

# Probabilistic Search and Track with Multiple Mobile Agents

Savvas Papaioannou, Panayiotis Kolios, Theocharis Theocharides,  
Christos G. Panayiotou and Marios M. Polycarpou

**Abstract**—In this paper we are interested in the task of searching and tracking multiple moving targets in a bounded surveillance area with a group of autonomous mobile agents. More specifically, we assume that targets can appear and disappear at random times inside the surveillance region and that their positions are random and unknown. The agents have a limited sensing range, and due to sensor imperfections they receive noisy measurements from the targets. In this work we utilize the theory of random finite sets (RFS) to capture the uncertainty in the time-varying number of targets and their states and we propose a decision and control framework, in which the mode of operation (i.e. *search* or *track*) as well as the mobility control action for each agent, at each time instance, are determined so that the collective goal of searching and tracking is achieved. Extensive simulation results demonstrate the effectiveness and performance of the proposed solution.

## I. INTRODUCTION

The deployment of multiple autonomous drones or unmanned aerial vehicles (UAVs) in rescue situations is an application scenario that has recently attracted considerable interest [1]. Indeed the recent advances in small UAVs and the cost reduction of electronic components have spurred an unprecedented interest in autonomous multi-agent systems for group missions [2]. Of particular interest in this work and motivating scenario is the use of a team of autonomous mobile agents in search and rescue missions. More specifically, we believe that a team of autonomous rescue drones can provide an important aid in many search and rescue operations by efficiently searching the area for victims or targets of interest and by accurately tracking the discovered targets until further assistance is available.

In this work, a probabilistic approach based on multi-object state estimation using random finite sets (RFS) [3] is derived to jointly address the problem of searching and tracking multiple moving targets by a team of autonomous mobile agents. This is a challenging problem since a) the mobile agents exhibit a limited sensing range thus there is a need for efficient searching, b) due to sensor imperfections the agents receive noisy measurements and false alarms which requires efficient algorithms for estimating the number of targets and their states and c) finally the problem becomes even more challenging due to the tradeoff which arises on how many resources to allocate in each task i.e. *searching* or *tracking* at each time instance so that the joint search-and-track objective is achieved as best as possible. The contributions of this work are the following:

- Provides a novel probabilistic framework for the problem of joint search-and-track with a team of autonomous

mobile agents by accounting the uncertainty in the number of targets, the target dynamics, the target detection process and the limited sensing range for target search.

- Develops a novel decision (i.e. the agents can switch between *search* and *track* mode during the mission) and control (i.e. controlling the movement of all agents) algorithm that optimizes the overall search-and-track objective.
- Demonstrates the performance of the proposed approach through extensive simulated scenarios.

The rest of the paper is organized as follows. Section II reviews the existing literature on searching and tracking by single and multiple agents, demonstrating in the way the novelty of this work. Section III outlines the problem and introduces the proposed system architecture. Section IV provides a review on stochastic filtering and random finite sets (RFS) and Section V develops the problem formulation and describes the details of the proposed approach. Finally, Section VII conducts an extensive performance analysis and Section VIII concludes the paper and discusses future work.

## II. RELATED WORK

Coordinated teams of agents unlock significantly greater capabilities than what is possible by soloing. Taking this fact into consideration, existing literature has looked into several challenging problems that multi-agent systems can successfully tackle including the use of drone swarms in emergency response operations [4]. The coverage/search problem is among the most cited problems to tackle whereby a fleet of airborne agents need to spread across an area over the shortest time interval for situational awareness or when searching for particular objects. The task assignment, scheduling and path planning problems, considering physical resource constraints such as the total number of agents, their battery levels and their communication ranges have also been considered lately, [5] [6] [7]. Another challenging problem is that of tracking of single or multiple targets. For this problem, the majority of strategies seek to minimize the tracking error for the detected targets [8] [9] [10].

In this work our focus is on the joint search-and-track problem which arises especially during and in the aftermath of large-scale disasters. Searching is necessary to detect targets before accurately tracking them, thus efficient multi-agent searching techniques are of high importance. In addition in many real-world search-and-track scenarios the ability to keep track of multiple targets as accurately as possible is essential. Interesting works on the problem of search-and-track are in [11] and [12]. In [11] the authors develop a probabilistic approach for the search-and-track problem however, they consider only the single-target single-agent

The authors are with the KIOS Research and Innovation Centre of Excellence, and the Department of Electrical and Computer Engineering, University of Cyprus, {papaioannou.savvas, pkolios, ttheocharides, christosp, mpolycar}@ucy.ac.cy

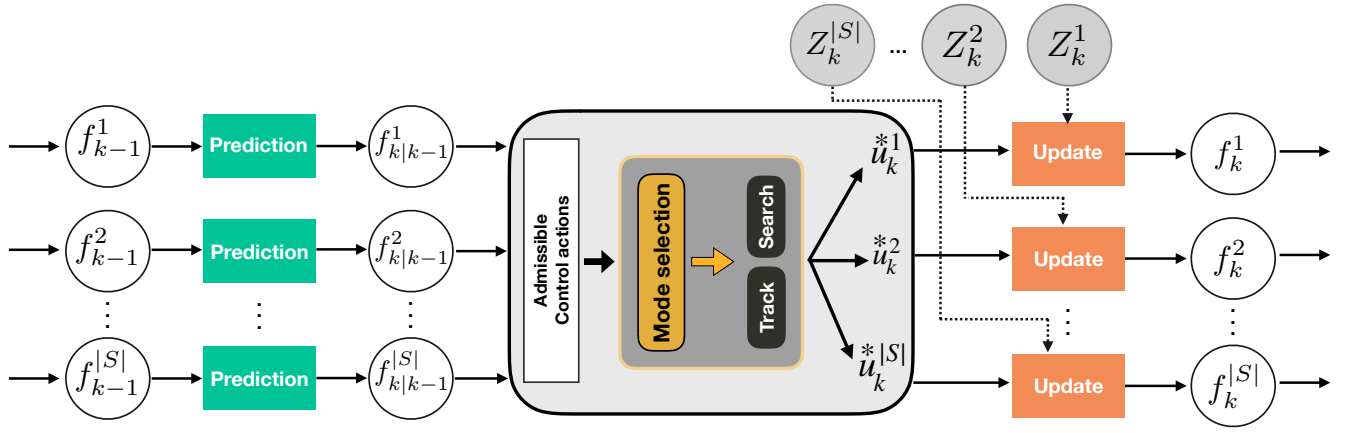


Fig. 1. The figure illustrates the proposed system architecture. The proposed approach uses the following recursion: a) for each agent  $j$  the predictive multi-object probability distribution  $f_{k|k-1}^j$  is first computed from the posterior  $f_{k-1}^j$  in the  $k$ th time-step, b) the number of targets and their states are extracted from each predictive multi-object distribution, c) a decision and control problem is then solved for all agents to decide the mode of operation and control actions, d) finally the agents apply the optimal control actions and move to their new states where they receive target measurements and compute the posterior multi-object density  $f_k^j$  through the measurement update step.

scenario. Similarly, a more recent work [12] investigates the same problem and proposes a model-predictive control (MPC) framework for a single-agent. Initial works on the problem of multi-agent search-and-track are discussed in [13], [14] but only for stationary targets. The work in [15] mainly focuses on path planning algorithms for searching and tracking without considering the problems of false alarms and data association. Finally, a more recent work in [16] focuses on the scalability issue when a large number of agents are present and proposes a computationally efficient clustering algorithm.

Complementary to the aforementioned studies, this work advances the state-of-the-art, by firstly proposing a probabilistic multi-agent framework which accounts for the uncertainties in the number of targets, target dynamics, measurement models and false-alarms and secondly by dynamically adjusting the mode of operation (i.e. search or track) and by determining the control actions of multiple agents so that the collective search-and-track objective is achieved.

### III. PROBLEM DEFINITION AND SYSTEM OVERVIEW

In this work we assume that a bounded region of interest needs to be searched for potential targets by a group of mobile agents. The targets can enter and exit the surveillance region at random times (i.e. the number of targets is not fixed and needs to be estimated) and their positions are unknown. In addition, the agents are equipped with imperfect sensors and thus they detect targets within their limited sensing range with probability less than one. Moreover, the agents receive noisy measurements from the targets and at random times they receive spurious measurements (e.g. false alarms/clutter) as well. The resulting problem we would like to address is the following:

*At every discrete time-step, a team of mobile agents must jointly decide their mode of operation (i.e. search or track) and select their mobility control actions in order to find the best tradeoff between a) maximizing the area coverage and b) accurately tracking all detected targets.*

The proposed approach is based on the framework of multi-object (or multi-target) filtering with random finite sets [3]. This framework extends the stochastic filtering theory [17] to allow us to jointly estimate both the number and the states of an unknown number of objects (or targets) from a set of measurements in the presence of clutter. In this section we provide an overview of the proposed system architecture, as shown in Fig. (1).

