Application and Comparison of Multi-Robot Exploration Methods for Radioactive Source Localization

Albert Chan, Long Kiu Chung, Yixian Li, Aaron Wong

***Abstract*— In situations involving routine radiological operations, lost radioactive sources, decommissioning and cleanup of nuclear facilities, and radiological terrorist events, efficient and accurate source localization algorithms for unmanned aerial vehicles (UAV) mounted with radiation detectors are important for mitigating potential risks of human exposure to radioactive materials. While most literature focuses on localization methods for a single UAV, there has also been an increase of interest in applying these techniques for multiple agents, since the resulting policy of distributed, multi-robot control is scalable to a variety of situations, robust to individual robot failure, and modular. Multi robot exploration also allows for faster exploration compared to single robot methods, allowing for the radiation source to be localized efficiently. In this paper, we implemented and compared a frontier-based method and an information-theoretic control policy for a network of UAVs applied to the radioactive source localization problem to determine which is best suited to the particular problem of radioactive source localization. We found that the frontier-based algorithm more effectively finds the sources and efficiently explores the map. While the information-theoretic algorithm did not perform as well, it shows promise for further investigation.**

I. Introduction

Robot exploration is the task of efficiently controlling the robots to map the environment accurately. Due to its central role in any mobile robotics application, the field of single-agent autonomous exploration has been a very active field of research. Over the past two decades, the field has seen an increase in interest in applying such exploration techniques to systems with multiple robots as well, since the resulting distributed multi-robot control policy is scalable to a variety of networks, robust to individual robot failure, modular, and usually faster compared to single-robot methods. Existing methods for multi-robot exploration can be generally divided into two categories: *frontier-based methods* and *information-theoretic-based methods*.

A. Frontier-Based Methods

Frontier-based approaches are based on the concept of occupancy grid maps and frontiers. An occupancy grid divides the environment into cells, with each cell of the map representing the probability that the represented area is an obstacle. The goal of frontier-based methods usually centers around estimating the true occupancy grid map (*true grids*) by filling up an estimated occupancy grid map (*evidence grids*) by exploring the environment with multiple robots. With no prior knowledge of the true grids, the cells in the evidence grids can be initialized with an occupancy probability of 0.5. Such cells are dubbed as *unknown cells*. The cells are then updated with new occupancy probability values as exploration progresses. Cells with occupancy probabilities less than 0.5 are called *open cells*, which means such cells are safe to travel to. Similarly, cells with occupancy probabilities greater than 0.5 are called *occupied cells*, which represent obstacles such as walls that the robots should avoid. *Frontier cells*, then, are all the known open cells that immediately neighbours unknown cells. This concept is illustrated in Fig. 1.

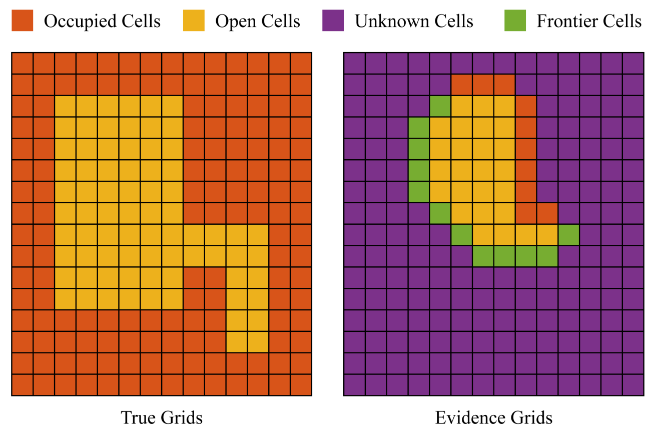


Fig. 1: Illustration of the concept of occupancy grid maps and frontier cells. The example true grids represent a room surrounded by walls. The evidence grids represent information known by the robots.

The central idea in frontier based exploration methods is to move toward the frontier, allowing the robot to recursively sense new regions and build its world map. A cost function will inform which location on the frontier each robot should explore. Research work in this field has centered around designing these heuristic cost functions to intelligently explore the map, trading off factors such as the distance to the frontier against the potential information gain of reaching an unexplored state while avoiding obstacles and respecting robot constraints. This problem is further complicated for multi-robot systems, which need to coordinate exploration trajectories with each other.

The concept of frontiers in a multi-robot exploration context was first proposed by the seminal paper by [1]. However, the paper merely combines the explored maps of each agent and applies the frontier exploration algorithm on the combined map for each agent. This heuristic can result

All authors are with Stanford University, Stanford, CA. Albert Chan, Long Kiu Chung, and Yixian Li are with the Department of Mechanical Engineering. Aaron Wong is with the Department of Aeronautics and Astronautics.

in each agent of the system navigating to the same frontier, which causes a loss in efficiency.

The work by [2] built upon this result by introducing the visibility probability $P(d)$, which is the probability that the robot's sensors will sense an object at distance d. While exploring, the robot will build this probability distribution from collected sensor data, and use it to inform U, the predicted utility from reaching the goal. A low visibility probability will lead to a high utility, incentivizing robots to explore more open areas, and avoid staying close to obstacles, while also disincentivizing the same robot from exploring the same region.

