Bi-objective trail-planning for a robot team orienteering in a hazardous environment

Cory M. Simon*[1], Jeffrey Richley[2], Lucas Overbey‡[2], and Darleen Perez-Lavin†[2]

[1]School of Chemical, Biological, and Environmental Engineering. Oregon State University. Corvallis, OR. USA.
[2]Naval Information Warfare Center Atlantic. Charleston, SC. USA.

Corresponding author:
Cory M. Simon[1]

Email address: cory.simon@oregonstate.edu

ABSTRACT

Teams of mobile [aerial, ground, or aquatic] robots have applications in resource delivery, patrolling, information-gathering, agriculture, forest fire fighting, chemical plume source localization and mapping, and search-and-rescue. Robot teams traversing hazardous environments—with e.g. rough terrain or seas, strong winds, or adversaries capable of attacking or capturing robots—should plan and coordinate their trails in consideration of risks of disablement, destruction, or capture. Specifically, the robots should take the safest trails, coordinate their trails to cooperatively achieve the team-level objective with robustness to robot failures, and balance the reward from visiting locations against risks of robot losses.

Herein, we consider bi-objective trail-planning for a mobile team of robots orienteering in a hazardous environment. The hazardous environment is abstracted as a directed graph whose arcs, when traversed by a robot, present known probabilities of survival. Each node of the graph offers a reward to the team if visited by a robot (which e.g. delivers a good to or images the node). We wish to search for the Pareto-optimal robot-team trail plans that maximize two [conflicting] team objectives: the expected (i) team reward and (ii) number of robots that survive the mission. A human decision-maker can then select trail plans that balance, according to their values, reward and robot survival. We implement ant colony optimization, guided by heuristics, to search for the Pareto-optimal set of robot team trail plans. As a case study, we illustrate with an information-gathering mission in an art museum.

1 INTRODUCTION

1.1 Applications of a team of mobile robots

Mobile [aerial [58], ground [20], or aquatic [19]] robots equipped with sensors, actuators, and/or cargo have applications in agriculture [79, 4, 63], commerce [96], the delivery of goods [22], search-and-rescue [73, 76], chemical, biological, radiological, or nuclear incident response [68, 46], environmental monitoring [34, 44, 102], safety monitoring in industrial chemical plants [87, 38], forest fire monitoring and fighting [64], wildlife conservation [52], patrolling [3], target tracking [74], and military surveillance and reconnaissance.

Path-planning for mobile robots in static, dynamic, or unknown environments under various objectives (e.g., take least-cost paths) and constraints (e.g., collision avoidance) [53, 77, 98, 23] is important for good performance on tasks, efficiency, safety, and autonomy.

Often, we wish for a team of mobile robots to coordinate their paths in an environment to cooperatively achieve a shared, team-level objective [70, 71]. Compared to a single robot, a team can increase spatial coverage, decrease the time to achieve the objectives, and make achievement of the objectives robust to the failure of robots [80, 10].

For example, in the (NP-hard) team orienteering [40] problem (TOP) [16, 43, 91], a team of robots are mobile in an environment modeled as a graph (nodes: locations; edges: spatial connections between locations), and each node gives a reward to the team if visited by a robot. The TOP is to plan the paths of the robots, from a source to destination node, to gather the most rewards as a team under a travel budget for each robot. Loosely, the TOP problem, which can be formulated as an integer program, combines aspects of the classic knapsack problem (selecting the nodes from which to collect rewards, under the travel budget) and the traveling salesman problem (finding the shortest path that visits these nodes) [91].

1.2 Teams of mobile robots orienteering in hazardous environments

In some applications, robots move in a hazardous environment [89] and incur risks of failure, destruction, disablement, and/or capture. The hazards could originate from dangerous terrain, rough seas, strong winds, heat, radiation, corrosive chemicals, or mines—or, an adversary with the capability to attack, disable, destroy, or capture robots [1].

Robots traversing a hazardous environment should plan and coordinate their trails in consideration of risks of failure. First, robots should take the safest trails to visit their destination(s). Second, the robots should coordinate their trails to make achievement of the team objective resilient to robot failures [103]. A *resilient* team of robots [72] (i) anticipates failures and makes risk-aware plans that endow the team with *robustness*—the ability to withstand failures with minimal concession of the objective, and/or (ii) changes their trail plans during the mission, adapting in response to realized robot failures, to recoup the otherwise anticipated loss in the objective. Third, the trail plans must balance the rewards gained from visiting different locations against the risks incurred by the robots to reach those locations.

Models and algorithms have been developed for path-planning for robot teams orienteering in hazardous environments abstracted as graphs [103]. In the Team Surviving Orienteers problem (TSOP) [50, 49, 51], each node of the graph offers a reward to the team when visited by a robot, and each edge-traversal by a robot incurs a probability of destruction. The objective in the TSOP is to plan the paths of the robots (from a source to destination node) to maximize the expected team reward under the constraint that each robot survives the mission with a probability above a set threshold. (In the *offline* setting, the paths of the robots are set at the beginning of the mission, then followed without adaptation. In the *online* setting, the paths are updated during the mission in response to realized robot failures.) Relatedly, the Foraging Route with the Maximum Expected Utility problem [31] is to plan the foraging route of a robot collecting rewards in a hazardous environment, but the rewards are lost if the robot is destroyed before returning to the source node to deposit the goods it collected. In the Robust Multiple-path Orienteering Problem (RMOP) [84], similarly, each node gives a reward to the team only if a robot visits it *and* survives the mission to deposit the reward, and the paths of the K robots are planned to maximize the team reward under the worst-case attacks of $\alpha < K$ of the robots by an adversary. The offline version of RMOP constitutes a two-stage, sequential, one-step game with perfect information: (1) the robot team chooses a set of paths then (2) the adversary, knowing these paths, chooses the subset of robots to attack and destroy. The optimal path plans for the robots must trade (i) redundancy in the nodes visited, to give robustness against attacks, and (ii) coverage of many nodes, to collect many rewards. In the hazardous orienteering problem [78, 66], robots incur a risk of destruction when visiting a node, and the reward is received only if the robot returns safely to the depot node. The time-bomb knapsack problem aims to maximize the expected profit of the items in the knapsack under a knapsack capacity when each item has a probability of exploding and causing the profit to be lost [65]. Notably, Subramanian (1997) was among the first to model stochastic hazards along a path in a graph [83]. Other work involving robot path-planning in hazardous environments includes maximizing coverage of an area containing threats to robots [57, 99], handling adversarial attacks on the sensors of the robots [60, 105, 62, 104], gathering information in an environment with unknown hazards [81], finding the optimal formation for a robot team [82], and multi-robot patrolling under adversarial attacks [45].

1.3 Our contribution

Herein, our contribution is: (1) framing and intuiting a bi-objective variant of the offline TSOP [50, 49, 51], the bi-objective team orienteering in hazardous environments (BOTOHE) problem, then (2) specifying a bi-objective ant colony optimization algorithm [47], guided by heuristics, to search for the Pareto-optimal set of robot-team trail plans, then (3) finally, solving and analyzing several examples of BOTOHE problem instances for insights.

In the BOTOHE problem, a team of robots are mobile in a hazardous environment, modeled as a directed graph whose arcs present known probabilities of destruction to robots traversing them. Each node of the graph offers a reward to the team if a robot visits it. The BOTOHE problem is to plan the closed trails of the robots to maximize two team-level objectives: the expected (1) rewards accumulated by the team via visiting nodes and (2) number of robots that survive the mission. (We focus on the offline setting, corresponding with a lack of communication with/between robots after the mission executes.) See Fig. 1a.

Three interesting features of the BOTOHE are: (1) for the survival objective, robots a) risk-aversely avoid visiting dangerous subgraphs despite rewards offered by nodes therein and b) take the safest closed trails to visit the nodes assigned to them; (2) for the reward objective, the robots a) daringly visit dangerous subgraphs to attempt collection of the rewards offered by nodes therein, b) visit the lower-risk and higher-reward subgraphs earlier in the mission, and c) build node-visit redundancy into their trail plans (i.e., multiple robots visit the same node) to make the team-reward robust to the loss of robots during the mission; and (3) comparing (1a) and (2a), the two objectives are inherently conflicting, as the robots must risk their survival while taking their trails to visit nodes and collect rewards.