Let  $S = \{1, 2, \dots, |S|\}$  be the set of available mobile agents. Each agent  $j$  can operate in either *search* or *track* mode and in each mode the agent is subject to a set of admissible control actions, denoted by  $\mathbb{U}_k^j$  for time  $k$ . When a subset of agents  $\alpha \subseteq S$  is in *search* mode the collective objective over these agents is the maximization of the covered area. In other words the agents should traverse the surveillance area in such a way so that the coverage is maximized. On the other hand, when a subset of agents  $\alpha^c \subseteq S$  (where  $\alpha^c$  denotes the complement of  $\alpha$  with respect to  $S$ ) is in *track* mode, the objective is the maximization of the tracking performance, i.e. the number of targets and their states inside the surveillance area should be estimated as accurately as possible. The cost of having a subset of agents in *search* or *track* mode is given by the objective functions  $\xi_{\text{search}}(\alpha)$  and  $\xi_{\text{track}}(\alpha^c)$ , respectively. The proposed controller solves the following:

*At each time-step  $k$  it finds the subsets of agents  $\alpha \subseteq S$  and  $\alpha^c \subseteq S$  for searching and tracking, respectively, and the appropriate control actions ( $u_k^j \in \mathbb{U}_k^j \forall j$ ) for each agent so that the total cost of searching and tracking  $\xi_{\text{search}}(\alpha) + \xi_{\text{track}}(\alpha^c)$  is minimized.*

In order to achieve this, the proposed system takes the following steps: Each agent maintains a multi-object probability distribution which jointly accounts for the uncertainty in the number of targets and their states (i.e. position and velocity) inside its sensing range (this multi-object probability distribution is a generalization of the probability density function on random finite sets). Let this posterior distribution at time  $k-1$  be denoted by  $f_{k-1}^j$  for agent  $j$ . In the first step,

the predictive density  $f_{k|k-1}^j$  is obtained through a prediction step, as shown in Fig. 1. Then the predicted number of targets and their states are extracted from the predictive density and based on these predictions, the proposed controller finds the combination of hypothesized control actions ( $u_k^j \in \mathbb{U}_k^j \forall j$ ) as well as the mode of operation of each agent which minimizes the total cost of the search-and-track objective. The optimal hypothesized control actions  $u_k^{*1}, \dots, u_k^{*|S|}$  are then applied to the agents. The agents move to their new states at time  $k$  where they can receive the measurements sets  $Z_k^1, \dots, Z_k^{|S|}$  from the targets (if any). These measurement sets at time  $k$  are then used in the measurement update step i.e., each agent uses the received measurements to compute the posterior multi-object distribution  $f_k^j$  from which the posterior number of targets and their states can be estimated.

The same procedure is repeated recursively over time. We should note here that the number of targets (and their states) is being estimated recursively over time and that based on this estimate the mode of operation should be carefully chosen. In the next section, we provide a brief overview on single and multi-object stochastic filtering which will help us set the basis of the proposed approach.

#### IV. BACKGROUND ON STOCHASTIC FILTERING

In this section we provide a brief overview on the theory of random finite sets (RFS) and stochastic filtering that will be used in this framework. A more detailed description can be found in [3], [18].

##### A. Stochastic Filtering

In stochastic filtering, we are interested in the posterior probability density of the hidden state  $x_k$  at time  $k$  given all measurements  $z_{1:k} = (z_1, \dots, z_k)$  up to time  $k$ , denoted by:

$$p_k(x_k | z_{1:k}) \quad (1)$$

In this paper, variable  $x_k$  denotes the state of a single target at time  $k$ . Given an initial density on the target state  $p_0(x_0)$ , the posterior density at time  $k$  can be computed using the Bayes recursion [18] as:

$$p_{k|k-1}(x_k | z_{1:k-1}) = \int \pi_{k|k-1}(x_k | x_{k-1}) p_{k-1}(x_{k-1} | z_{1:k-1}) dx_{k-1} \quad (2)$$

$$p_k(x_k | z_{1:k}) = \frac{g_k(z_k | x_k) p_{k|k-1}(x_k | z_{1:k-1})}{\int g_k(z_k | x_k) p_{k|k-1}(x_k | z_{1:k-1}) dx_k} \quad (3)$$

where Eqn. (2) and (3) are referred to as the prediction and measurement update steps, respectively. In the prediction step, the Chapman-Kolmogorov equation is used to calculate the predictive density  $p_{k|k-1}(x_k | z_{1:k-1})$  i.e. the predicted probability density function (pdf) of  $x_k$  based on the measurements up to time  $k-1$ . The measurement update step applies the Bayes rule to update the predicted pdf using the current measurement  $z_k$  and obtain the posterior distribution  $p_k(x_k | z_{1:k})$  for the state  $x_k$ . The function  $\pi_{k|k-1}(x_k | x_{k-1})$  models the uncertainty of the current state given the previous state and is referred to as the transitional density and the function  $g_k(z_k | x_k)$  describes the distribution

of measurements given the current state and is referred to as the measurement likelihood function. Finally, once the posterior density is obtained, the state  $x_k$  can be extracted using for instance the expected a posteriori (EAP) or the maximum a posteriori (MAP) estimators.

##### B. Random Finite Sets

A random finite set [3], [19], [20] (RFS) is a finite-set-valued random variable, which differs from a random vector in two ways: a) the number of elements in a RFS is random and b) the order of the elements in a RFS is irrelevant. More specifically, a RFS  $X = \{x_1, x_2, \dots, x_n\}$  is completely specified by a) its cardinality distribution  $\rho(n) = p(|X| = n)$  which defines the probability over the number of elements in  $X$  and b) by a family of conditional joint probability distributions  $p(x_1, \dots, x_n | n)$ ,  $x_1, \dots, x_n \in \mathcal{X}$  over the distribution of its elements. The belief density or otherwise known multi-object/multi-target pdf  $f(X)$  of the RFS  $X$  is given by:  $f(X) = f(\{x_1, \dots, x_n\}) = n! \rho(n) p(x_1, \dots, x_n | n)$  and the notion of integration is given by the set-integral which is defined as:

$$\int f(X) \delta X = f(\emptyset) + \sum_{n=1}^{\infty} \frac{1}{n!} \int f(\{x_1, \dots, x_n\}) d_{x_1} \dots d_{x_n} \quad (4)$$

That said, it has been shown [21] that the stochastic filtering recursion of Eqn. (2)-(3) can be extended for RFSs in order to allow for the recursive computation of the multi-object posterior density  $f_k(X_k | Z_{1:k})$  of RFS  $X_k$  given RFS measurements  $Z_{1:k}$ . In this recursion the random variables  $x_k$  and  $z_k$  are represented by the RFSs'  $X_k$  and  $Z_k$ , the functions  $\pi_{k|k-1}(x_k | x_{k-1})$  and  $g_k(z_k | x_k)$  are replaced by the multi-object transition density and the multi-object likelihood function respectively and finally the integrals have become set integrals as defined above. This recursion (i.e. the multi-object stochastic filtering) however, has in general no analytic solution, and because of its combinatorial complexity is usually intractable. Thus, approximations exist which propagate only summary statistics, instead of the full posterior density.

A prominent approach which is summarized here is the multi-Bernoulli filter [22], which is based on the approximation of the multi-object posterior density  $f_k(X_k | Z_{1:k})$  as multi-Bernoulli pdf. In essence this approach only propagates the multi-Bernoulli parameters  $\{(r_k^i, p_k^i(x))\}_{i=1}^{v_k}$  in time, using a predictor-corrector scheme with computational complexity  $O(mn)$ , where  $m$  is the number of measurements and  $n$  is the number of targets.