Another drawback of these works is that they still rely on communication with a central planner in order to generate exploration trajectories. This is addressed in future works such as [3] on market-based approaches to multi robot exploration. Market-based approaches are fundamentally similar to frontier-based approaches, as they all have a common utility or cost function to inform how they explore, but the "market economy" provides a different framework for robot coordination such that they do not need to interact with a central planner. In market-based approaches, each robot starts with a list of potential sites on the frontier to explore, while computing costs and utility to reach each of these points. The robot will then communicate with nearby neighbours and "auction" off tasks that the neighbour can perform at lower cost, transferring the task to a different robot. This method bypasses the need for a centralized planner.

Further research of this methodology has also resulted in more complex and detailed heuristics, such as in [4]. This work attempts to use range information to train a linear classifier for semantic place labelling of the sensed environment, enabling the planner to determine between a corridor or a room. A Hidden Markov Model is then used to estimate spacial relationships between locations, significantly improving the exploration time need for large multi-robot systems.

[4] also begins the introduction of Machine-Learning-based (ML) techniques in robot exploration, extracting higher fidelity data and introducing more variables to better predict the utility of a given exploration frontier. Another application of ML in multi robot exploration is to also learn intelligent coordination patterns between robots. This research is captured in [5], who proposes MacDec-MADDRQN, a macro-action-based decentralized multi-agent double deep recurrent Q-net. This work builds on a Decentralized Partially Observable Markov Decision Process (Dec-POMDP) to learn cooperative interactions between robots using a double Q-net framework, and is applied in [6], resulting in significantly higher exploration rates and better overall cooperation compared to traditional utility function based approaches.

B. Information-Theoretic-Based Methods

Information-theoretic approaches to robot exploration are based on the pioneering work in [7], which provides the mathematical structure of quantifying the information contained in observing a random variable. The problem can be formulated as follows: given a system of mobile robots, starting in some initial configuration, and equipped with sensors which provide measurements of a random field representing the state of the environment, control the robots such that the uncertainty in environment state is reduced From an information-theoretic standpoint, this corresponds to reducing the entropy of the belief over the environment state.

The entropy of a probability distribution is defined as the negative expectation of the log-likelihood.Considering the mean conditional entropy in the light of Bayesian belief updating gives an expression of mutual information. Most information-theoretic controllers leverage mutual information, as it is an a priori measurement of potential information gain from an observation.

This problem was first approached as a problem of sensor management, which seeks optimal placement of static sensors to best sense the environment while minimizing sensor effort. [8] uses Bayesian belief updates and mutual information to optimize sensor placement for object detection. This method was extended to decentralized networks of sensors in [9], which uses individual sensor measurements to generate a fused Bayesian belief. The problem is further expanded to the control of multiple dynamic sensors in [10] which develops a controller to increase mutual information achieved by the sensors.

A shortcoming of mutual-information is that every robot needs knowledge of measurements and beliefs across the entire network of robots, limiting the decentralized nature of mutual-information controllers. To remedy this, [11] uses a consensus algorithm to disseminate the belief state, allowing distributed hypothesis testing. The use of consensus is extended in [12] to compute mutual information and generate a distributed control law based on approximations of the mutual information gradient.

[13] applied information-theoretic control to a realistic multi-robot exploration situation. In an unknown discretized environment, there were probabilistic hazard spots and event spots. A group of robots were tasked to circumvent the hazard spots and explore the event spots as soon as possible. By taking a gradient step derived from the mutual information, the robots were positioned to maximize the expected information gain. However, computing the gradient of mutual information is not computationally feasible, and either the environment state or the gradient must be approximated in implementation.

II. Problem Formulation

In situations involving routine radiological operations, lost radioactive sources, decommissioning and cleanup of nuclear facilities, and radiological terrorist events, efficient and accurate source localization algorithms for unmanned aerial vehicles (UAV) with radiation detectors are important for mitigating potential risks of human exposure to radioactive materials.

While many single-agent methods have been proposed for such source localization problem, the same cannot be said about multi-agent approaches. As most single-agent approaches are myopic and computationally intractable for multiple agents, extension of these methods to multi-agent scenarios are not straightforward either [14].

Our project aims to apply and compare between the effectiveness of *frontier-based method* and *information-theoretic-based method* on the source localization problem.

A. Problem Statement

For a two-dimensional (2D) open field with multiple obstacles and n radioactive sources, located at $(x_{0,1}, y_{0,1}), \cdots, (x_{0,n}, y_{0,n})$, the radiation reading $f(x,y)$ at point (x,y) in the field is given by

$$f(x,y) = \sum_{i=1}^{n} \frac{\lambda_i}{(x - x_{0,i})^2 + (y - y_{0,i})^2} \tag{1}$$

where λ_i is the proportional constant of the i^{th} radioactive source. This assumes no attenuation or reflection of the radiation off the ground or the obstacles [15].

Given a fully connected network of k robots equipped with range sensors of radius r_{max} and radiation detectors, we deploy them to such two-dimensional field without knowledge of n beforehand. The robots never fail, can fully traverse the field, and are controlled from a central source which has full access to the history of robot states, range sensor measurements, and radiation measurements as obtained in equation (1). The goal of the problem is to find the location of all radiation sources in the field $(x_{0,1}, y_{0,1}), \cdots, (x_{0,n}, y_{0,n})$ as quickly and accurately as possible.