To handle the conflict between the reward and survival objectives in the BOTOHE, we search for the Pareto-optimal set [69, 11] of robot-team trail plans. By definition, a Pareto-optimal trail plan cannot be altered to give a higher expected reward without lowering the expected number of robots that survive—and vice versa. Fig. 1b illustrates. We then can present the Pareto-optimal set to a human decision-maker to select the team trail plan that strikes some tradeoff between the two objectives, based on their value of rewards vs. robot survival.

To search for the Pareto-optimal set of robot-team trail plans in the BOTOHE, we specify and implement a bi-objective ant colony optimization (ACO) algorithm [47], guided by heuristics. We adopt ACO [32, 7, 6] because it is naturally-suited for searching for trail sets on graphs, For illustration, we solve and analyze a BOTOHE problem instance for information-gathering in an art museum. Through ablation studies, we quantify the contribution of the greedy heuristics and pheromone to the search efficiency.

2 THE BI-OBJECTIVE TEAM ORIENTEERING IN HAZARDOUS ENVIRONMENTS (BOTHE) PROBLEM

2.1 Problem setup

Here, we frame the Bi-Objective robot-Team Orienteering in Hazardous Environments (BOTOHE) problem.

A homogenous team of K robots are mobile in an environment modeled as a directed graph $G = (\mathscr{V}, \mathscr{E})$. Each node $v \in \mathscr{V}$ represents a location in the environment (e.g., a room in a building). Each arc $(v, v') \in \mathscr{E}$, an ordered pair of distinct nodes, represents a one-way, direct spatial connection (e.g. a door- or hall-way) to travel from node v to node v'. *Mobility* of the robots implies they may walk on the graph G, i.e. sequentially hop from a node v to another node v' via traversing arc $(v, v') \in \mathscr{E}$. All K robots begin at a base node $v_b \in \mathscr{V}$. Regarding reachability, we assume G is strongly connected—but, not necessarily complete.

Owing to unpredictable and/or uncertain hazards in the environment, a robot incurs a probability of destruction of $1 - \omega(v, v')$ when, beginning at node v, it attempts to traverse arc $(v, v') \in \mathscr{E}$ to visit node

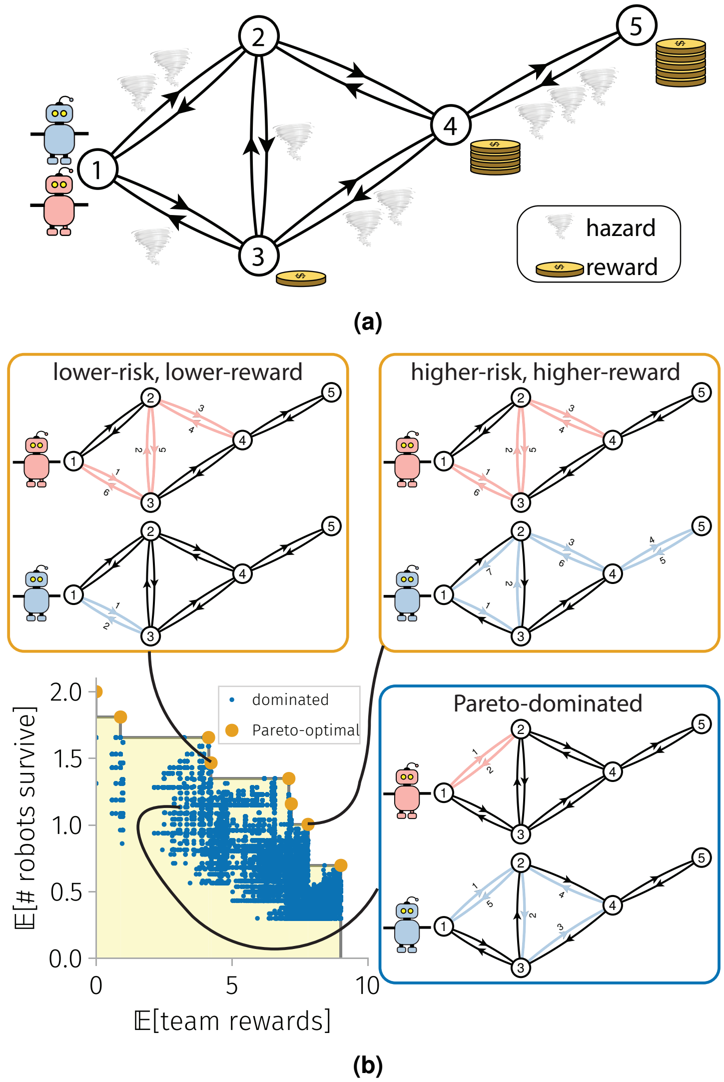


Figure 1. The bi-objective team orienteering in hazardous environments (BOTHE) problem. (a) A team of robots are mobile on a directed graph whose nodes offer a reward to the team if visited by a robot and arcs present a probability of destruction to robots that traverse them (tornado: 1/10 probability of destruction). The task is to plan the trails of the robots to maximize the expected reward collected and the expected number of robots that survive. (b) Pareto-optimal and -dominated robot-team trail plans scattered in objective space, with two Pareto-optimal plans and one Pareto-dominated plan shown.

v'. Each outcome (survival or destruction) is an independent event. The arc survival probability map $\omega : \mathscr{E} \to (0,1]$ is known, static over the course of the mission, and not necessarily symmetric (owing to e.g., [directional] air or water currents or bright sunlight or an adversary with limited range nearer one node than another).

Each node $v \in \mathscr{V}$ of the graph G offers a reward $r(v) \in \mathbb{R}_{\geq 0}$ to the robot team if visited by one or more robots over the course of the mission. The reward $r(v)$ quantifies the utility gained by the team when a robot e.g. delivers a resource to node v, takes an image of node v and transmits it back to the command center, or actuates some process (e.g. turns a valve) at node v. The total reward collected by the team is additive among nodes of the graph. Note, (1) even if a robot is destroyed after leaving a node, the harvested reward from that node is still accumulated by the team; and (2) multiple visits to a node by the same or distinct robot(s) do not give marginal reward over a single visit.

To collect rewards in this hazardous environment, the team of robots must plan a set[1] of closed, directed trails $\mathscr{P} := \{\rho_1, ..., \rho_K\}$ on the graph G to follow. A *directed trail* [21, 97] is an ordered list of nodes $\rho = (\rho[0], \rho[1], ..., \rho[|\rho|])$ where (i) $\rho[i] \in \mathscr{V}$ is the ith node in the trail, (ii) an arc exists from each node in ρ to the subsequent node, i.e., $(\rho[i-1], \rho[i]) \in \mathscr{E}$ for $1 \leq i \leq |\rho|$, (iii) $|\rho|$ is the number of arcs traversed in the trail, and (iv) the arcs traversed in the trail are unique, i.e. each arc in the multiset $\{(\rho[i-1], \rho[i])\}_{i=1}^{|\rho|}$ has a multiplicity of one. Note, unlike a path, the nodes in a trail are not necessarily distinct [95]. A *closed* trail begins and ends at the same node, i.e. $\rho[0] = \rho[|\rho|]$, which, here, $= v_b$. Each trail ρ_k belonging to the robot-team trail plan $\mathscr{P}$ constitutes a *plan* because robot k may be destroyed in the process of following ρ_k and thus not *actually* visit all nodes in ρ_k. A robot *survives* the mission if it visits all nodes in its planned trail and returns to the base node (without getting destroyed).

We wish to design the robot-team trail plan $\mathscr{P}$ to maximize two objectives: (1) the expected team reward, $\mathbb{E}[R]$, and (2) the expected number of robots that survive the mission, $\mathbb{E}[S]$. (Both R and S (i) are random variables owing to the stochasticity of robot survival while trail-following and (ii) depend on the robot-team trail plan $\mathscr{P}$ owing to different rewards among nodes and dangers among arcs.) That is, the BOTHE problem is:

$$\max_{\mathscr{P}=\{\rho_1,...,\rho_K\}} (\mathbb{E}[R(\mathscr{P})], \mathbb{E}[S(\mathscr{P})]) \tag{1}$$

when given the directed graph G as a spatial abstraction of the environment, the homogenous team of K mobile robots starting at the base node v_b, the node reward map $r : \mathscr{V} \to \mathbb{R}_{\geq 0}$, and the arc survival probability map $\omega : \mathscr{E} \to (0, 1]$.