More specifically, let the multi-object posterior density at time  $k$  be approximated by a multi-Bernoulli distribution  $f_k = \{(r_k^i, p_k^i(x))\}_{i=1}^{v_k}$  where  $v_k$  is the maximum number of hypothesized targets at time  $k$ ,  $r_k^i$  denotes the probability that the  $i$ th target actually exists and  $p_k^i(x)$  is the spatial density of the  $i$ th target. The predictive multi-object density at time  $k$  is given by:  $f_{k|k-1} = f_{k|k-1}^{\text{persist}} \cup f_k^{\text{birth}}$  where  $f_{k|k-1}^{\text{persist}} = \{(r_{k|k-1}^i, p_{k|k-1}^i(x))\}_{i=1}^{v_{k-1}}$  describes the set of persisting targets from the previous time-step and  $f_k^{\text{birth}} = \{(r_k^i, p_k^i(\cdot))\}_{i=1}^k$  are the parameters of a multi-Bernoulli RFS for the new-born targets at time  $k$ . The distribution of

persisting targets  $f_{k|k-1}^{\text{persist}}$  is given by:

$$r_{k|k-1}^i = r_{k-1}^i \langle p_{k-1}^i, p_S \rangle \quad (5)$$

$$p_{k|k-1}^i(x) = \frac{\langle \pi_{k|k-1}(x), p_{k-1}^i p_S \rangle}{\langle p_{k-1}^i, p_S \rangle} \quad (6)$$

where the notation  $\langle f, g \rangle$  is defined as  $\int f(x)g(x)dx$  and  $p_S(x)$  denotes the probability of target survival. Given the predictive multi-object density  $f_{k|k-1} = \{(r_{k|k-1}^i, p_{k|k-1}^i(x))\}_{i=1}^{v_{k|k-1}}$  (where the term  $v_{k|k-1}$  denotes the number of hypothesized Bernoulli components after accounting for the target births) and the new measurement set  $Z_k = \{z_k^1, \dots, z_k^{m_k}\}$  then the posterior multi-object distribution at time  $k$  is approximated by the multi-Bernoulli distribution  $f_k = f_k^{\text{legacy}} \cup f_k^{\text{meas}}$  where  $f_k^{\text{legacy}} = \{(r_k^i, p_k^i(x))\}_{i=1}^{v_{k|k-1}}$  contains legacy targets i.e. updates of predicted targets, assuming that no measurements were collected from them and given by [22]:

$$r_k^i = r_{k|k-1}^i \frac{1 - \langle p_{k|k-1}^i, p_D \rangle}{1 - r_{k|k-1}^i \langle p_{k|k-1}^i, p_D \rangle} \quad (7)$$

$$p_k^i(x) = p_{k|k-1}^i(x) \frac{1 - p_D(x)}{1 - \langle p_{k|k-1}^i, p_D \rangle} \quad (8)$$

where  $p_D(x)$  is the probability of detecting a target with state  $x$ .  $f_k^{\text{meas}} = \{(r_k(z), p_k(x, z))\}_{z \in Z_k}$  contains new targets i.e. joint updates of all predicted targets using each of the measurements separately and given by:

$$r_k(z) = \frac{\sum_{i=1}^{v_{k|k-1}} \frac{r_{k|k-1}^i (1 - r_{k|k-1}^i) \langle p_{k|k-1}^i, L_k^z \rangle}{(1 - r_{k|k-1}^i \langle p_{k|k-1}^i, p_D \rangle)^2}}{\kappa_k(z) + \sum_{i=1}^{v_{k|k-1}} \frac{r_{k|k-1}^i \langle p_{k|k-1}^i, L_k^z \rangle}{1 - r_{k|k-1}^i \langle p_{k|k-1}^i, p_D \rangle}} \quad (9)$$

$$p_k(x, z) = \frac{\sum_{i=1}^{v_{k|k-1}} \frac{r_{k|k-1}^i}{1 - r_{k|k-1}^i} p_{k|k-1}^i L_k^z(x)}{\sum_{i=1}^{v_{k|k-1}} \frac{r_{k|k-1}^i}{1 - r_{k|k-1}^i} \langle p_{k|k-1}^i, L_k^z \rangle} \quad (10)$$

where  $L_k^z(x) = g_k(z|x)p_D(x)$ .

## V. SYSTEM MODEL

In this section we discuss the details of the proposed approach. More specifically in this paper we make the following modeling assumptions:

### A. Single Target Dynamics

Lets assume first that the target's state vector  $x_k \in \mathcal{X}$  evolves in time according to the following equation:

$$x_k = \zeta_k(x_{k-1}) + v_k \quad (11)$$

where the possibly non-linear function  $\zeta_k$  which is usually referred to as the dynamical model, describes the evolution of the state vector as a first order Markov process with transitional density  $\pi(x_k|x_{k-1}) = p_v(x_k - \zeta_k(x_{k-1}))$ . The random process  $v_k \in \mathbb{R}^{n_x}$ , which is referred to as the process noise, is independent and identically distributed (IID) according to the probability density function  $p_v$ . The role of the process noise is to model random disturbances in

the evolution of the state. In this paper the state vector  $x_k$  is composed of position and velocity components i.e.  $x_k = [x, \dot{x}, y, \dot{y}]^\top$  in Cartesian coordinates.

### B. Agent dynamics

Let  $S = \{1, 2, \dots, |S|\}$  be the set of all mobile agents that operate in a discrete-time setting. At time  $k$ , the 2D surveillance area  $\mathcal{A} \subset \mathbb{R}^2$  is monitored by  $|S|$  mobile agents with states  $s_k^1, s_k^2, \dots, s_k^{|S|}$  taking values in  $\mathcal{A}$ . Each agent is subject to the following dynamics with bounded control inputs:

$$s_k^j = s_{k-1}^j + \begin{bmatrix} l_1 \Delta_R \cos(l_2 \Delta_\theta) \\ l_1 \Delta_R \sin(l_2 \Delta_\theta) \end{bmatrix}, \quad \begin{matrix} l_2 = 0, \dots, N_\theta \\ l_1 = 0, \dots, N_R \end{matrix} \quad (12)$$

where  $s_{k-1}^j = [s_x^j, s_y^j]^\top$  denotes the position (i.e. xy-coordinates) of the  $j$ th agent at time  $k-1$ ,  $\Delta_R$  is the radial step size,  $\Delta_\theta = 2\pi/N_\theta$  and the parameters  $(N_\theta, N_R)$  control the number of possible control actions. We denote the set of all admissible control vectors (or actions) of agent  $j$  at time  $k$  with  $\mathbb{U}_k^j = \{s_k^{j,1}, s_k^{j,2}, \dots, s_k^{j,|\mathbb{U}_k^j|}\}$  as computed by Eqn. (12).

### C. Agent sensing model

The ability of an agent to sense its 2D environment is modeled by the function  $p_D(x_k, s_k)$  that measures the probability that a target with state  $x_k$  at time  $k$  is detected by an agent with state (i.e. 2D position coordinates)  $s_k = [s_x, s_y]^\top \in \mathcal{A} \subseteq \mathbb{R}^2$ . More specifically the detection probability of a target with state  $x_k$  at position  $p_k = Hx_k$  (where  $H$  is a matrix that extracts the  $xy$ -coordinates of a target from its state vector) from an agent with state  $s_k$ , is given by:

$$p_D(x_k, s_k) = \begin{cases} p_D^{\text{max}} & \text{if } d_k < R_0 \\ \max\{0, p_D^{\text{max}} - \eta(d_k - R_0)\} & \text{if } d_k \geq R_0 \end{cases} \quad (13)$$

where  $d_k = \|Hx_k - s_k\|_2$  denotes the Euclidean distance between the agent and the target,  $p_D^{\text{max}}$  is the detection probability for targets that reside within  $R_0$  distance from the agent's position and finally  $\eta$  captures the reduced effectiveness of the agent to detect distant targets. In addition, when the agent detects a target, it receives a measurement vector  $z_k \in \mathcal{Z}$  (composed of range and bearing measurements) which is related to the target state as follows:

$$z_k = h_k(x_k, s_k) + w_k \quad (14)$$

where the non-linear function  $h_k$  projects the state vector to the observation space. The random process  $w_k \in \mathbb{R}^{n_z}$  is IID, independent of  $v_k$  and distributed according to  $p_w$ . The probability density of receiving measurement  $z_k$  for a target with state  $x_k$  is given by the measurement likelihood function  $g(z_k|x_k, s_k) = p_w(z_k - h_k(x_k, s_k))$ .