III. Method

A. Frontier-Based Method

We applied a *frontier-based method* on the source localization problem in a *state machine* formulation, with the robots being in either *exploration* mode or *source-tracking* mode. When the radiation reading $f(x_i, y_i)$ of a robot i at position (x_i, y_i) is larger than some threshold γ (e.g. background radiation level, or when the signal-to-noise ratio is high enough), and (x_i, y_i) is far enough from sources that have already been found and other robots in source-tracking mode, robot i would remain in *source-tracking* mode until the corresponding source has been found. Otherwise, all robots defaults to operating in *exploration* mode. Mathematically, if $s = 1$ represents exploration mode, and $s = 0$ represents source-tracking mode, for robot i,

$$\begin{cases} s_i = 0 & \text{if } f(x_i, y_i) > \gamma, \\ & \|[x_i - x_{0,j}, y_i - y_{0,j}]\|_2 > \varepsilon_k \\ & \|[x_i - x_\ell, y_i - y_\ell]\|_2 > \varepsilon_m \\ & \forall \{ j = 1, \cdots, n_k, \\ & s_\ell = 0, \\ & \ell = 1, \cdots, k \\ & \ell \neq i \}, \\ s_i = 1 & \text{otherwise.} \end{cases} \tag{2}$$

where n_k is the number of sources found, ε_k and ε_m are the distance threshold for the found sources and for the other robots in source-tracking mode respectively.

1) Source-Tracking Mode: When a robot enters source-tracking mode, any single-agent source localization method, such as [15], can be utilized. For simplicity, we proposed a very simple gradient ascent method for the robots in this mode to follow.

At each step, we consider the robot i at (x_i, y_i) for four action states up $a_1 = (x_i, y_i + 1)$, down $a_2 = (x_i, y_i - 1)$, left $a_3 = (x_i - 1, y_i)$, and right $a_4 = (x_i + 1, y_i)$. For each action j, we designed a cost function $g(a_j)$, which is calculated as the gradient from the current grid to the action grid. If the radiation measurement at the action grid $f(a_j)$ is not known but the measurement at the opposite action grid $f(a_j^*)$ is known, the cost is calculated as the negative gradient from the current grid to the opposite action grid. Otherwise, the cost for the action defaults to 0. The robot then chooses the action with the least cost until a source is found. The algorithm is detailed mathematically in Algorithm 1.

Algorithm 1: Gradient ascent algorithm for source tracking mode.

while *corresponding source is not found* **do**
  for $j = 1, \cdots, 4$ **do**
    if $j \mod 2 = 0$ **then**
      $a_j^* = a_{j-1}$
    else
      $a_j^* = a_{j+1}$
    if $f(a_j)$ *is known* **then**
      $g(a_j) = f(x_i, y_i) - f(a_j)$
    else if $f(a_j^*)$ *is known* **then**
      $g(a_j) = f(a_j^*) - f(x_i, y_i)$
    else
      $g(a_j) = 0$
  $(x_i, y_i) := \arg\min_{a_j} g(a_j)$;

The source-tracking mode repeats until the corresponding source has been found. For simplicity, we define a found source as a grid with higher reading than all of its neighbouring grid. That is, a source is found at $(x_{0,j}, y_{0,j})$ iff $f(x_{0,j}, y_{0,j}) > f(x_{0,j} + \Delta x, y_{0,j} + \Delta y) \forall \{\Delta x, \Delta y \in \{-1, 0, 1\}, |\Delta x| \neq |\Delta y|\}$.

2) Exploration Mode: In exploration mode, we apply a modified version of the coordinated multi robot exploration method proposed by [16]. This method is a utility-function-based, centralized, sequential planner for exploration. The cost to go V is computed via Dijkstra's shortest path algorithm [17] with a Manhattan Distance heuristic.

Algorithm 2 details the method for calculating the optimal frontier-to-go for each robot.

The $P(d)$ in Algorithm 2, where $d = ||t - t'||$, is computed as:

Algorithm 2: Calculation of optimal frontier-to-go.

Determine set of frontier points;
Compute the cost to go V_t^i for a given robot i to reach a frontier point t;
Set the utility of all frontier cells to 1;
for *each robot i* **do**
  Determine a frontier cell t which satisfy: $t = \arg\max(U_t - \beta V_t^i)$
Reduce Utility of other frontiers t' as follows: $U_{t'} = U_t - \sum_{i=1}^{n} P(||t-t'||)$;

$$P(d) = \begin{cases} 1 - d/r_{max} & \forall\{d \leq r_{max}, \\ & \text{line from } t \text{ to } t' \text{ is unobstructed}\}, \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

At a high level, this algorithm returns the frontier point t with the maximum utility for the minimum cost. The utility of frontier points t' that are within line of sight of t and are within the sensing range are reduced to disincentivize robots from exploring the same frontier. At each step, (x_i, y_i) is updated as the next-step action space via a Dijkstra's shortest path planner.

B. Information-Theoretic-Based Method

We developed a naive approach for multi-robots exploration by utilizing mutual information in a local region. In the following subsections, we will break down the specific details of the approach into measurement model, posterior update with new state representation, action selection from a locally defined mutual information, and the overall algorithm.