Because the bi-objective optimization problem in eqn. 1 presents a conflict between designing the robot-team trail plan to maximize the expected reward *and* the number of surviving robots, we seek the *Pareto-optimal set* [69, 11] of team-robot trail plans. The conflict is: (1) to maximize survival, a risk-averse robot team would not visit a dangerous region even if large rewards were contained therein, sacrificing reward for survival; (2) to maximize reward, a daring robot team would visit a dangerous region even if only small rewards were contained therein, sacrificing survival for reward. Hence, a utopian robot-team trail plan that simultaneously maximizes *both* reward and survival objectives is unlikely to exist; the ultimate team trail plan selected for the mission must strike some tradeoff between reward and survival. By definition, a *Pareto-optimal* [69, 11] robot-team trail plan $\mathscr{P}^*$ cannot be altered to increase the survival objective $\mathbb{E}[S(\mathscr{P}^*)]$ without compromising (decreasing) the reward objective $\mathbb{E}[R(\mathscr{P}^*)]$—and vice versa. See Fig. 1b and the formal definition below. The Pareto-optimal set of team trail plans is valuable for presenting a *portfolio* of team trail plans to a human decision-maker, who ultimately invokes their values—i.e., makes a team-reward vs. robot-survival tradeoff—by selecting a good team trail plan from the Pareto set.

Formal definition of Pareto-optimality and Pareto-front. Plan $\mathscr{P}^*$ belongs to the Pareto-optimal set of plans if and only if no other plan $\mathscr{P}'$ *Pareto-dominates* it. By definition, a plan $\mathscr{P}'$ Pareto-dominates plan $\mathscr{P}^*$ if and only if both:

$$\left(\mathbb{E}[R(\mathscr{P}')] \geq \mathbb{E}[R(\mathscr{P}^*)]\right) \wedge \left(\mathbb{E}[S(\mathscr{P}')] \geq \mathbb{E}[S(\mathscr{P}^*)]\right) \tag{2}$$

$$\left(\mathbb{E}[R(\mathscr{P}')] > \mathbb{E}[R(\mathscr{P}^*)]\right) \vee \left(\mathbb{E}[S(\mathscr{P}')] > \mathbb{E}[S(\mathscr{P}^*)]\right). \tag{3}$$

So, the Pareto-dominating plan $\mathscr{P}'$ is not worse than plan $\mathscr{P}^*$ for reward nor for survivability, and is better for one or both of them. If a plan $\mathscr{P}'$ were to Pareto-dominate another plan $\mathscr{P}^*$, one would objectively never choose plan $\mathscr{P}^*$ over plan $\mathscr{P}'$, regardless of how they relatively value the two objectives. The *Pareto front* is the set of objective vectors $\{(\mathbb{E}[R(\mathscr{P}^*)], \mathbb{E}[S(\mathscr{P}^*)])\}$ associated with the Pareto-optimal set of team trail plans $\{\mathscr{P}^*\}$.

Comparison with TSOP. The BOTHE problem formulation follows the offline TSOP [50] with three modifications: we (i) omit the constraints that each robot survives above a threshold probability[2], (ii) allow robots to follow trails instead of restricting to paths, as paths prevent robots from visiting a given node more than once and thus from e.g., harvesting reward from a leaf node having one in- and one out-degree involving the same node, and (iii) have two objectives instead of one and seek the Pareto-optimal set of team trail plans.

[1] Since the robot team is homogenous, we consider the team trail plan as permutation-invariant and thus treat it as a set not a list.
[2] i.e., we accept if one [uncrewed] robot bears much more risk of destruction than another during the mission.

2.2 Probability distributions and expectations of $R(\mathscr{P})$ and $S(\mathscr{P})$

We now write formulas for the two objectives—the expectations of the team reward $R(\mathscr{P})$ and of number of robots that survive the mission $S(\mathscr{P})$—as a function of the robot-team trail plan $\mathscr{P}$. These formulas follow from the directed graph G, arc survival probability map ω, and node reward map r. Fig. 2 illustrates our notation.

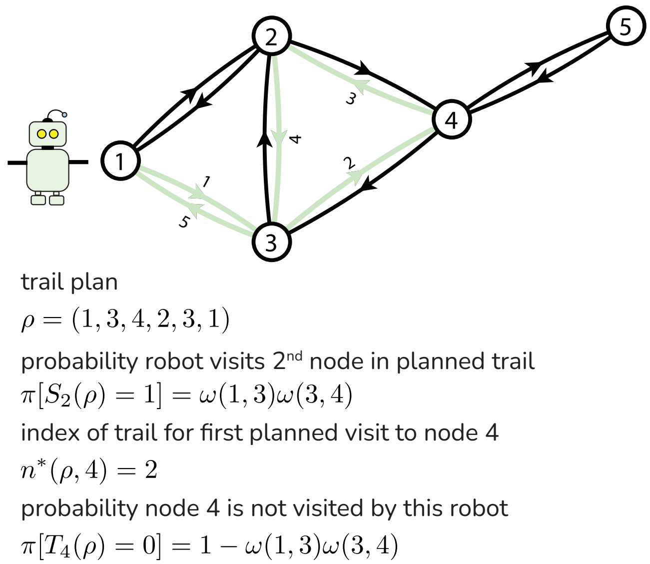


Figure 2. Illustrating notation for a particular robot's trail plan.

2.2.1 The survival of a single robot along its followed trail

Central to computing both $\mathbb{E}[S(\mathscr{P})]$ and $\mathbb{E}[R(\mathscr{P})]$ is the probability that a robot survives to reach a given node in its planned trail. Let $S_n(\rho)$ for $0 \le n \le |\rho|$ be the Bernoulli random variable that is one if a robot following trail ρ survives to visit the nth node in the trail and zero otherwise. For the event of survival, the robot must survive traversal of *all* of the first n arcs in its trail to visit node $\rho[n]$. So, since the survival of a robot over each arc traversal attempt is an independent event, the probability that a robot following trail ρ successfully visits node $\rho[n]$ is the product of the survival probabilities of the first n arc-hops in the trail:

$$\pi[S_n(\rho) = 1] = \prod_{i=1}^{n} \omega(\rho[i-1], \rho[i]) = 1 - \pi[S_n(\rho) = 0]. \tag{4}$$

The second equality follows because the complement of the event of survival is destruction.

2.2.2 Expected number of robots that survive

Now, we write a formula for the expected number of robots that survive the mission, $\mathbb{E}[S(\mathscr{P})]$. The event of survival of each robot is independent of the other robots. Consequently, the number of robots that survive the mission, S, is the sum of the Bernoulli random variables indicating the survival of each robot over its planned trail:

$$S(\mathscr{P} = \{\rho_1, ..., \rho_K\}) = \sum_{k=1}^{K} S_{|\rho_k|}(\rho_k). \tag{5}$$

Thus, S follows the Poisson-Binomial distribution [88]. Seen from eqn. 5 and the linearity of the expectation operator, the expected number of robots that survive the mission is:

$$\mathbb{E}[S(\mathscr{P} = \{\rho_1, ..., \rho_K\})] = \sum_{k=1}^{K} \mathbb{E}[S_{|\rho_k|}(\rho_k)] = \sum_{k=1}^{K} \pi[S_{|\rho_k|}(\rho_k) = 1] \tag{6}$$

with $\pi[S_{|\rho_k|}(\rho_k) = 1]$ given in eqn. 4.

2.2.3 Expected team reward

Now, we write a formula for the expected team reward, $\mathbb{E}[R(\mathscr{P})]$.

First, we calculate the probability that a robot following a given trail ρ does not visit a given node $v \in \mathscr{V}$. Let the Bernoulli random variable $T_v(\rho)$ be one if a robot following trail ρ visits node v and zero if it does not. If node v is not in the planned trail ρ, $T_v(\rho) = 0$ with certainty. Now, suppose node v is in the planned trail ρ. Let n^* be the index in the trail where the robot plans its first visit to node v:

$$n^*(\rho, v) = \min_{\substack{n \in \{0,...,|\rho|\} \\ \rho[n]=v}} n. \tag{7}$$

Then, $T_v(\rho)$ is equal to the random variable $S_{n^*}(\rho)$ because the robot visits node v if and only if it survives its first n^* arc-hops to first land on node v. So, the probability node v is not visited by a robot following trail ρ is:

$$\pi[T_v(\rho) = 0] = \begin{cases} 1 & v \notin \rho \\ \pi[S_{n^*(\rho,v)}(\rho) = 0] & v \in \rho \end{cases} \tag{8}$$

with $\pi[S_{n^*}(\rho) = 0]$ calculated using eqn. 4.