### D. Multi-Target Dynamics and Sensor Observation Models

Suppose now that multiple targets  $x_k^i$  exists inside the surveillance region at a particular time instance  $k$  where we assume that the targets are independent of each other. The targets can spawn from anywhere in the surveillance area and target births and deaths occur at random times. This means that at each time  $k$ , there exist  $n_k$  target states  $x_k^1, x_k^2, \dots, x_k^{n_k}$ , each taking values in the state space  $\mathcal{X}$

where both the number of targets  $n_k$  and their individual states  $x_k^i, \forall i$  are random and time-varying. The group of target states at time  $k$  can now be modeled as the RFS  $X_k = \{x_k^1, x_k^2, \dots, x_k^{n_k}\} \in \mathcal{F}(\mathcal{X})$  where  $\mathcal{F}(\mathcal{X})$  denotes the space of all finite subsets of  $\mathcal{X}$ . Note here that  $n_k$  is the true but unknown number of targets at time  $k$  which needs to be estimated along with the target states  $x_k^i, \forall i$ .

Let  $x_{k-1} \in X_{k-1}$  denote the state vector of a single target, and  $X_{k-1}$  the multi-target state (i.e. RFS) at time  $k-1$ . The multi-target dynamics is given by:

$$X_k = \left[ \bigcup_{x_{k-1} \in X_{k-1}} \Psi(x_{k-1}) \right] \cup \Gamma_k \quad (15)$$

where  $\Psi(x_{k-1})$  is a Bernoulli RFS [23] which models the evolution of the set from the previous state, with parameters  $r = p_{S,k}(x_{k-1})$  and  $p(x_k) = \pi(x_k|x_{k-1})$ . Thus a target with state  $x_{k-1}$  continues to exist at time  $k$  with surviving probability  $p_{S,k}(x_{k-1})$  and moves to a new state  $x_k$  with transition probability  $\pi(x_k|x_{k-1})$ . Otherwise the target dies with probability  $1 - p_{S,k}(x_{k-1})$ . The term  $\Gamma_k$  denotes the multi-Bernoulli RFS of spontaneous births [22].

An agent can receive  $m_k$  measurements  $z_k^1, z_k^2, \dots, z_k^{m_k}$ , from the targets and clutter, each taking values in the measurement space  $\mathcal{Z}$ , where  $m_k$  and  $z_k^i, \forall i$  are random and time varying. Thus the group of received measurements by an agent at time  $k$  can be modeled as the RFS  $Z_k = \{z_k^1, z_k^2, \dots, z_k^{m_k}\} \in \mathcal{F}(\mathcal{Z})$  where  $\mathcal{F}(\mathcal{Z})$  denotes the space of all finite subsets of  $\mathcal{Z}$ . As a consequence the agent  $j$  with state  $s_k^j \in \mathcal{A}$  receives measurements at time  $k$ ,  $z_k^{j,i} \in Z_k^j$  where  $Z_k^j$  is the RFS multi-target measurement set described by:

$$Z_k^j = \left[ \bigcup_{x_k^j \in X_k^j} \Theta(x_k^j) \right] \cup K_k \quad (16)$$

where  $\Theta(x_k^j)$  is a Bernoulli RFS which models the target generated measurements with parameters  $(p_D(x_k^j, s_k^j), g_k(z_k^{j,i}|x_k^j, s_k^j))$ . Thus a target with state  $x_k^j$  at time  $k$  is detected by the  $j$ th agent with state  $s_k^j$  with probability  $p_D(x_k^j, s_k^j)$  (see Eqn. (13)) and generates a measurement  $z_k^{j,i}$  with likelihood  $g_k(z_k^{j,i}|x_k^j, s_k^j)$  (see Eqn. (14)) or is missed with probability  $1 - p_D(x_k^j, s_k^j)$  and generates no measurements. The term  $K_k$  is a Poisson RFS [23] which models the set of false alarms or clutter received by the sensor at time  $k$  with intensity function  $\kappa_k(z_k^{j,i}) = \lambda f_c(z_k^{j,i})$ , where in this paper  $f_c(\cdot)$  denotes the uniform distribution over  $\mathcal{Z}$ .

## VI. MULTI-AGENT DECISION AND CONTROL

To allow the agents to switch modes (i.e. *search* or *track*) dynamically during the mission and to optimize the joint search-and-track objective the following optimization problem is to be solved at each time-step  $k$ :

$$(P_1) \arg \min_{u_k^j, \forall j \in S} \left[ \xi_{\text{search}}^\alpha(\{u_k^j : j \in \alpha\}) + \sum_{\substack{j=1 \\ j \neq \alpha}}^{|S|} \xi_{\text{track}}^j(u_k^j, Z_{k|k-1, u_k^j}^j) \right] \quad (17)$$

subject to  $\alpha \in \mathbb{P}(S)$

$$u_k^j \in \mathcal{U}_k^j, \forall j \in S$$

$$Z_{k|k-1, u_k^j}^j = Z_{\text{PIMS}}(\hat{X}_{k|k-1}^j, u_k^j), \forall j \in S$$

$$\|u_k^i - u_k^j\| > d_{\min}, \forall i \neq j$$

where  $\alpha$  is the decision variable for the agents in search mode and  $\mathbb{P}(S) = \{\emptyset, \{1\}, \{2\}, \dots, \{|S|\}, \{1, 2\}, \dots\}$  denotes the power set of  $S$ . The function  $\xi_{\text{search}}^\alpha(\{u_k^j : j \in \alpha\})$  takes as input a set of control actions for the agents in search mode and returns the total cost of searching and the function  $\xi_{\text{track}}^j(u_k^j, Z_{k|k-1, u_k^j}^j)$  returns the cost of tracking when the control  $u_k^j$  is taken and the predicted multi-target measurement set  $Z_{k|k-1, u_k^j}^j$  is supposed to have been received by agent  $j$ .  $\hat{X}_{k|k-1}^j$  denotes the predicted multi-target state for time  $k$  from agent  $j$  based on time  $k-1$ . Note that the cost of tracking depends on the received measurements which in turn depend on the control  $u_k^j$  that is applied to agent  $j$  (i.e. depending on the control action taken, different measurements will be received). The measurement generator function  $Z_{\text{PIMS}}(\hat{X}_{k|k-1}^j, u_k^j)$  generates the predicted ideal measurement set based on  $\hat{X}_{k|k-1}^j$  and  $u_k^j$ . Finally, the constraint  $\|u_k^i - u_k^j\| > d_{\min}, \forall i \neq j$  does not allow any two agents to get closer than  $d_{\min}$  from each other. This is by design since in this work there is no benefit of two agents to search the same area or perform tracking of the same target. We should note here that the optimization of Eqn. (17) partitions the agents into two groups i.e. search or track and finds the joint control actions so that the joint search-and-track cost is minimum. Finally, each agent  $j$  uses the multi-Bernoulli recursion Eqn. (6) - (10) to compute posterior multi-object density at time  $k$  given by  $f_k^j(X_k|Z_{1:k}) = \{(r_k^{j,i}, p_k^{j,i}(\cdot))\}_{i=1}^{v_k^j}$  from where the estimated number of targets  $\hat{n}_k^j$  and the multi-target state  $\hat{X}_k^j$  can be computed (see VI-A). The total estimated number of targets at time  $k$  in the surveillance area is then computed as  $\hat{n}_k = \sum_j \hat{n}_k^j$  and the multi-target state which accounts for all targets in the area as  $\hat{X}_k = \bigcup_j \hat{X}_k^j$ .

### A. Multi-Agent Tracking Objective

The tracking objective derived hereafter allows the proposed system to determine which controls to assign to each agent in order to optimize the tracking performance. Lets assume that at time  $k$  the multi-object target distribution estimated by agent  $j$  is approximated by a multi-Bernoulli pdf  $f_k^j = \{(r_k^{j,i}, p_k^{j,i}(x))\}_{i=1}^{v_k^j}$ . From this distribution the expected a posteriori (EAP) estimate of number of targets  $\hat{n}_k^j$  can be computed as  $\hat{n}_k^j = \text{round}(\sum_i r_k^{j,i})$  and the multi-target state  $\hat{X}_k^j$  can then be computed by taking the  $\hat{n}_k^j$  components with the highest probabilities of existence and extracting the single target states (i.e.  $\hat{x}_k$ ) from their individual posterior densities as  $\hat{X}_k^j = \bigcup_{i=1}^{\hat{n}_k^j} \{\hat{x}_k = \arg \max p_k^{j,i}(x_k)\}$ .

Observe that the sensor measurements  $Z_k^j$  play a crucial role in the measurement update step i.e. Eqn. (9)-(10) for the computation of the posterior distribution. In addition, the

quality of the received measurements depend on the state  $s_k^j$  (i.e. the control action taken) of the agent through the functions  $p_D(x_k^j, s_k^j)$  and  $g_k(z_k^j | x_k^j, s_k^j)$ . Since the state  $s_k^j$  of an agent is assumed to be controllable via Eqn. (12), we come to the following conclusion: the control action affects the received measurements which in turn affects the estimate of the target state. Thus in order to increase the tracking accuracy it suffices to find an optimized control decision.