1) Measurement Model: To apply information-based control algorithms to the source localization problem, we had to create a measurement model distribution that related the state of the map to the radiation reading that each robot is receiving. Past literature on information-based robot exploration considers sensor models that operate at a range, such as cameras or LIDAR sensors [10], [13], which have well characterized measurement model distributions as a function of relative position. These measurement models thus assume a single measurement influences the state distribution in an area of the map. This is unrealistic in the field of radiation sensing, as the highly ionizing radiation particles have low penetration depth through the air.

We propose leveraging the radiation model of equation (1) to create an abstracted measurement model $\mathcal{Z}$ compatible with an information-theoretic control framework. Consider a robot that takes a measurement y_i of the radiation field at its current location. To account for model error, future measurements can be assumed to be corrupted by noise. A cumulative distribution function over the measurement at a remote location can be generated by considering the probability of the measurement over all possible source locations and power levels.

Theoretically, this could be achieved by assuming the state of the environment is the radiation power level of a potential source at each cell in the map. The cumulative distribution function over measurements can then be calculated from the law of total probability and the artificial noise model. In addition, taking a measurement can be used to update the belief over the state of the environment over the entire map through Bayesian belief updating. While this model is amenable to information-based control, it is infeasible for implementation, because these are infinite-support distributions, and updating the belief would scale exponentially with the size of the search space.

We simplify this model to produce a computationally tractable belief updating scheme. First, we reduce the state at each cell s to a binary value, whether a radiation source exists at that point or not. We also bound the remote states $\mathcal{S}$ that a measurement influences by a finite sensing radius. This is realistic as the distributions of measurements become increasingly uniform the further away the remote point is. In addition, we form a hypothesis test based on pre-computed probability distributions of remote measurements over a variety of radiation models. Unlike before, these distributions are independent of the prior belief, and are used to generate a hypothesis test that defines the measurement model $\mathcal{Z}$.

The new measurement model is a binary, probabilistic measurement drawn from the hypothesis test, described in equation (4). The measurement y is distributed by the false positive and negative rates of the hypothesis test. In addition, since the measurement is a simple function of the subset of states within the sensing region of the robot, the belief over the states in the sensing region can be updated computationally efficiently. This gives a viable measurement model $\mathcal{Z}$ that can be used in information-theoretic exploration.

$$p(y_i = 1 \mid \exists s_i = 1 \in \mathcal{S}) = 1 - \beta \tag{4}$$

$$p(y_i = 0 \mid s_i = 0 \,\forall\, s_i \in \mathcal{S}) = 1 - \alpha \tag{5}$$

2) Prior Transform and Posterior Update: It is important to note that from the method adopted, we only reason about a subset of j cells within the sensing region $\mathcal{S}_i$ from robot i's position, where we would update our posterior on. Since we are not assuming dependency between grid cells and the measurement model is defined over all sensing region instead of each grid cell, we need to transform our prior information from a space of $|\mathcal{S}_i|$ into a space of $2^{|\mathcal{S}_i|}$ to represent every possible state distribution in $\mathcal{S}_i$. Denote the new state representation and new state space as S' and $\mathcal{S}'$.

$$S' = [s_1, s_2, ..., s_j]$$

where $s_j \in \{0, 1\}$. For each S', the prior information is the joint probability of every individual's prior:

$$p(S') = \prod_{s_j \in \mathcal{S}} p(s_j) \tag{6}$$

We can now perform posterior update on the transformed prior given new measurement y_i:

$$p(S'|y_i) = \frac{P(y_i|S')P(S')}{\sum_{S_p \in \mathcal{S}'} P(y_i|S_p)P(S_p)} \tag{7}$$

To update the posterior over the map, we can transform the posterior back to the grid representation for each s_j:

$$p(s_j = 1|y_i) = \sum_{S_p \in \mathcal{S}'} p(S_p)\mathbb{1}(s_j = 1) \tag{8}$$

3) Mutual Information and Action selection: We are selecting an action a_i that takes us to $x_n = x_i + a_i$ that maximizes the mutual information $MI(x_n)$, a measure of how much more information about states within sensing range $S' \in \mathcal{S}'$ we will gain from a new measurement y_n from robot's next position x_n. It is formally defined as:

$$MI(x_n) = H(S'_n) - \overline{H}(S'_n|y_n) \tag{9}$$

where $H(S'_n)$ is a measurement of how uncertain we are about the sensing space from a new position. Mathematically it is defined as:

$$H(S'_n) = \sum_{S_p \in \mathcal{S}'} p(S_p)\log(S_p) \tag{10}$$

where $p(S_p)$ is the transformed prior calculated from equation (6).
$\overline{H}(S'_n|y_n)$ is a measurement of how uncertain we are about the sensing space from a new position given an unknown new measurement from that position. This is defined as:

$$\overline{H}(S'_n|y_n) = \sum_{\hat{y_n} \in \{0,1\}} p(\hat{y_n}) \sum_{S_p \in \mathcal{S}'} p(S_p|\hat{y_n}) \log p(S_p|\hat{y_n}) \tag{11}$$

$$p(\hat{y_n}) = \sum_{S_p \in \mathcal{S}'} p(\hat{y_n}|S_p)p(S_p) \tag{12}$$

where $p(\hat{y_n}|S_p)$ is the sensor model defined in equation (4), and $p(S_p|\hat{y_n})$ is the posterior from equation (8).
By choosing the action from all candidate actions that maximize the mutual information, the robots will naively go to the place where there are more uncertain regions around them.