Second, we calculate the probability that a given node $v \in \mathscr{V}$ is visited by one or more robots on a team following trail plans $\mathscr{P} = \{\rho_1, ..., \rho_K\}$; only then is the reward offered by that node, $r(v)$, accumulated by the team. Let $T_v(\mathscr{P})$ be the Bernoulli random variable that is one if one or more robots on the team with trail plans $\mathscr{P}$ visit node v and zero otherwise (i.e. if and only if zero of the robots visit v). The complement of the event $T_v(\mathscr{P}) = 1$ is that all K robots do not visit node v (independent events), so:

$$\pi[T_v(\{\rho_1, ..., \rho_K\}) = 1] = 1 - \prod_{k=1}^{K} \pi[T_v(\rho_k) = 0]. \tag{9}$$

with $\pi[T_v(\rho) = 0]$ in eqn. 8.

Finally, the total team reward collected by the robot-team following trail plan $\{\rho_1, ..., \rho_K\}$ is the sum of the rewards given to the team by each node:

$$R(\{\rho_1, ..., \rho_K\}) = \sum_{v \in \mathscr{V}} r(v) T_v(\{\rho_1, ..., \rho_K\}), \tag{10}$$

where the reward from node v, $r(v)$, is accumulated if and only if node v is visited by one or more robots (i.e. iff $T_v(\mathscr{P}) = 1$). The expected reward accumulated over the mission by a robot-team following trail plans $\{\rho_1, ..., \rho_K\}$ is:

$$\mathbb{E}[R(\{\rho_1, ..., \rho_K\})] = \sum_{v \in \mathscr{V}} r(v) \pi[T_v(\{\rho_1, ..., \rho_K\}) = 1] \tag{11}$$

with $\pi[T_v(\{\rho_1, ..., \rho_K\}) = 1]$ given in eqn. 9.

3 BI-OBJECTIVE ANT COLONY OPTIMIZATION TO SEARCH FOR THE PARETO-OPTIMAL ROBOT-TEAM TRAIL PLANS

To efficiently search for the Pareto-optimal set of robot-team trail plans defined by eqn. 1, we employ bi-objective (BO) ant colony optimization (ACO) [47]. ACO [32, 7, 6] is a meta-heuristic for combinatorial optimization problems framed as a search for an optimal path (or trail) on a graph. As a swarm intelligence method [7], ACO is inspired by the behavior of ants efficiently foraging for food [8]. See Box 1. As precedent, variants of the single-robot- and team-orienteering problem have been efficiently and effectively solved by ACO [54, 18, 92, 86, 17, 67] (as well as other [meta-]heuristics [39, 26, 15, 13]), but not the TSOP.

Box 1: Pheromone trail laying and following by foraging ants

Largely (but not exclusively [37, 25, 75]) via laying and following pheromone trails [25], some species of foraging ant colonies advantageously [30] can collectively select (i) the shortest path from the nest to a food source [41] and (ii) the highest quality food source among multiple options equidistant from the nest [5].

Pheromone, a relatively volatile chemical [27], serves as an olfactory cue for ants [56], allowing for *indirect* communication between ants in the colony. On the way back to the nest from a food source, ants deposit pheromone on the ground—an amount modulated by the quality of the food source [5]. Shorter paths to higher quality food sources accumulate pheromone more quickly. Since ants are recruited to and follow pheromone [5, 25], these paths recruit more ants and get reinforced. Via this positive feedback, the colony can collectively select the shortest path or highest-quality food source [48, 25, 7]. The pheromone trails laid by the colony form a *collective memory*—a local guide for ants to high-quality food sources and short paths to them [48]. Some species of ants can deposit multiple species of pheromone (from different glands) with e.g. different longevities [25], allowing for more complex indirect communication [48, 75].

As negative feedback mechanisms, (i) pheromone evaporates over time [48, 27, 90], allowing the ant colony to "forget" trails to exhausted food sources, and (ii) ants deposit less pheromone on trails a) with already-high pheromone concentrations [24] or b) leading to food sources already occupied by their nestmates [94].

Ants select among pheromone trails with some degree of stochasticity [29], which is beneficial for (i) continual exploration to find even shorter paths and even higher-quality food sources, (ii) exploiting multiple food sources in parallel, and (iii) plasticity in a dynamic environment [30, 85, 28, 35, 36].

The combination of positive feedback, negative feedback, and randomness in ants' pheromone-laying and -following can give rise to complex collective behavior despite simple interactions among decentralized individuals [9, 7, 41, 48, 36, 93].

In bi-objective ACO, we simulate a heterogenous colony of artificial ants walking on the graph G—under a loose analogy, foraging for food—over many iterations. At each iteration, each worker ant stochastically constructs a robot-team trail plan $\{\rho_1, ..., \rho_K\}$ robot-by-robot, arc-by-arc, biased by (i) the amounts of two species of pheromone on the arcs encoding the colony's past experiences and (ii) two heuristics that score the greedy, *a priori* appeal of each arc for each objective. As a division of labor, each worker ant specializes by searching for team trail plans belonging to a different region of the Pareto-front. Then, ants that found the Pareto-optimal team trail plans over that iteration deposit pheromone of each species on the arcs involved, proportionally to the value of the objective achieved by that plan. An elitist ant [33] maintains a set of global (i.e, over all iterations) Pareto-optimal team trail plans and also deposits pheromone on arcs. Finally, to prevent stagnation and promote continual exploration, a fraction of the pheromone evaporates. After many iterations, the ACO algorithm returns the [approximate[3]] Pareto-optimal set of robot-team trail plans maintained by the elitist ant.

3.1 The heterogenous artificial colony of ants, pheromone, and heuristics

Our heterogenous artificial ant colony consists of (i) N_{ants} worker ants and (ii) an elitist ant. Worker ant $i \in \{1, ..., N_{\text{ants}}\}$ in the colony is assigned a parameter $\lambda_i := (i-1)/(N_{\text{ants}}-1)$ dictating its balance of the two objectives when searching for Pareto-optimal team trail plans. A λ closer to zero (one) implies the ant prioritizes maximizing the expected reward $\mathbb{E}[R]$ (robot survivals $\mathbb{E}[S]$). So, different ants seek team trail plans belonging to different regions of the Pareto front. The elitist ant maintains the global-Pareto-optimal set of team trail plans.

Each arc $(v, v') \in \mathscr{E}$ of the graph is associated with (i) amounts of two distinct species of pheromone, $\tau_R(v, v')$ and $\tau_S(v, v')$, and (ii) heuristic scores $\eta_R(v, v')$ and $\eta_S(v, v')$. Both $\tau_{R,S}(v, v') > 0$ and $\eta_{R,S}(v, v') > 0$ score the promise of arc (v, v') for belonging to Pareto-optimal team trail plans that maximize $\mathbb{E}[R]$ and $\mathbb{E}[S]$, respectively, and guide worker ants' construction of robot-team trail plans. The pheromone is learned, reflecting the past collective experience/memory of the ant colony. Due to deposition and

[3]The ACO meta-heuristic is not guaranteed to find all Pareto-optimal solutions nor neglect to include a Pareto-dominated solution.

evaporation, the pheromone is dynamic over iterations. By contrast, the heuristic is static and scores the *a priori*, greedy/myopic appeal of each arc to accelerate the convergence of ACO.

3.2 Constructing robot-team trail plans

Each iteration, every worker ant stochastically constructs a team trail plan $\mathscr{P} = \{\rho_1, ..., \rho_K\}$ by sequentially allocating trails to the robots and [conceptually] following the closed trail the ant designs for each robot. Computing the objectives achieved under each plan via eqns. 6 and 11 gives data $\{(\mathscr{P}_i, \mathbb{E}[R(\mathscr{P}_i)], \mathbb{E}[S(\mathscr{P}_i)])\}_{i=1}^{N_{\text{ants}}}$ used to deposit pheromone (worker ant) and update the global Pareto-optimal set (elitist ant).