Let us denote the objective function we wish to minimize by  $\xi_{\text{track}}^j(u_k^j, Z_k^j)$ ,  $u_k^j \in \mathbb{U}_k^j$  if we were going to apply at time  $k$  the control action  $u_k^j$  and observe the measurement set  $Z_k^j$ . Notice that the objective function we wish to minimize depends on an unknown future measurement set, i.e. the measurement set will be receive at time  $k$  only after applying the control  $u_k^j$ . A common approach around this problem (as discussed in [24], [25]) is to take the statistical expectation of the objective function with respect to all possible values of future measurements i.e.  $\mathbb{E}[\xi_{\text{track}}^j(u_k^j, Z_k^j)]$  thus the optimization problem becomes  $u_k^{j*} = \arg \min \mathbb{E}[\xi_{\text{track}}^j(u_k^j, Z_k^j)]$ . Computing this expectation however, is computationally intensive because it requires the generation of an ensemble of measurement sets (pseudo-measurements)  $Z_k^j$  for each hypothesized control action. An alternative and computationally cheaper approach which we adopt in the paper uses the predicted ideal measurement set (PIMS) [26] i.e. the noise-free clutter-free measurement set which is most likely to be obtained when a particular control action is applied, denoted  $Z_{k|k-1}$  in this paper. Thus the control problem now becomes:

$$u_k^{j*} = \arg \min_{u_k^j \in \mathbb{U}_k^j} \xi_{\text{track}}^j(u_k^j, Z_{k|k-1}^j) \quad (18)$$

The above optimization can now be solved as follows: Using the posterior distribution  $f_{k-1}^j$  at time  $k-1$  we first compute the predictive density  $f_{k|k-1}^j = \{(r_{k|k-1}^{ji}(u_{k-1}^j), p_{k|k-1}^{ji}(\cdot, u_{k-1}^j))\}_{i=1}^{v_{k|k-1}^j}$  which is obtained without yet performing any control actions. The above notation is to explicitly show the dependence on the control action that was taken for the time-step  $k-1$ . Next we perform a pre-estimation step where the number of targets  $\hat{n}_{k|k-1}^j$  is first estimated from the predictive density by counting the Bernoulli components which exhibit a probability of existence  $r_{k|k-1}^{ji} > 0.5, \forall i \in \{1 \dots v_{k|k-1}^j\}$ . The states of those targets are then extracted from their spatial distributions to obtain  $\hat{X}_{k|k-1}^j$ . Then for a given hypothesized control action  $u_k^j \in \mathbb{U}_k^j$  the corresponding PIMS  $Z_{k|k-1, u_k^j}^j$  is generated as:

$$Z_{k|k-1, u_k^j}^j = Z_{k|k-1, u_k^j}^j \cup \{\arg \max_z g_k(z | \hat{x}_{k|k-1}^j, u_k^j)\}, \forall \hat{x}_{k|k-1}^j \in \hat{X}_{k|k-1}^j \quad (19)$$

Now for each pair  $(u_k^j, Z_{k|k-1, u_k^j}^j)$  a pseudo-update is performed using Eqn. (7) - (10) to compute the (pseudo) posterior density  $\hat{f}_k^j(X_k^j | Z_{k|k-1, u_k^j}^j, u_k^j)$  from where the number of targets is estimated. We consider the variance of this estimate as a measure of uncertainty which should be minimized. More specifically let  $\hat{f}_k^j =$

$\{(r_k^{ji}(u_k^j, Z_{k|k-1, u_k^j}^j), p_k^{ji}(\cdot, u_k^j, Z_{k|k-1, u_k^j}^j))\}_{i=1}^{v_{k, u_k^j}^j}$  then the expected number of targets  $\hat{n}_k^j$  and its variance  $\sigma_k^j$  is given by:

$$\hat{n}_k^j(u_k^j, Z_{k|k-1, u_k^j}^j) = \sum_{i=1}^{v_{k, u_k^j}^j} r_k^{ji}(u_k^j, Z_{k|k-1, u_k^j}^j) \quad (20)$$

$$\sigma_k^j(u_k^j, Z_{k|k-1, u_k^j}^j) = \sum_{i=1}^{v_{k, u_k^j}^j} r_k^{ji}(u_k^j, Z_{k|k-1, u_k^j}^j) \times (1 - r_k^{ji}(u_k^j, Z_{k|k-1, u_k^j}^j)) \quad (21)$$

In essence, by minimizing the cardinality variance  $\sigma_k^j$ , the agent  $j$  chooses the control action which results in the best estimate regarding the number of targets. This variance is maximized when  $r_k^{ji}(u_k^j, Z_{k|k-1, u_k^j}^j) = 0.5, \forall i$  thus we can define:

$$\xi_{\text{track}}^j(u_k^j, Z_{k|k-1, u_k^j}^j) = \begin{cases} \frac{4\sigma_k^j}{v_k^j} & , \text{ if } \hat{n}_k^j(u_k^j, Z_{k|k-1, u_k^j}^j) \geq 1 \\ 1 & , \text{ otherwise} \end{cases} \quad (22)$$

where we have dropped the dependence on the control actions on the right hand side for notational clarity. When the number of estimated targets is less than 1 i.e.  $\hat{n}_k^j < 1$  the tracking score attains its maximum value. Finally we define the multi-agent tracking objective function to be minimized as  $\sum_{j \in \alpha} \xi_{\text{track}}^j(u_k^j, Z_{k|k-1, u_k^j}^j)$  where  $\alpha \subseteq S$  denotes the set of agents in *tracking* mode.

## B. Multi-Agent Searching Objective

In the previous sub-section we have defined the objective function for the task of tracking for single and multiple agents. In this section we are going to do the same for the task of searching. Let us define the *search value* at time  $k$  for specific location  $p = (p_x, p_y) \in \mathcal{A}$  for the agent  $j$  just after control action  $u_k^j \in \mathbb{U}_k^j$  has been applied as follows:

$$\xi^j(p, u_k^j) = 1 - p_D(\tilde{p}, u_k^j) \quad (23)$$

where the function  $p_D(\tilde{p}, u_k^j)$  is the agent's sensing model and  $\tilde{p} = [p_x, 0, p_y, 0]$  constructs a location vector compatible with Eqn. (13). The idea of defining the search value as the complement of the probability of detection reflects the fact that locations with low probability of detection, should have high value for searching. As a consequence distant locations with respect to the agent's location appear to have increased search value i.e. these areas are worth exploring in order to find new targets. Assuming pairwise detection independence between all agents and all targets, the search value at location  $p$  when accounting for a set of agents  $\alpha \subseteq S$  is given by:

$$\xi^\alpha(p, \{u_k^j : j \in \alpha\}) = \prod_{j \in \alpha} \xi^j(p, u_k^j) \quad (24)$$

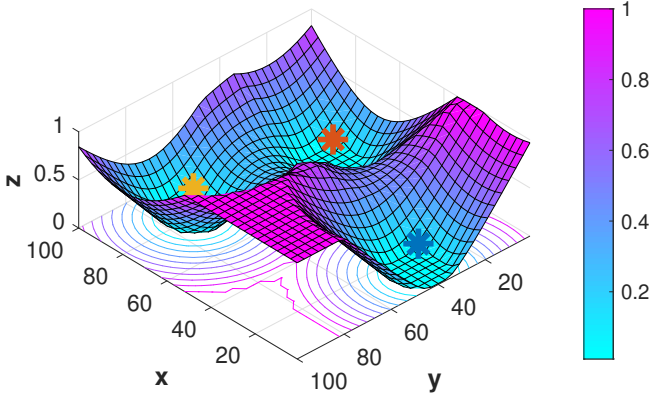


Fig. 2. The figure shows an illustrative example of the search value i.e. Eqn. (24) over a surveillance area of size 100m by 100m for 3 agents. The position of agents are marked with \*.

where  $u_k^j \in \mathbb{U}_k^j$ . The total search value over the surveillance area, i.e. the multi-agent search objective function is now given by:

$$\xi_{\text{search}}^\alpha(\{u_k^j : j \in \alpha\}) = \frac{1}{A} \int_{\mathcal{A}} \xi^\alpha(p, \{u_k^j : j \in \alpha\}) dp \quad (25)$$

where  $\mathcal{A} \subset \mathbb{R}^2$  and  $A$  is the total area of the 2D surveillance region.

