, ||x - y||)$ and Π_n is the set of all permutations of the set $\{1, \ldots, n\}$. The distance error component of the OSPA computes the minimum cost assignment between $\mathcal{X}$ and $\mathcal{Y}$, such that $x_i \in \mathcal{X}$ is matched to $y_{i'} \in \mathcal{Y}$ only when they are within a distance c of each other. In our experiments, we set $c = 2$, $p = 1$.

Based on the results in Section 5, we consider DecSTER-II as our best approach in this setting, labeled DecSTER in the following experiments.

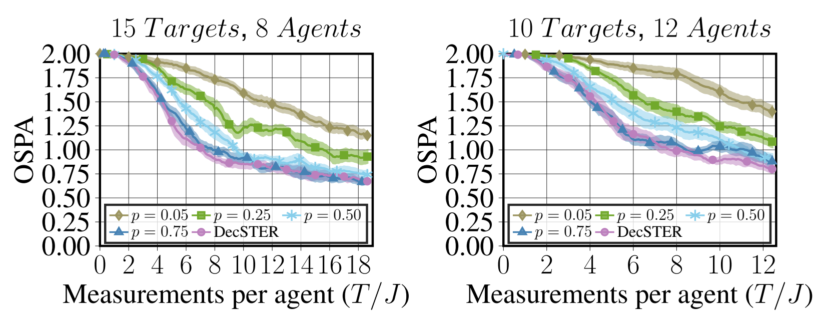


Figure 4: **Robustness to unreliable communication**. When agents communicate their actions and observations with decreasing probability p, DecSTER experiences a graceful deterioration in OSPA performance and agents require increasingly more measurements to estimate the number and locations of true targets in the search space.

Robustness to communication delays. Multi-agent systems benefit from leveraging observations shared by their teammates. Agents in our decentralized and asynchronous multi-agent setting benefit from any information shared by their teammates but they can continue searching for and tracking targets without waiting for such communication. Therefore, we now analyze the robustness of DecSTER under unreliable inter-agent communication. In simulation, we consider each agent chooses to communicate its own observation at time t, along with any prior observations it had not shared with its teammates, with a probability $p \in \{0.05, 0.25, 0.50, 0.75, 1\}$. The $p = 1$ setting corresponds to our description and analysis of DecSTER in Fig. 1. We observe a graceful decay in the OSPA performance with decreasing rates of inter-agent communication in Fig. 4, both when targets outnumber agents and vice versa. Compared to prior work in the centralized or distributed multi-agent tracking setting (Robin and Lacroix, 2016), DecSTER does not depend on synchronized communication within the team, thus agents can adapt and continue their search-and-tracking tasks even when communication is unreliable or unavailable.