A robot trail is constructed by iteratively applying a stochastic, partial-trail extension rule until the closed trail is complete. Suppose an ant with objective-balancing parameter λ is constructing the closed trail for robot k, ρ_k, and currently resides at node $v = \rho_k[i]$. Namely, the ant has selected (i) the trails for the previous $k-1$ robots, $(\rho_1, ..., \rho_{k-1})$, and (ii) an incomplete/partial trail for robot k, $\tilde{\rho}_k = (v_b, \rho_k[1], ..., \rho_k[i] = v)$, giving its first i arc-hops. The probability of next hopping to node v' across arc $(v, v') \in \mathscr{E}$ not yet traversed in the partial trail $\tilde{\rho}_k$ is:

$$\pi(v' \mid \rho_1, ..., \rho_{k-1}, \tilde{\rho}_k) \propto \left[\tau_R(v, v')\eta_R(v, v'; \rho_1, ..., \rho_{k-1}, \tilde{\rho}_k)\right]^{1-\lambda} \left[\tau_S(v, v')\eta_S(v, v')\right]^{\lambda}. \tag{12}$$

The partial trail is more likely to be extended with $v' =: \rho_k[i+1]$ if arc (v, v') has more pheromone $\tau_{R,S}(v, v')$ and/or greedy heuristic appeal $\eta_{R,S}(v, v')$—with more or less emphasis on the reward or survivability pheromone/heuristic depending on the ant's λ. Iteratively extending the partial trail using eqn. 12, the ant completes the trail for robot k after it traverses the self-loop of the base node, (v_b, v_b). Then, the ant begins trail construction for robot $k+1$ if $k < K$ or completes its team trail plan $\mathscr{P}$ if $k = K$.

Heuristics. For the survivability objective, we score the desirability of arc (v, v') with the probability of the robot surviving traversal of that arc, $\eta_S(v, v') := \omega(v, v')$—myopic because it does not consider survivability of arcs later in the trail. For the reward objective, we greedily score the desirability of arc (v, v') with the expected marginal reward the team receives by robot k visiting node v' next, which is $r(v')$ if (i) none of the previous $k-1$ robots visit node v', (ii) node v' is not planned to be visited earlier in the trail of robot k, and (iii) robot k survives its hop to node v' and zero otherwise:

$$\eta_R(v, v'; \rho_1, ..., \rho_{k-1}, \tilde{\rho}_k) := \pi[T_{v'}(\{\rho_1, ..., \rho_{k-1}\}) = 0)]\mathscr{I}[v' \notin \tilde{\rho}_k]\omega(v, v')r(v'), \tag{13}$$

with $\mathscr{I}(\cdot)$ the indicator function—myopic because it does not account for rewards robot k could collect further along the trail. Note, to prevent either heuristic from being exactly zero (resulting in *never* selecting that arc), we add a small number ε to each heuristic.

3.3 Pheromone update

At the end of each iteration, we update the pheromone maps τ_R and τ_S to capture the experience of the ants in finding Pareto-optimal robot-team trail plans, better-guide the ants' trail-building in the next iteration, and prevent stagnation (premature convergence).

The pheromone update rule is:

$$\tau_{R,S}(v, v') \leftarrow (1-\rho)\tau_{R,S}(v, v') + \Delta\tau_{R,S}(v, v') \text{ for } (v, v') \in \mathscr{E}, \tag{14}$$

with $\rho \in (0, 1)$ the evaporation rate (a hyperparameter) and $\Delta\tau_{R,S}(v, v')$ the amount of new pheromone deposited on arc (v, v') by the ants.

Evaporation, accomplished by the first term in eqn. 14, removes a fraction of the pheromone on every arc. This negative feedback mechanism prevents premature convergence to suboptimal trails and encourages continual exploration.

Deposition, accomplished by the second term in eqn. 14, constitutes indirect communication to ants in future iterations about which arcs tend to belong to Pareto-optimal team trail plans and the objectives achieved under those plans. First, the ants *collaborate* by (i) among the worker ants, comparing the solutions constructed *this* iteration to obtain the *iteration*-Pareto-optimal set of plans and (ii) worker ants sharing the solutions constructed this iteration with the elitist ant, who then updates the *global*-Pareto-optimal set. Second, the worker ants with an iteration-Pareto-optimal team trail plan execute their constructed plan while depositing pheromone on the arcs. Third, the elitist ant executes all global-Pareto-optimal team trail plans while depositing pheromone on the arcs. Each ant deposits both reward and survival pheromone in proportion to the reward and survival objectives, respectively, achieved under the plan they are following. (The objective is analogous with food quality). Specifically, amalgamating the iteration- and global-Pareto-optimal team trail plans into a multiset $\{(\mathscr{P}_p^*, \mathbb{E}[R(\mathscr{P}_p^*)], \mathbb{E}[R(\mathscr{P}_p^*)])\}_{p=1}^P$, arc (v, v') receives pheromone:

$$\Delta\tau_{R,S}(v,v') := \frac{1}{P}\sum_{p=1}^{P}\mathbb{E}[R(\mathscr{P}_p^*), S(\mathscr{P}_p^*)]\sum_{k=1}^{K}\sum_{i=1}^{|\rho_k^{(p)}|}\mathscr{I}\left[(v,v') = (\rho_k^{(p)}[i-1], \rho_k^{(p)}[i])\right]. \tag{15}$$

with $\mathscr{I}(\cdot)$ the indicator function. The sums are over, left-to-right, Pareto-optimal plans, robot trails in those plans, and arcs in those robot trails. The latter two sums count the number of times the arc (v, v') appears in $\mathscr{P}_p^*$. By construction, arcs that receive the most reward (survival) pheromone frequently belong to trails in the Pareto-optimal team trail plans with high reward (survival).

We initialize the pheromone on all arcs with $\tau_{R,S}(v, v\prime) = 1$, to allow the heuristic to completely guide the first iteration.

3.4 Area indicator for Pareto-set quality

From iteration-to-iteration, we measure the quality of the global [approximate] set of Pareto-optimal robot-team trail plans tracked by the elitist ant using an area indicator [14, 42]—loosely, the area in objective space enclosed between the origin and the [approximated] Pareto-front. Formally, the quality q of a Pareto set $\{\mathscr{P}_1^*, ..., \mathscr{P}_P^*\}$ is the area of the union of rectangles in 2D objective space:

$$q(\{\mathscr{P}_1^*, ..., \mathscr{P}_P^*\}) := \left|\bigcup_{p=1}^{P}\{o \in \mathbb{R}^2 : o \geq 0 \land o \leq (\mathbb{E}[R(\mathscr{P}_p^*)], \mathbb{E}[S(\mathscr{P}_p^*)])\}\right| \tag{16}$$

illustrated with the shaded yellow area in Fig. 1b.

4 RESULTS

We now implement bi-objective ant colony optimization to search for Pareto-optimal robot-team trail plans for an instance of bi-objective team orienteering in a hazardous environment. We visualize the pheromone trail and some of the Pareto-optimal solutions to gain intuition. We also conduct an ablation study to quantify the importance of the pheromone and heuristics for guiding the search. Our Julia code to reproduce our results and modify and build upon with more complexity is available at `github.com/SimonEnsemble/BO_ACO_TOHE`.

4.1 Problem setup

A team of $K = 3$ mobile robots are assigned an information-gathering mission in the San Diego Museum of Art. We selected this art museum for (i) its rich connectivity between galleries, giving an interesting example, and (ii) its small size, allowing us to visualize, interpret, and intuit the solution to a BOTOHE problem on it. We spatially model the art museum as a directed graph $G = (\mathscr{V}, \mathscr{E})$. The set of 27 nodes $\mathscr{V}$ represents the 23 art galleries (rooms), the outside entrance to the building (base node v_b), the main entrance rooms on the first and second floors, and the stairway. The set of arcs $\mathscr{E}$ represents direct passages/doorways between the rooms.

Suppose traversing the art museum is hazardous for the robots, owing to (i) adversarial security guards that (a) seek to prevent the robots from imaging the art and (b) possess the ability attack and/or capture the robots, and (ii) obstacles that the robots could (a) crash into or (b) become entangled in. To model risks of destruction or capture, we assign survival probabilities $\omega(i, j)$ for the arc(s) (i) traversing the staircase of 0.8, (ii) inside and in/out of the main entrance room of 0.9, (iii) on the right side of the first floor of 0.97, (iv) on the left side of the first floor of 0.95, and (v) on the second floor of 0.9.

Suppose, when a robot visits an art gallery in the museum, it images the art there and transmits this image back to the command center, providing utility to the command center. The utility of each image to the command center is scored by the node reward map r assigning rewards of (i) 2/3 for large galleries, (ii) 1/3 for medium-sized galleries, (iii) 1 for small galleries that we suppose contain the most valuable art, and (iv) 1/10 for five galleries that are in corners of the museum or behind the stairway.

292 The two objectives of the command center are to plan the trails of the robots in the art museum to maximize the (1) expected reward, via robots visiting art galleries, imaging them, then transmitting the images back to the command center, and (2) expected number of robots that return from the mission.