### C. Implementation

The optimization problem ( $P_1$ ) in Eqn. (17) requires the enumeration of all possible combinations of the the power set  $\mathbb{P}(S)$  and joint control actions between agents, in order to select the optimal modes of operation and controls which result in the overall minimization of the search-and-track objective. Clearly, this search space grows exponentially with the number of agents and the number of controls, resulting in a very hard combinatorial problem. Instead of solving this problem directly, we use the intuition behind ( $P_1$ ) to derive a heuristic algorithm that can solve instances of the problem in real time. To achieve this we approximate the mode selection as follows: At each time-step  $k$  an agent  $j$  can switch to tracking mode when  $\sum_i r_k^{ji}(u, Z) \geq 1$  i.e. an agent switches to tracking mode only when the expected number of targets inside its sensing range is at least 1. Otherwise, the agent is in searching mode. In essence, this approximation helps us reduce the search space by taking into account information regarding the existence of targets in the area. The agents start by default in search mode and take joint mobility control decisions which minimize the search objective expressed in Eqn. (25). Once an agent detects the presence of targets it switches to tracking mode and the control actions are due to the tracking objective function. An agent  $j$  can switch back to search mode when  $\sum_i r_k^{ji}(u, Z) < 1$ . With the above heuristics however the system cannot attain the optimal behavior provided by the joint objective expressed in Eqn. (17). For instance, an agent in tracking mode does not have the potential to switch to search mode while tracking targets. However, this approximation enables simplified switching that can be followed logically by the agents and which performs reasonably well in many scenarios.

## VII. EVALUATION

### A. Simulation Setup

In order to evaluate the performance of the proposed approach we have conducted several numerical experiments involving various number of agents and targets for a variety of scenarios. In each experiment we compare the proposed approach with the ground-truth either qualitatively or quantitatively. In this section we include a representative sample of the conducted numerical simulations.

More specifically, we assume that the targets maneuver in an area of 100m  $\times$  100m and that the single target state at time  $k$  is described by  $x_k = [x, \dot{x}, y, \dot{y}]^T$  i.e. position and velocity components in  $xy$ -direction. For the target dynamics i.e. Eqn. (11) we assume that the target motion is piecewise linear according to the near constant velocity model with the process noise being Gaussian. Thus the single target transitional density is given by  $\pi(x_k|x_{k-1}) = \mathcal{N}(x_k; Fx_{k-1}, Q)$  where:

$$F = \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad Q = \begin{bmatrix} T/3 & T/2 & 0 & 0 \\ T/2 & T & 0 & 0 \\ 0 & 0 & T/3 & T/2 \\ 0 & 0 & T/2 & T \end{bmatrix}$$

with sampling interval  $T = 1$ s. The target survival probability from time  $k-1$  to time  $k$  is constant  $p_{s,k}(x_{k-1}) = 0.99$  and does not depend on the target's state. Once an agent detects a target it receives bearing and range measurements thus the measurement model of Eqn. (14) is given by:

$$h_k(x_k, s_k) = \left[ \arctan\left(\frac{s_y - y}{s_x - x}\right), \|s_k - Hx_k\|_2 \right]$$

where  $H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$ . The single target likelihood function is then given by  $g(z_k|x_k, s_k) = \mathcal{N}(z_k; h_k(x_k, s_k), \Sigma^T \Sigma)$  and  $\Sigma$  is defined as  $\Sigma = \text{diag}(\sigma_\phi, \sigma_\zeta)$ . The standard deviations ( $\sigma_\phi, \sigma_\zeta$ ) are range dependent and given by:

$$\begin{aligned} \sigma_\phi &= \phi_0 + \beta_\phi \|s_k - Hx_k\|_2 \\ \sigma_\zeta &= \zeta_0 + \beta_\zeta \|s_k - Hx_k\|_2^2 \end{aligned}$$

with  $\phi_0 = \pi/180$ rad,  $\beta_\phi = 10^{-5}$ rad/m,  $\zeta_0 = 1$ m and  $\beta_\zeta = 3 \times 10^{-3}$ m $^{-1}$ . Moreover the agent receives spurious measurements (i.e. clutter) with fixed Poisson rate  $\lambda_k = 5$  uniformly distributed over the measurement space. The agent's sensing model parameters take the following values  $p_D^{\max} = 0.99$ ,  $\eta = 0.003$  and  $R_0 = 10$ m. The agent's dynamical model has radial displacement  $\Delta_R = 2$ m,  $N_R = 2$  and  $N_\theta = 8$  which gives a total of 17 control actions, including the initial position of the agent. Finally, in order to handle the non-linear measurement model we have implemented a Sequential Monte Carlo (SMC) version [22] of the multi-Bernoulli filter.

### B. Results

**Representative search and track scenario:** We begin the evaluation of the proposed technique with a representative search-and-track scenario with 3 agents and 3 targets which takes place during 100 time-steps. This scenario is depicted in Fig. 3 during 4 consecutive time periods i.e.  $k = 1..17$ ,

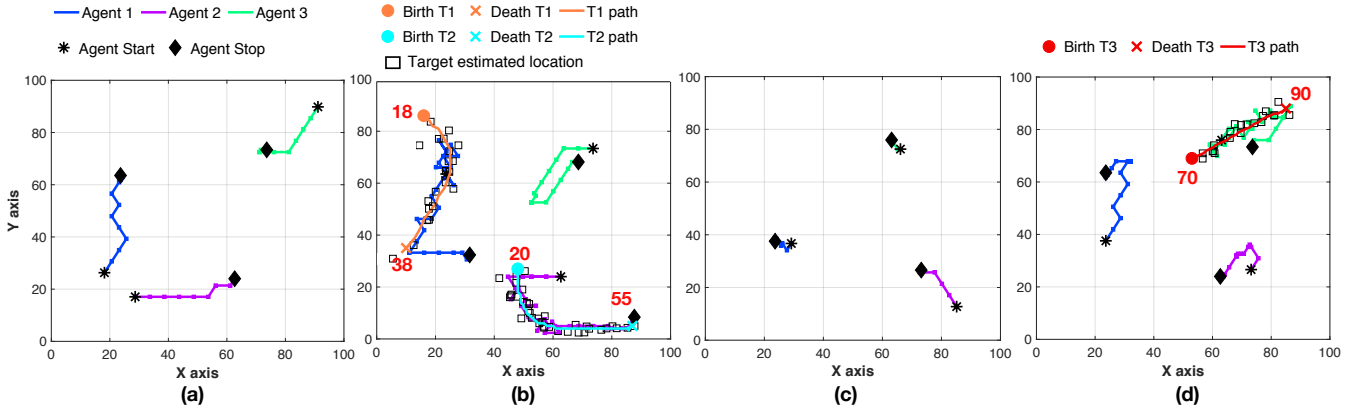


Fig. 3. The figure shows the maneuvers of 3 agents for the task of search-and-track in a simulated scenario during 4 time periods i.e. (a) - (d). Agent start and stop positions are marked with  $\star$  and  $\diamond$ , respectively. Target birth and death positions are marked with  $\bullet$  and  $\times$  respectively and their corresponding birth/death times are shown in red.

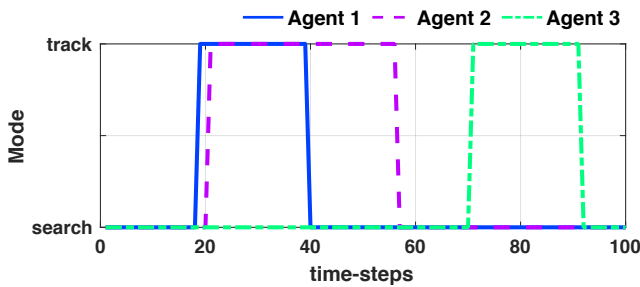


Fig. 4. The figure shows the mode of operation i.e. search or track for the 3 agents of the simulated scenario shown in Fig. 3

$k = 18..56$ ,  $k = 57..69$  and  $k = 70..100$ . In this scenario, 3 agents enter at  $k = 1$  the surveillance area of size  $100\text{m} \times 100\text{m}$  at the locations marked with  $\star$  in Fig. 3a with coordinates  $[18, 26]$ ,  $[28, 17]$  and  $[90, 89]$  for agents 1, 2 and 3, respectively. The target birth/death times are  $k = 18/38$ ,  $k = 20/55$  and  $k = 70/90$  for targets 1, 2 and 3, respectively. The target birth locations (marked with  $\bullet$ ) are  $[16, 86]$ ,  $[48, 27]$  and  $[53, 69]$  for targets 1, 2 and 3 respectively and their corresponding death locations (marked with  $\times$ ) are  $[10, 35]$ ,  $[87, 5]$  and  $[85, 88]$  as shown in Fig. 3. The trajectories of the agents and the targets during this experiment are shown in Fig. 3a - Fig. 3d for periods  $k = 1..17$ ,  $k = 18..56$ ,  $k = 57..69$  and  $k = 70..100$  respectively. At each time-step the proposed system partitions the agents into the two modes (i.e. *search* or *track*) and selects the control actions for all agents that optimize the overall search-and-track objective as explained in the previous sections.