4) General Algorithm: The general algorithm for the information-based robot exploration strategy we implemented is presented in Algorithm 3. This algorithm sequentially chooses the action for each robot that maximizes the expected information gain of the states over the entire map. Because the robots have a limited sensing region, measurements only affect the belief in that region, so the total expected information gain is equivalent to the mutual information due to the measurement model over the sensing region.

Considering the optimization over the action space $\mathcal{A}$ of each robot individually reduces the optimality of this algorithm, specifically when multiple robots share portions of their sensing regions. However, this reduces the optimization space for k robots from $|\mathcal{A}|^k$ to $k|\mathcal{A}|$. In addition, the joint optimization is not possible for a distributed network of robots, which is a possible further direction for this work.

Algorithm 3: Information-based robot exploration using mutual information

```
Initialize the belief over the environment state with a
  uniform prior p(s_m = 1) = p_0 for all grid cells in
  the map.
for each time step t do
    for each robot i do
        Generate y_i^t ∈ {0,1} from Z of sensing region
          around the robot position x_i^t;
        Convert the belief representation: S → S' (6);
        Perform Bayesian Update of the belief (7);
        Convert the belief representation: S' → S and
          update posterior (8);
        a_i^t = argmax_{a_i ∈ |A|} MI(x_i^t + a_i) (9);
        x_i^{t+1} = x_i^t + a_i^t;
```

In Algorithm 3, we define the action space $\mathcal{A} = \{up, down, left, right\}$ for the simplified quadcopter robots. Mutual information at each candidate next state $x_n = x_i + a_i$ is used to choose the optimal action to take. The prior p_0 of each cell represents the apriori estimate of the density of radiation sources in the map. While the number of radiation sources is not known before exploration, if the density is assumed to be low, as is usually the case, a lower prior leads to better performance.

This algorithm is written in a sequential, centralized manner for ease of implementation on a centralized server computer, but this algorithm could be extended to the distributed case for deployment on a network of independent robots. A consensus algorithm, such as linear consensus protocol or that used in [12] could be used to propagate local belief updates over the network such that the robots agree on the belief at each point in the map.

IV. Results

We performed simulations in MATLAB on a laptop with 4 cores, 8 GB RAM, and no GPU. The simulations were performed in a grid-based 10 m × 10 m environment, with 3 sources located at (1.65 m, 8.35 m), (6.65 m, 3.35 m), and (7.35 m, 8.35 m) with $\lambda_1 = \lambda_2 = \lambda_3 = 150$ mSv m^2 h^{-1}. The drones are equipped with $r_{max} = 0.9$ m perfect range sensors and radiation detectors that report the radiation reading at each step. The action space of the robots are limited to up, down, left, and right, and we assume the methods are centralized and sequential. We performed simulations for 1, 2, 3, 4, 6, and 9 robots starting at (50 m, 50 m) and 2 robots starting on the sides at (4.95 m, 8.35 m) and (4.95 m, 1.75 m) for a 100 × 100 grid and a 30 × 30 grid with frontier-based method, and a 30 × 30 grid with information-theoretic-based method.

A. Frontier-Based Method

For all simulations in frontier-based method, we use $\beta = 2$, $\gamma = 2$ mSv h^{-1}, and $\varepsilon_k = \varepsilon_m = 2r_{max}$. In all of our simulations, our robot is able to efficiently explore the region and localize the sources. Fig. 2 shows the trajectories of each robot in localizing the source for a 100 $\times$ 100 grid, with 4 robots starting at $(50, 50)$.

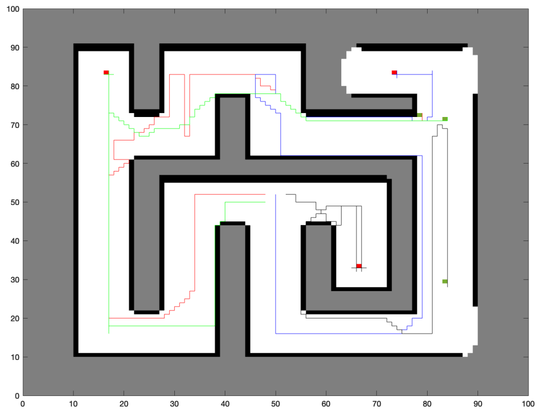
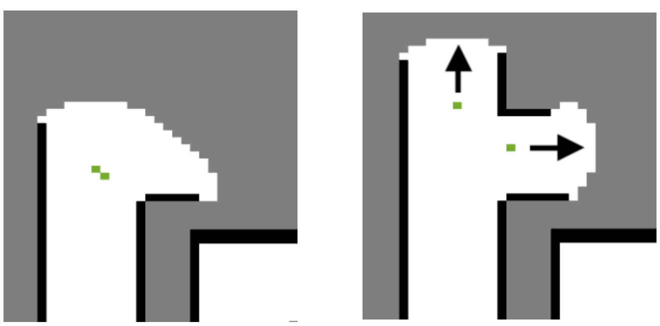


Fig. 2: Robot trajectories for a 100 $\times$ 100 grid, with 4 robots starting at $(50, 50)$.

Fig. 3 shows the benefit of the utility function reduction in the same scenario. Initially, the two robots close to each other and exploring the same frontier. When a fork appears, the algorithm assigns the first robot to explore straight up, then discounts the utility function to incentivize the second robot to go to the right. This leads to fast and efficient exploration of the map.