We visualize this BOTOHE problem instance in Fig. 3. The topology of the directed graph is shown, with a layout reflecting the spatial location of the rooms in the San Diego Museum of Art. The nodes on the first and second floor are grouped together. The base node is marked by the three robots near it. The arc survival probability map ω is visualized by the color assigned to each arc. The stairway is the most dangerous arc. The node reward map r is visualized by the color assigned to each node.

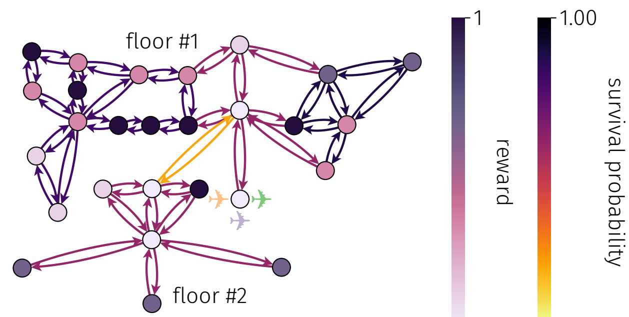


Figure 3. Our TOHE problem instance. The directed graph represents the two-floor San Diego art museum with distinct rooms (nodes) connected by doorways or a stairway (arcs). The arcs are colored according to robot survival probabilities. The nodes are colored according to rewards offered to the team when a robot visits. The three robots (planes) initially reside at the base node.

4.2 Computational results

We now employ our bi-objective ant optimization algorithm to search for the Pareto-optimal robot-team trail plans for our problem instance. We use a colony of $N_{ants} = 100$ artificial ants, a pheromone evaporation rate of $\rho = 0.04$, and 10 000 iterations. We initialize the pheromone maps with one unit of pheromone on each arc. The runtime is $\sim$5 min on an Apple M1 machine.

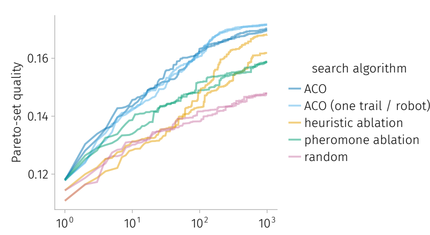
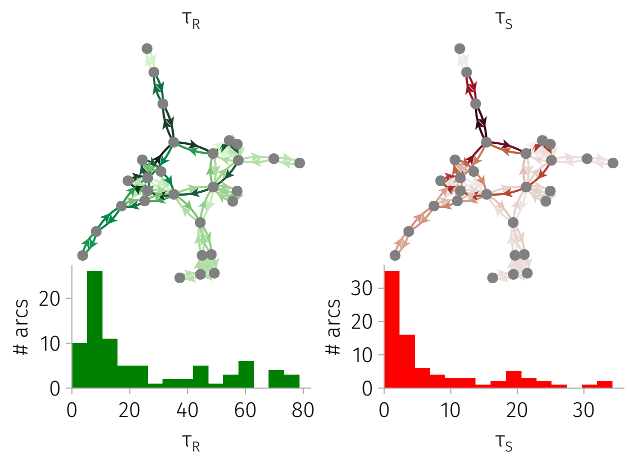


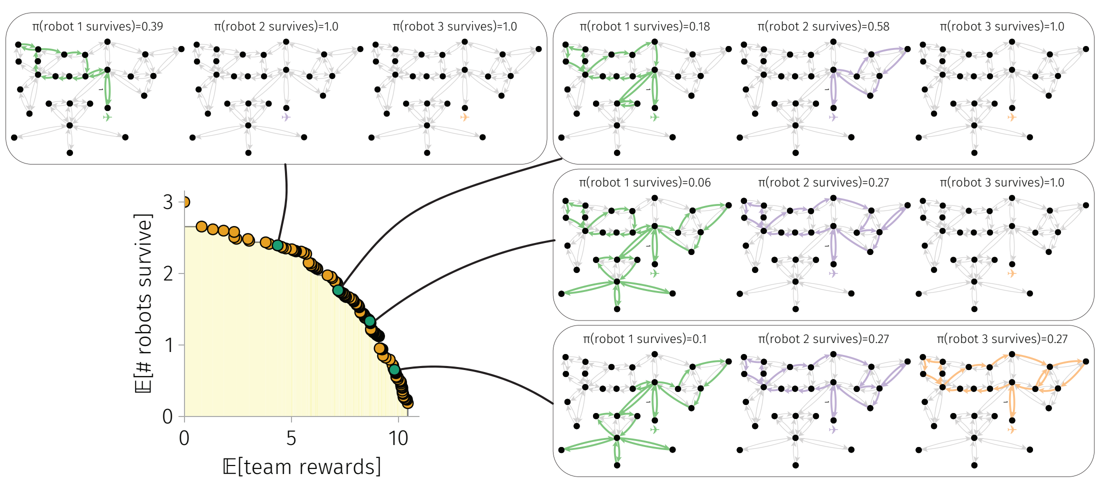
Figure 4. Progress of BO-ACO on our TOHE problem instance and comparison with baselines. The quality of the Pareto-set of robot team trail plans, measured via the area indicator, is shown as a function of the number of iterations of ACO—for ordinary ACO, ACO without heuristics, and ACO without pheromone.

Fig. 4 shows the quality (area indicator) of the Pareto-set (q in eqn. 16) over iterations. As the ACO algorithm progresses, the quality of the Pareto-set improves, but with diminishing returns. The saturation of the progress over iterations indicates satisfactory convergence.

ACO searches for Pareto-optimal robot-team trail plans, guided by (i) the static heuristics $\eta_{R,S}$ that greedily score the appeal of traversing a given arc and (ii) the dynamic pheromone maps $\tau_{R,S}$ that encapsulate the memory of the ant colony over previous iterations. Next, we quantify how each of these components are contributing to the effectiveness of ACO by ablating each. First, we run ACO where heuristics are not used to bias the ants towards *a priori* promising arcs by setting $\eta_{R,S}(v,v') = 1$ for all arcs $(v,v') \in \mathscr{E}$. Second, we run ACO where pheromone is not used to bias the ants towards promising arcs based on the history of the colony's search by setting $\tau_{R,S}(v,v') = 1$ for all arcs $(v,v') \in \mathscr{E}$. Fig. 4 compares the search efficiency of these ablated runs with ordinary ACO (where both heuristics and pheromone are used). Compared to ordinary ACO, the search efficiency diminishes for each ablation study. Thus, both heuristic and pheromone contribute to the search efficiency. However, the search efficiency drops more dramatically when the heuristic is ablated. Thus, the heuristic is more helpful than the pheromone in terms of finding a quality Pareto-set in few iterations.



(a) Pheromone trails



(b) Pareto-optimal solutions

Figure 5. Analysis of the BO-ACO solution to our TOHE problem instance. (a) The pheromone trails and distribution of the amount of pheromone on the arcs at the end of the ACO algorithm. (b) The [approximate] Pareto-front of robot-team team trail plans at the end of the ACO algorithm with four select plans shown. The indicator of the Pareto-set quality is the area of the highlighted, yellow region.

At the end of the ACO algorithm, Fig. 5a visualizes the pheromone maps τ_R and τ_S. Note the clock-wise vs. counter-close-wise preferences in some cycles of high pheromone to pursue high-reward nodes earlier in the trails to maximize expected reward.

Fig. 5b shows the [approximate] Pareto-front of 167 solutions found by BO-ACO. Walking down the Pareto-front trades larger expected team reward for a smaller expected number of robots that return from the mission. The key motivation for treating the *bi-objective* TOHE problem is to present this Pareto-front to a human decision-maker who ultimately chooses a robot-team trail plan that balances reward with robot survival.

Finally, Fig. 5b shows four Pareto-optimal trail plans belonging to different regions of the Pareto-front. As we move down the front, plans offer larger expected reward but lower robot survival. In the first plan, only a single robot enters the museum to image galleries offering the highest rewards on the left side of the first floor—avoiding the more dangerous right side of the second floor and the very dangerous staircase to the second floor. In the second plan, two robots are utilized: the first plans to cover much of the left side of the first floor, while the second plans to cover much of the right side of the first floor. In the third plan, also utilizing two robots, the first robot plans to traverse most of the first floor and the second floor later. The second robot traverses much of the first floor. This redundancy in first-floor coverage builds robustness into the plans: even if the first robot gets destroyed early on, the second robot can still image most of the galleries on the first floor. Finally, the fourth plan uses all three robots with much redundancy to achieve high expected reward. Still, the robots avoid taking the risk to enter the bottom left corner of the first floor, whose galleries offer only a small reward.