Figure 3a shows the trajectories of the 3 agents during period  $k = 1..17$ . During this period the number of estimated targets is 0 so the system assigns all agents to search mode for the whole period. The objective now is to increase coverage by optimizing the search objective i.e. Eqn. 24. As we can observe from the figure agent 1 and agent 2 start at  $k = 1$  to move away from each other in order to decrease the overlap in their sensing ranges. In particular, the locations of the 3 agents at the end of this period are  $[23, 63]$ ,  $[62, 23]$  and  $[73, 73]$  respectively. This formation (i.e. agent positions)

increases the area coverage since it reduces the sensing range overlap. The modes of operation for the 3 agents for this scenario is shown more clearly in Fig. 4.

Now, at time  $k = 18$  target 1 appears in the area as shown in Fig. 3b which is being detected by agent 1 at the next time-step  $k = 19$ . During time-steps  $k = 19..23$  agent 1 switches to tracking mode and moves towards target 1 to intercept as shown in the figure. Between time-steps  $k = 24..38$  agent 1 tracks target 1. At the same time target 2 appears in the surveillance region at time-step  $k = 20$ . This target is subsequently detected and tracked by agent 2. Now, during time-steps  $k = 21..38$  agents 1 and 2 are in track mode which leaves only agent 3 in search mode. For this reason agent 3 moves toward the center of the surveillance area to increase coverage. However, at time-step  $k = 39$  target 1 dies and thus agent 1 switches to search mode as well. At this point agent 1 and agent 3 are in search mode while agent 2 is in track mode. As we can observe from the figure at  $k = 39$  agent 3 moves away from the center of the surveillance region in order to jointly optimize coverage along with agent 1. Meanwhile agent 2 tracks target 2 until  $k = 55$  as shown in the figure.

Next, at  $k = 56$  target 2 dies, which makes agent 2 to switch to search mode. Thus during the 3rd period i.e.  $k = 57..69$  all agents are in search mode. From Fig. 3c we can see that agents 1 and 3 make slight adjustments in their positions whereas agent 2 moves away from the edge of the surveillance region in order to jointly optimize with the other agents the area coverage.

Finally, during the 4th period at  $k = 70$  target 3 appears which is subsequently detected and tracked by agent 3 during time-steps  $k = 71..90$  as is shown in Fig. 3d. Meanwhile, agents 1 and 2 remain in search mode and move towards the diagonal of the square surveillance area which results in maximum coverage. To conclude the experiment, at  $k = 91$  agent 3 switched back to search mode after target 3 dies and all agents optimize the area coverage as is shown in the figure.

**Multi-agent searching:** The next set of experiments aims to investigate in more detail the behavior of our system when all agents are in search mode. More specifically, the



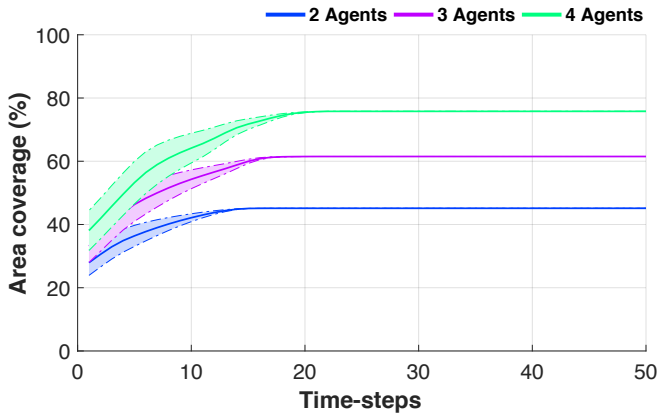


Fig. 5. The figure shows the percentage of covered surveillance area over time for different number of agents.

desired behavior is for the agents to jointly find the best configuration i.e. select the appropriate controls such that the total search value of the surveillance area is minimized or the area coverage is maximized. To investigate this we have performed the following procedure. For a given number of agents we have conducted 50 Monte-Carlo (MC) trials where the agents are spawned from random locations (uniformly distributed) inside the surveillance area of size  $100\text{m} \times 100\text{m}$ . Then we let our system to run for 50 time steps and we monitor the percentage of covered area. Our intuition is that the search value should decrease over time since the agents would try to move to locations where their sensing range overlap is minimized which in turn will result in an increase in the area coverage over time. In other words we would like to have the agents spread as much as possible inside the surveillance region to increase coverage. Figure 5 shows the average area coverage over time for 2, 3 and 4 agents (with  $p_D^{\max} = 0.99$ ,  $\eta = 0.03$  and  $R_0 = 10\text{m}$ ). As we can observe from this figure the search value is decreasing over time which is evident by the increase of the covered area over time. The agents are trying to cover as much area as possible and minimize the overlap in their sensing ranges. It is also clear that as the number of deployed agents inside the surveillance area increases the covered area increases as well.

In order to investigate further the experiment conducted above we have been monitoring the controls assigned to the agents during the search. At the end of the experiment the final agent positions for the 3 configurations i.e. 2, 3 and 4 agents are shown in Fig. 6. Interestingly, we can see that the agents have converged to a formation which results in maximum coverage. For instance from Fig. 6a we see that when the number of agents is 2 the configuration which results in the maximum coverage is the diagonal of the square surveillance area. On the other hand when we have 3 and 4 agents, the optimal configuration forms a triangle (Fig. 6b) and a square respectively (Fig. 6c). This behavior however depends on the agent's sensing model (i.e. Eqn. (13)). It is not unlikely to have other formations given a different sensing model. This is something quite interesting that we will investigate in the future. In addition, it would be very interesting to see that happens when not all agents possess

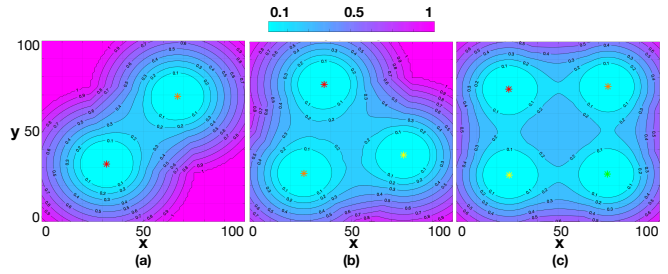


Fig. 6. The figure shows the search value over the whole surveillance area for different number of agents, i.e. (a) 2 agents, (b) 3 agents and (c) 4 agents. As we can observe the agents take positions which result in the maximum area coverage (the figure shows the last time-step of a 50 time-step simulation experiment).

the same sensing model. This is also left for future work.

We should note here that in the previous experiment no targets were initiated during the simulation. This was done in order to investigate the behavior of the system in the search mode. Additionally, the number of targets and their initial locations are not known a priori in this work, instead there are being estimated at run-time. Thus during the previous experiment, the proposed system correctly assigned all agents in search mode since no targets were present.

**Search and track performance:** Finally, our last experiment was conducted in order to demonstrate the search-and-track performance of the proposed system. For this experiment we have used the optimal sub-pattern assignment (OSPA) metric [27] which is widely used to measure the accuracy of multi-object filters as an indicator of the search-and-track performance. In this experiment which takes place during 80 time-steps, 8 targets are spawned at the following locations  $[1, 97]$ ,  $[4, 89]$ ,  $[3, 4]$ ,  $[7, 9]$ ,  $[97, 4]$ ,  $[95, 4]$ ,  $[95, 94]$  and  $[97, 85]$  for targets 1 to 8 respectively and die at  $[43, 67]$ ,  $[39, 57]$ ,  $[50, 41]$ ,  $[42, 46]$ ,  $[62, 49]$ ,  $[54, 36]$ ,  $[55, 66]$  and  $[61, 55]$  as shown in Fig. 7a. The target birth/death times are  $1/80$  for all targets. With this setup 2, 3 and 4 agents are spawned simultaneously at time-step  $k = 1$  from the center of the surveillance area (i.e. the agents are spawned randomly inside the bounding box  $[40..60] \times [40..60]$ ) and the system runs for 80 time-steps. We perform 30 MC trials for each configuration (i.e. 2, 3 and 4 agents) and we measure the OSPA error. Figure 7b shows the average OSPA error (with parameters  $c = 100, p = 2$ ) for the 3 configurations.