(a) (b)

Fig. 3: Notice how the robot is incentivized to explore different frontiers because of the discounting when transitioning from (a) to (b).

B. Information-Based Method

For all simulations in information-based method, we use $p_0 = 0.1$, probability parameter $\beta = 0.05$, $\alpha = 0.05$, with a sensing range that includes the 13 grid cells surrounding the robot ($|\mathbb{S}| = 13$). Fig. 4 is an example of robots exploring in a 30*30 grid-like environment. The source posterior map (on the left) is initialized at low values for each state, with robots (marked in green) updating the map information collectively. Whenever a robot finds a source, the source location will be reflected in posterior map with yellow or orange color.

In the exploration from Fig. 4, we see that a robot is stuck at a place in the middle. In some other simulations with fewer number of robots being employed in the environment, there can often be failures to explore all the sources. This is because the mutual information we are reasoning about is locally defined, we can get stuck between two actions if action space is limited.

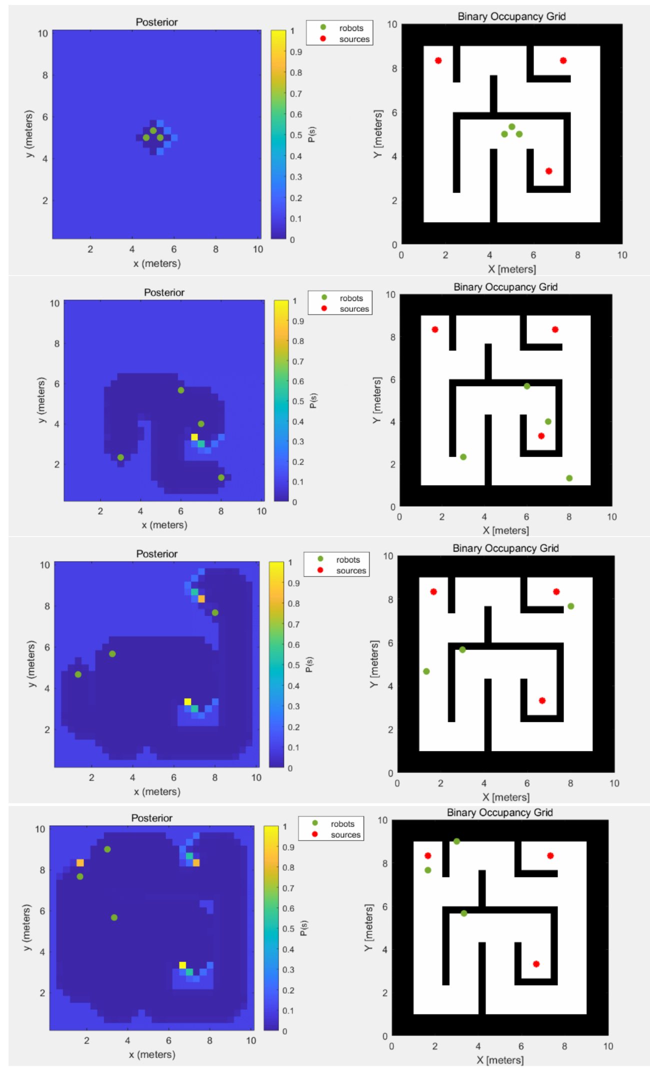


Fig. 4: Robot exploration with posterior map update for 30 $\times$ 30 grid, 4 robots at (50 m, 50 m)

C. Comparison

1) Effect of Grid Size and Number of Robots: A plot of the number of iterations needed to locate all 3 sources versus the number of robots starting at (50 m, 50 m) for information-theoretic method and both grid sizes in frontier-based method is shown in Fig. 5. As also observed in [16], the number of iterations needed plateaus as the number of robots increases, since the limitation gradually shifts to the iterations required to reach the source. The number of iterations needed for 30 $\times$ 30 grid is significantly less than that for 100 $\times$ 100 grid, which makes sense as, at each iteration, each drone moves a distance of $0.\dot{3}$ m on the 30×30 grid but moves only a distance of 0.1 m on the 100×100 grid. Thus, the drone on the 100×100 grid would take a longer time and more iterations to reach the source. While the number of iterations needed for the frontier-based method and the information-theoretic-based method do not differ by much, the info-theoretic method fails to locate the source for the case with 1, 2, and 3 robots starting at (50 m, 50 m).

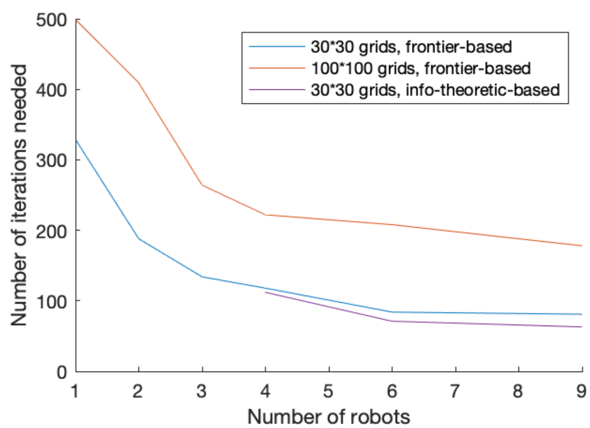


Fig. 5: Number of iterations needed to locate all 3 sources versus the number of robots starting at (50 m, 50 m) for frontier method with different grid sizes and for information-theoretic method.