5 DISCUSSION

Many applications for mobile robot teams—ranging from information-gathering, resource-delivery, chemical plume source localization, forest fire fighting, to resource delivery—involve traversing environments with hazards e.g., corrosive chemicals, attacking adversaries, obstacles, or rough terrain/seas. For each application, the robots must coordinate their trails for the team-level objective in a risk-aware manner: select the subset of locations to visit, assign the safest trails to the robots, balance reward and risk to the robots, and build redundancy into the plans to make the objective robust to the failure of robots on the team.

Heavily inspired by the Team Surviving Orienteers Problem [50, 49, 51], we posed the bi-objective team orienteering in a hazardous environment (BOTHE) problem. By finding the Pareto-optimal set of robot-team trail plans, we can present them to a human decision-maker who ultimately chooses the robot trail plans for the mission according to how he or she values the expected reward collected by the robot team compared with the expected number of robots that survive the dangerous mission.

We employed bi-objective ant colony optimization to search for the Pareto-optimal team trail plans. Despite lacking theoretical guarantees to find the Pareto-optimal set, ACO was effective and can be readily adapted to handle extensions to the TOHE problem.

Future work. For the BOTOHE we have posed, we wish to (i) tackle the online version with ACO, where the robots adapt their planned trails during the mission, in response to observed failures of robots, (ii) devise local search methods to improve the robot trails the ants found, accelerating the convergence of ACO [32], (iii) benchmark non-sequential methods to allocate trails to robots with ACO [54], and (iv) employ multi-colony ant optimization [47].

Interesting and practical extensions of robot-team orienteering in adversarial/hazardous environments abstracted a graphs include treating (some of these ideas from Ref. [50]): (i) a heterogenous team of robots with different (a) capabilities to harvest rewards from nodes and (b) survival probabilities for each arc traversal owing to e.g. stealth; (ii) more complicated reward structures, e.g., time-dependent, stochastic, non-additive (correlated [100]), multi-category, or multi-visit rewards; (iii) fuel/battery constraints of the robots and nodes representing refueling, recharging, or battery-switching stations [2, 55, 59, 101]; (iv) constraints on the rewards a robot can harvest e.g. for resource delivery applications where each robot holds limited cargo capacity [22]; (v) non-binary surviving states of the robots due to various levels of damage; (vi) non-independent events of robots surviving arc-traversals; (vii) risk metrics different from the expected value [61].

Another interesting direction is to learn/modify the survival probabilities associated with the edges of the graph from data over repeated missions (an inverse problem [12]). Specifically, suppose we are uncertain about the survival probability $\omega(i,j)$ of each arc (i,j). Within a Bayesian inference framework, we may impose a prior distribution on each $\omega(i,j)$. Then, when a robot survives or gets destroyed during a mission, we update the prior distribution. The trail-planning of the robots over sequential missions may then balance (a) exploitation to harvest the most reward and take what appear to be, under uncertainty, the safest trails and (b) exploration to find even safer trails.

Finally, instead of abstracting the environment as a graph, one could path-plan for robots in a continuous space with obstacles.

ACKNOWLEDGEMENTS

CMS acknowledges the Office of Naval Research Summer Faculty Research Program.

REFERENCES

[1] Agmon, N. (2017). Robotic strategic behavior in adversarial environments. In *IJCAI*, pages 5106–5110.

[2] Asghar, A. B., Shi, G., Karapetyan, N., Humann, J., Reddinger, J.-P., Dotterweich, J., and Tokekar, P. (2023). Risk-aware recharging rendezvous for a collaborative team of UAVs and UGVs. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5544–5550. IEEE.

[3] Basilico, N. (2022). Recent trends in robotic patrolling. *Current Robotics Reports*, 3(2):65–76.

[4] Bawden, O., Kulk, J., Russell, R., McCool, C., English, A., Dayoub, F., Lehnert, C., and Perez, T. (2017). Robot for weed species plant-specific management. *Journal of Field Robotics*, 34(6):1179–1199.

[5] Beckers, R., Deneubourg, J.-L., and Goss, S. (1993). Modulation of trail laying in the ant Lasius niger (hymenoptera: Formicidae) and its role in the collective selection of a food source. *Journal of Insect Behavior*, 6:751–759.

[6] Blum, C. (2005). Ant colony optimization: Introduction and recent trends. *Physics of Life Reviews*, 2(4):353–373.

[7] Bonabeau, E., Dorigo, M., and Theraulaz, G. (1999). *Swarm Intelligence: From Natural to Artificial Systems*. Number 1. Oxford university press.

[8] Bonabeau, E., Dorigo, M., and Theraulaz, G. (2000). Inspiration for optimization from social insect behaviour. *Nature*, 406(6791):39–42.

[9] Bonabeau, E., Theraulaz, G., Deneubourg, J.-L., Aron, S., and Camazine, S. (1997). Self-organization in social insects. *Trends in Ecology & Evolution*, 12(5):188–193.

[10] Brambilla, M., Ferrante, E., Birattari, M., and Dorigo, M. (2013). Swarm robotics: a review from the swarm engineering perspective. *Swarm Intelligence*, 7:1–41.

[11] Branke, J., Deb, K. D., Miettinen, K., and Słowiński, R. (2008). *Multiobjective optimization: Interactive and evolutionary approaches*. Springer Science & Business Media.

[12] Burton, D. and Toint, P. L. (1992). On an instance of the inverse shortest paths problem. *Mathematical Programming*, 53:45–61.

[13] Butt, S. E. and Cavalier, T. M. (1994). A heuristic for the multiple tour maximum collection problem. *Computers & Operations Research*, 21(1):101–111.

[14] Cao, Y., Smucker, B. J., and Robinson, T. J. (2015). On using the hypervolume indicator to compare Pareto fronts: Applications to multi-criteria optimal experimental design. *Journal of Statistical Planning and Inference*, 160:60–74.

[15] Chao, I.-M., Golden, B. L., and Wasil, E. A. (1996a). A fast and effective heuristic for the orienteering problem. *European Journal of Operational Research*, 88(3):475–489.

[16] Chao, I.-M., Golden, B. L., and Wasil, E. A. (1996b). The team orienteering problem. *European Journal of Operational Research*, 88(3):464–474.

[17] Chen, J., Zeng, S., and Xu, X. (2022). Environment-aware path planning for UAV-assisted search and rescue. In *2022 14th International Conference on Wireless Communications and Signal Processing (WCSP)*, pages 922–927. IEEE.

[18] Chen, Y.-H., Sun, W.-J., and Chiang, T.-C. (2015). Multiobjective orienteering problem with time windows: An ant colony optimization algorithm. In *2015 Conference on Technologies and Applications of Artificial Intelligence (TAAI)*, pages 128–135. IEEE.

[19] Choi, H.-T. and Yuh, J. (2016). Underwater robots. *Springer Handbook of Robotics*, pages 595–622.

[20] Chung, W. and Iagnemma, K. (2016). Wheeled robots. *Springer Handbook of Robotics*, pages 575–594.

[21] Clark, J. and Holton, D. A. (1991). *A First Look at Graph Theory*. World Scientific.

[22] Coelho, L. C., Cordeau, J.-F., and Laporte, G. (2014). Thirty years of inventory routing. *Transportation Science*, 48(1):1–19.

[23] Costa, M. M. and Silva, M. F. (2019). A survey on path planning algorithms for mobile robots. In *2019 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)*, pages 1–7. IEEE.

[24] Czaczkes, T. J., Grüter, C., Ellis, L., Wood, E., and Ratnieks, F. L. (2013). Ant foraging on complex trails: route learning and the role of trail pheromones in lasius niger. *Journal of Experimental Biology*, 216(2):188–197.

[25] Czaczkes, T. J., Grüter, C., and Ratnieks, F. L. (2015). Trail pheromones: an integrative view of their role in social insect colony organization. *Annual Review of Entomology*, 60:581–599.

[26] Dang, D.-C., Guibadj, R. N., and Moukrim, A. (2013). An effective pso-inspired algorithm for the team orienteering problem. *European Journal of Operational Research*, 229(2):332–344.

[27] David Morgan, E. (2009). Trail pheromones of ants. *Physiological Entomology*, 34(1):1–17.

[28] Deneubourg, J. L., Aron, S., Goss, S., Pasteels, J., and Duerinck, G. (1986). Random behaviour, amplification processes and number of participants: how they contribute to the foraging properties of ants. *Physica D: Nonlinear Phenomena*, 22(1-3):176–186.