The first observation to make is with respect to the searching behavior of our system. Since all agents start from about the center of the surveillance area whereas all targets start from the perimeter there is a period of searching during the first time-steps of the experiment. Indeed, during the first time-steps all agents are in search mode and try to maximize the area coverage. This is shown in Fig. 7b from the OSPA error which takes the maximum value during the beginning of the test for all agent configurations. When the number of targets being tracked is 0 the OSPA error reaches its maximum value which is  $100\text{m}$  in this case. Moreover, we observe that with 3 and 4 agents the search is faster compared to the 2 agent case. With 2 agents the first target(s) seems to be detected at around  $k = 15$ , whereas 3 and 4 agents detect the first targets around  $k = 5$ . This is

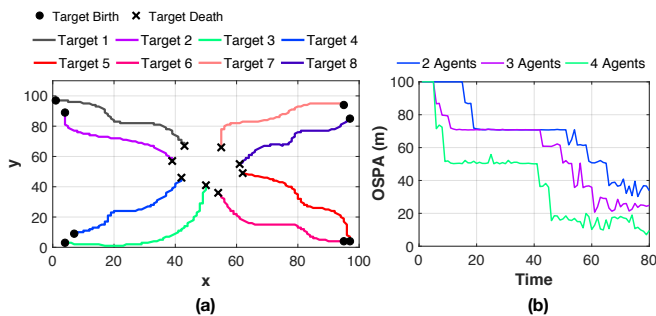


Fig. 7. The figure demonstrates the search-and-track performance of the proposed technique through the OSPA error.

reasonable and verifies our intuition that as the number of agents increases the search performance should increase as well. The second observation from this experiment is with respect to the tracking capacity of the proposed system. As we can observe from the OSPA error as the number of agents increases the tracking accuracy increases as well. Since each agent is capable of tracking multiple targets the overall tracking accuracy can be increased by positioning the agents in locations which will result in maximum target coverage. This is achieved in this work by minimizing the sensing overlap between agents as is shown in this experiment.

### VIII. CONCLUSION AND FUTURE WORK

In this paper we have studied the problem of searching and tracking of an unknown and time-varying number of targets with a team of mobile agents. We have presented a novel probabilistic framework based on random finite sets that accounts for the uncertainty in the number of targets, target dynamics and measurement models. The proposed approach allows the agents to optimize the search-and-track objective by dynamically switching their mode of operation at each time-step during the mission according to the acquired information. Finally, we have demonstrated the performance of the proposed decision and control algorithm through extensive numerical simulations. Future work will investigate the extension of this approach in 3D environments and study the problem of searching and tracking in highly dynamic environments with obstacles and occlusions.

### ACKNOWLEDGMENTS

This work is supported by the European Union Civil Protection under grant agreement No 783299 (SWIFTERS), by the European Union's Horizon 2020 research and innovation programme under grant agreement No 739551 (KIOS CoE) and from the Republic of Cyprus through the Directorate General for European Programmes, Coordination and Development.

### REFERENCES

- [1] D. Floreano and R. J. Wood, "Science, technology and the future of small autonomous drones," *Nature*, vol. 521, pp. 460 EP –, 05 2015.
- [2] Y. Rizk, M. Awad, and E. W. Tunstel, "Decision making in multiagent systems: A survey," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 10, no. 3, pp. 514–529, Sep. 2018.
- [3] R. P. S. Mahler, *Advances in Statistical Multisource-Multitarget Information Fusion*. Norwood, MA, USA: Artech House, Inc., 2014.

- [4] C. Kyrkou, S. Timotheou, P. Kolios, T. Theocharides, and C. Panayiotou, "Drones: Augmenting our quality of life," *IEEE Potentials*, vol. 38, no. 1, pp. 30–36, Jan 2019.
- [5] V. Mersheeva and G. Friedrich, "Multi-uav monitoring with priorities and limited energy resources," 2015.
- [6] A. Khan, E. Yanmaz, and B. Rinner, "Information exchange and decision making in micro aerial vehicle networks for cooperative search," *IEEE Transactions on Control of Network Systems*, vol. 2, no. 4, pp. 335–347, Dec 2015.
- [7] G. Avellar, G. Pereira, L. Pimenta, and P. Iscold, "Multi-uav routing for area coverage and remote sensing with minimum time," *Sensors*, vol. 15, no. 11, p. 27783–27803, Nov 2015.
- [8] C. Li, L. Dong, and S. K. Nguang, "Cooperative control of multi-agent systems with variable number of tracking agents," *IET Control Theory Applications*, vol. 11, no. 12, pp. 1922–1927, 2017.
- [9] R. Dutta, L. Sun, and D. Pack, "A decentralized formation and network connectivity tracking controller for multiple unmanned systems," *IEEE Transactions on Control Systems Technology*, vol. PP, no. 99, pp. 1–8, 2017.
- [10] S. Shahrapour and A. Jadbabaie, "An online optimization approach for multi-agent tracking of dynamic parameters in the presence of adversarial noise," in *2017 American Control Conference (ACC)*, May 2017, pp. 3306–3311.
- [11] F. Bourgault, T. Furukawa, and H. F. Durrant-Whyte, "Optimal search for a lost target in a bayesian world," in *Field and Service Robotics: Recent Advances in Research and Applications*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 209–222.
- [12] C. Liu and J. K. Hedrick, "Model predictive control-based target search and tracking using autonomous mobile robot with limited sensing domain," in *2017 American Control Conference (ACC)*, May 2017, pp. 2937–2942.
- [13] G. M. Hoffmann, S. L. Wasl, and C. J. Tomlin, "Distributed cooperative search using information-theoretic costs for particle filters with quadrotor applications," in *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, 2006, pp. 21–24.
- [14] G. M. Hoffmann and C. J. Tomlin, "Mobile sensor network control using mutual information methods and particle filters," *IEEE Transactions on Automatic Control*, vol. 55, no. 1, pp. 32–47, Jan 2010.
- [15] R. R. Pitre, X. R. Li, and R. Delbalzo, "Uav route planning for joint search and track missions: An information-value approach," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 3, pp. 2551–2565, JULY 2012.
- [16] C. K. Peterson, "Dynamic grouping of cooperating vehicles using a receding horizon controller for ground target search and track missions," in *2017 IEEE Conference on Control Technology and Applications (CCTA)*, Aug 2017, pp. 1855–1860.
- [17] Z. Chen, "Bayesian filtering: From kalman filters to particle filters, and beyond," *Technical report. Hamilton, ON, Canada: Adaptive Syst. Lab., McMaster University*, 2003.
- [18] D. Simon, *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*. Wiley-Interscience, 2006.
- [19] R. Mahler, "Statistics 102 for multisource-multitarget detection and tracking," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 3, pp. 376–389, June 2013.
- [20] R. P. S. Mahler, "Statistics 101 for multisensor, multitarget data fusion," *IEEE Aerospace and Electronic Systems Magazine*, vol. 19, no. 1, pp. 53–64, Jan 2004.
- [21] —, *Statistical Multisource-Multitarget Information Fusion*. Norwood, MA, USA: Artech House, Inc., 2007.
- [22] B. T. Vo, B. N. Vo, and A. Cantoni, "The cardinality balanced multi-target multi-bernoulli filter and its implementations," *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 409–423, Feb 2009.
- [23] B. Ristic, B. T. Vo, B. N. Vo, and A. Farina, "A tutorial on bernoulli filters: Theory, implementation and applications," *IEEE Transactions on Signal Processing*, vol. 61, no. 13, pp. 3406–3430, July 2013.
- [24] B. Ristic, B. N. Vo, and D. Clark, "A note on the reward function for phd filters with sensor control," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, no. 2, pp. 1521–1529, April 2011.
- [25] M. Beard, B. T. Vo, B. N. Vo, and S. Arulampalam, "Sensor control for multi-target tracking using cauchy-schwarz divergence," in *2015 18th International Conference on Information Fusion (Fusion)*, July 2015, pp. 937–944.
- [26] R. Mahler, "Multitarget sensor management of dispersed mobile sensors," *Series on Computers and Operations Research, World Scientific (2004)*, 2004.
- [27] 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, pp. 3447–3457, Aug 2008.