2) Effect on computation time: Table I compares the computation time of both algorithms for a variety of robot network sizes. The frontier based method is approximately twice as efficient as the information based method. This is due to the high complexity of calculating the mutual information and the large state space that the information based method is based on. For both, computation time is approximately linear with the number of robots, which is expected as both are sequential approaches.

It is worth noting that computation time for the frontier method increases with the grid size of the map, as a finer resolution increases the number of frontier points to reason over. For the information-theoretic approach, the computation time is expected to grow exponentially with finer resolution. With a fixed sensing range, there will be more states included ($|\mathbb{S}|$ increases), and the total representation of the transformed state space $|\mathbb{S}'| = 2^{|\mathbb{S}|}$ exponentially increases. This brings computational difficulties in reasoning about every state representation in equation (6) (7) (8) .In our experiment, the problem becomes computationally intractable if our sensing range has incorporated number of states $|\mathbb{S}| > 16$

3) Discussions: Our frontier based method consistently finds all the sources in the environment and outperforms info-based method on computation effort in same environment setting. Generally, it can be applied to finer resolution within a tractable computational effort. Even though, there are couple of drawbacks.

One challenge of frontier based methods is the design of the utility function. In particular, consider the case depicted in figure 5. In this particular case, a pair of robots are moving to the left and to the right, meeting in the middle. Because this particular hallway is connected, the robots coming together is suboptimal exploration since there is more wasted movement. To solve this a utility function could be designed that penalizes robots from getting too close to each other. However if the hallway was non-connected, then the exploration would still be optimal, and a utility function that penalizes robots from getting too close together may increase exploration time for these cases. The takeaway from this is that in general it is hard to design utility functions for frontier exploration, and that manual tuning is needed for each application.

	2 on the sides	2 at the center	4 at the center
30×30, info	8.7	8.4	17.0
30×30, frontier	4.1	3.5	6.6
100×100, frontier	22.9	26.8	42.7

TABLE I: Computation time needed in seconds per 100 iterations for different methods, grid sizes, and starting position.

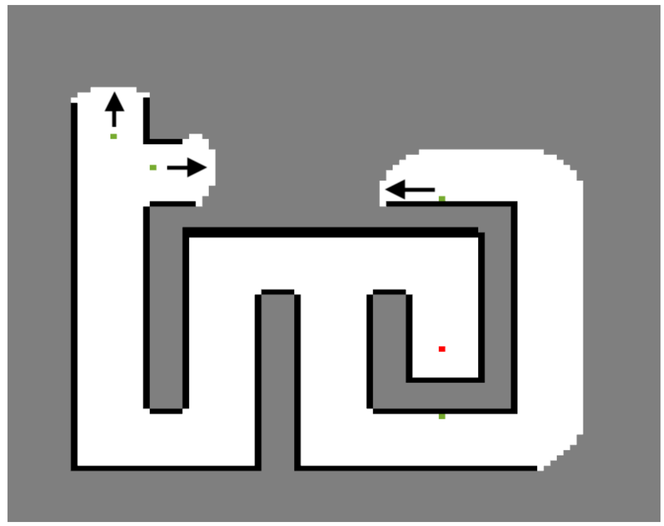


Fig. 6: Frontier Exploration Heuristic

Another weakness of the implementation is the centralized and sequential nature of our approach. Centralized methods are not robust to communication dropout, and sequential planners introduce suboptimality since the frontier selection is done for each robot sequentially, rather than in consideration of the positon of all robots. These issues can be solved through Multi-Robot Task Allocation (MRTA) based methods, effectively solving the travelling salesman problem. However this approach requires significant computational cost, since all permutations of robots and frontiers need to have their utility computed in order to allocate a frontier.

Our information-theoretic-based approach is a naive, simple-to-implement method for robot exploration by only requiring the robot to look at a probability space in a local area of the map. With more robots being employed, the searching becomes more successful and efficient. But there are downsides.

The primary limitation of our info-based approach is computational time. For an environment with coarser resolution, we can effectively run the simulation. However, if we segment the environment into more grids, more states will be included in sensing range and it quickly becomes computationally infeasible to update the posterior and calculate the mutual information. One way to address this issue is to assume some dependency between grid cell that are close to each other. For example, one can assume sources are far from each other so we could represent the state distribution in the sensing range as $|\mathcal{S}|$ one-hot vectors instead of $2^{|\mathcal{S}|}$ different distributions. However, doing this will complicate map posterior update because the probability change of each cell also affects other cells in the surrounding.

The other deficiency of our information-theoretic algorithm is that it acts greedily with respect to the mutual information gain in the next step and does not consider multi-step action sequences. This can cause the robots to move redundantly and get stuck, if the mutual information changes are nearly equal in multiple adjacent positions. This can be somewhat alleviated if random noise is used as a a tie-breaker, but still leads to suboptimal behavior.

The information-theoretic-based method also only updates the state belief over a small region of the map immediately surrounding the robot. This means that the robots only consider changes in entropy in the immediate surrounding region of candidate states. Because the algorithm is sequential to reduce computational expenses, there is potentially a more optimal group action by considering how multiple simultaneous measurements can reduce the total entropy the most. Also, implicit in this model is more independence between measurements than reality. Consider a robot that receives a positive measurement in a particular state and increases its belief distribution over $\mathcal{S}^t$ accordingly. It then moves one unit over and gets a negative measurement in a region $\mathcal{S}^{t+1}$. The overlapping sensing region $\mathcal{S}^t \cap \mathcal{S}^{t+1}$ will have appropriately decreased probability of being a source, but this implies that the excluded region $\mathcal{S}^t \setminus \mathcal{S}^{t+1}$ should have increased probability; future measurements affect the updates due to previous measurements. This would require a history of all measurements and associated sensing regions to properly account for, which increases the memory requirements of the algorithm immensely.

V. Conclusions

We applied two different multi-robot exploration strategies to the problem of radiation source localization. The frontier based method utilizes two different exploration modes in order to switch between exploring the map and localizing radiation sources. In contrast, the information based method utilizes an abstracted, remote sensing measurement model which infers the location of nearby sources from repeated measurements in different locations.

We tested each on a variety of initial conditions and maps, and found that for sufficiently large fleets of UAVs, both the frontier and information based methods performed nearly as well. However, the frontier exploration algorithm successfully pinpoints sources with small networks of robots. The information-theoretic strategy struggles when there are few UAVs as the individual controllers can get stuck because they only reason about local changes.

Going forward, both algorithms show potential in the field of radiation source localization. Both methods could be extended for decentralized, distributed multi-robot exploration in future work. Also, the effects of robot dynamics or measurement noise on each algorithm could be analyzed for increased robustness in practical application.

References

[1] B. Yamauchi, "Frontier-based exploration using multiple robots," in *Proceedings of the second international conference on Autonomous agents*, 1998, pp. 47–53.

[2] W. Burgard, M. Moors, D. Fox, R. Simmons, and S. Thrun, "Collaborative multi-robot exploration," in *Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings (Cat. No. 00CH37065)*, IEEE, vol. 1, 2000, pp. 476–481.

[3] R. Zlot, A. Stentz, M. B. Dias, and S. Thayer, "Multi-robot exploration controlled by a market economy," in *Proceedings 2002 IEEE international conference on robotics and automation (Cat. No. 02CH37292)*, IEEE, vol. 3, 2002, pp. 3016–3023.

[4] C. Stachniss, Ó. Martinez Mozos, and W. Burgard, "Efficient exploration of unknown indoor environments using a team of mobile robots," *Annals of Mathematics and Artificial Intelligence*, vol. 52, no. 2, pp. 205–227, 2008.

[5] Y. Xiao, J. Hoffman, T. Xia, and C. Amato, "Learning multi-robot decentralized macro-action-based policies via a centralized q-net," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, 2020, pp. 10 695–10 701.

[6] A. H. Tan, F. P. Bejarano, and G. Nejat, "Deep reinforcement learning for decentralized multi-robot exploration with macro actions," *arXiv preprint arXiv:2110.02181*, 2021.

[7] C. E. Shannon, "A mathematical theory of communication," *The Bell system technical journal*, vol. 27, no. 3, pp. 379–423, 1948.

[8] A. Cameron and H. Durrant-Whyte, "A bayesian approach to optimal sensor placement," *The International Journal of Robotics Research*, vol. 9, no. 5, pp. 70–88, 1990.

[9] A. Makarenko and H. Durrant-Whyte, "Decentralized data fusion and control in active sensor networks," in *Proceedings of the Seventh International Conference on Information Fusion*, Citeseer, vol. 1, 2004, pp. 479–486.

[10] B. Grocholsky, "Information-theoretic control of multiple sensor platforms," 2002.

[11] R. Olfati-Saber, E. Franco, E. Frazzoli, and J. S. Shamma, "Belief consensus and distributed hypothesis testing in sensor networks," in *Networked Embedded Sensing and Control*, Springer, 2006, pp. 169–182.

[12] B. J. Julian, M. Angermann, M. Schwager, and D. Rus, "A scalable information theoretic approach to distributed robot coordination," in *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, IEEE, 2011, pp. 5187–5194.

[13] M. Schwager, P. Dames, D. Rus, and V. Kumar, "A multi-robot control policy for information gathering in the presence of unknown hazards," in *Robotics research*, Springer, 2017, pp. 455–472.

[14] S. Kemp, A. Torgesen, J. Rogers, and J. How, "Optimal uav trajectory planning for radiological seearch," in *2020 ETI Summer Meeting.*, 2020.

[15] L. K. Chung, A. J. Kent, M. A. Cooney, J. D. Noey, K. J. Liebler, and K. J. Kearfott, "Simulations and experimental verifications of an algorithm for radiation source mapping and navigational path generation," *Health physics*, vol. 120, no. 6, pp. 648–660, 2021.

[16] W. Burgard, M. Moors, C. Stachniss, and F. E. Schneider, "Coordinated multi-robot exploration," *IEEE Transactions on robotics*, vol. 21, no. 3, pp. 376–386, 2005.

[17] K. Karur, N. Sharma, C. Dharmatti, and J. E. Siegel, "A survey of path planning algorithms for mobile robots," *Vehicles*, vol. 3, no. 3, pp. 448–468, 2021.