[29] Deneubourg, J. L., Aron, S., Goss, S., and Pasteels, J. M. (1990). The self-organizing exploratory pattern of the Argentine ant. *Journal of Insect Behavior*, 3:159–168.

[30] Deneubourg, J.-L., Pasteels, J. M., and Verhaeghe, J.-C. (1983). Probabilistic behaviour in ants: a strategy of errors? *Journal of Theoretical Biology*, 105(2):259–271.

[31] Di, K., Zhou, Y., Yan, F., Jiang, J., Yang, S., and Jiang, Y. (2022). A foraging strategy with risk response for individual robots in adversarial environments. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 13(5):1–29.

[32] Dorigo, M., Birattari, M., and Stutzle, T. (2006). Ant colony optimization. *IEEE Computational Intelligence Magazine*, 1(4):28–39.

[33] Dorigo, M., Maniezzo, V., and Colorni, A. (1996). Ant system: optimization by a colony of cooperating agents. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 26(1):29–41.

[34] Dunbabin, M. and Marques, L. (2012). Robots for environmental monitoring: Significant advancements and applications. *IEEE Robotics & Automation Magazine*, 19(1):24–39.

[35] Dussutour, A., Beekman, M., Nicolis, S. C., and Meyer, B. (2009). Noise improves collective decision-making by ants in dynamic environments. *Proceedings of the Royal Society B: Biological Sciences*, 276(1677):4353–4361.

[36] Edelstein-Keshet, L., Watmough, J., and Ermentrout, G. B. (1995). Trail following in ants: individual properties determine population behaviour. *Behavioral Ecology and Sociobiology*, 36:119–133.

[37] Evison, S. E., Petchey, O. L., Beckerman, A. P., and Ratnieks, F. L. (2008). Combined use of pheromone trails and visual landmarks by the common garden ant Lasius niger. *Behavioral Ecology and Sociobiology*, 63:261–267.

[38] Francis, A., Li, S., Griffiths, C., and Sienz, J. (2022). Gas source localization and mapping with mobile robots: A review. *Journal of Field Robotics*, 39(8):1341–1373.

[39] Gavalas, D., Konstantopoulos, C., Mastakas, K., and Pantziou, G. (2014). A survey on algorithmic approaches for solving tourist trip design problems. *Journal of Heuristics*, 20(3):291–328.

[40] Golden, B. L., Levy, L., and Vohra, R. (1987). The orienteering problem. *Naval Research Logistics (NRL)*, 34(3):307–318.

[41] Goss, S., Aron, S., Deneubourg, J.-L., and Pasteels, J. M. (1989). Self-organized shortcuts in the Argentine ant. *Naturwissenschaften*, 76(12):579–581.

[42] Guerreiro, A. P., Fonseca, C. M., and Paquete, L. (2020). The hypervolume indicator: Problems and algorithms. *arXiv preprint arXiv:2005.00515*.

[43] Gunawan, A., Lau, H. C., and Vansteenwegen, P. (2016). Orienteering problem: A survey of recent variants, solution approaches and applications. *European Journal of Operational Research*, 255(2):315–332.

[44] Hernandez Bennetts, V., Lilienthal, A. J., Neumann, P. P., and Trincavelli, M. (2012). Mobile robots for localizing gas emission sources on landfill sites: is bio-inspiration the way to go? *Frontiers in Neuroengineering*, 4:20.

[45] Huang, L., Zhou, M., Hao, K., and Hou, E. (2019). A survey of multi-robot regular and adversarial patrolling. *IEEE/CAA Journal of Automatica Sinica*, 6(4):894–903.
[46] Hutchinson, M., Liu, C., Thomas, P., and Chen, W.-H. (2019). Unmanned aerial vehicle-based hazardous materials response: Information-theoretic hazardous source search and reconstruction. *IEEE Robotics & Automation Magazine*, 27(3):108–119.
[47] Iredi, S., Merkle, D., and Middendorf, M. (2001). Bi-criterion optimization with multi colony ant algorithms. In *Evolutionary Multi-Criterion Optimization: First International Conference, EMO 2001 Zurich, Switzerland, March 7–9, 2001 Proceedings 1*, pages 359–372. Springer.
[48] Jackson, D. E. and Ratnieks, F. L. (2006). Communication in ants. *Current Biology*, 16(15):R570–R574.
[49] Jorgensen, S., Chen, R. H., Milam, M. B., and Pavone, M. (2017). The matroid team surviving orienteers problem: Constrained routing of heterogeneous teams with risky traversal. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5622–5629. IEEE.
[50] Jorgensen, S., Chen, R. H., Milam, M. B., and Pavone, M. (2018). The team surviving orienteers problem: routing teams of robots in uncertain environments with survival constraints. *Autonomous Robots*, 42:927–952.
[51] Jorgensen, S. and Pavone, M. (2024). The matroid team surviving orienteers problem and its variants: Constrained routing of heterogeneous teams with risky traversal. *The International Journal of Robotics Research*, 43(1):34–52.
[52] Kamminga, J., Ayele, E., Meratnia, N., and Havinga, P. (2018). Poaching detection technologies—a survey. *Sensors*, 18(5):1474.
[53] Karur, K., Sharma, N., Dharmatti, C., and Siegel, J. E. (2021). A survey of path planning algorithms for mobile robots. *Vehicles*, 3(3):448–468.
[54] Ke, L., Archetti, C., and Feng, Z. (2008). Ants can solve the team orienteering problem. *Computers & Industrial Engineering*, 54(3):648–665.
[55] Khuller, S., Malekian, A., and Mestre, J. (2011). To fill or not to fill: The gas station problem. *ACM Transactions on Algorithms (TALG)*, 7(3):1–16.
[56] Knaden, M. and Graham, P. (2016). The sensory ecology of ant navigation: from natural environments to neural mechanisms. *Annual Review of Entomology*, 61:63–76.
[57] Korngut, Y. and Agmon, N. (2023). Multi-robot heterogeneous adversarial coverage. In *2023 International Symposium on Multi-Robot and Multi-Agent Systems (MRS)*, pages 100–106. IEEE.
[58] Leutenegger, S., Hürzeler, C., Stowers, A. K., Alexis, K., Achtelik, M. W., Lentink, D., Oh, P. Y., and Siegwart, R. (2016). Flying robots. *Springer Handbook of Robotics*, pages 623–670.
[59] Liao, C.-S., Lu, S.-H., and Shen, Z.-J. M. (2016). The electric vehicle touring problem. *Transportation Research Part B: Methodological*, 86:163–180.
[60] Liu, J., Zhou, L., Tokekar, P., and Williams, R. K. (2021). Distributed resilient submodular action selection in adversarial environments. *IEEE Robotics and Automation Letters*, 6(3):5832–5839.
[61] Majumdar, A. and Pavone, M. (2020). How should a robot assess risk? towards an axiomatic theory of risk in robotics. In *Robotics Research: The 18th International Symposium ISRR*, pages 75–84. Springer.
[62] Mayya, S., Ramachandran, R. K., Zhou, L., Senthil, V., Thakur, D., Sukhatme, G. S., and Kumar, V. (2022). Adaptive and risk-aware target tracking for robot teams with heterogeneous sensors. *IEEE Robotics and Automation Letters*, 7(2):5615–5622.
[63] McAllister, W., Osipychev, D., Chowdhary, G., and Davis, A. (2018). Multi-agent planning for coordinated robotic weed killing. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 7955–7960. IEEE.
[64] Merino, L., Caballero, F., Martínez-de Dios, J. R., Maza, I., and Ollero, A. (2012). An unmanned aircraft system for automatic forest fire monitoring and measurement. *Journal of Intelligent & Robotic Systems*, 65:533–548.
[65] Monaci, M., Pike-Burke, C., and Santini, A. (2022). Exact algorithms for the 0–1 time-bomb knapsack problem. *Computers & Operations Research*, 145:105848.
[66] Montemanni, R. and Smith, D. H. (2025). An exact algorithm for the hazardous orienteering problem. *Networks*.
[67] Montemanni, R., Weyland, D., and Gambardella, L. (2011). An enhanced ant colony system for the team orienteering problem with time windows. In *2011 international symposium on computer science*

and society, pages 381–384. IEEE.

[68] Murphy, R. R., Peschel, J., Arnett, C., and Martin, D. (2012). Projected needs for robot-assisted chemical, biological, radiological, or nuclear (CBRN) incidents. In *2012 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*, pages 1–4. IEEE.

[69] Pardalos, P. M., Žilinskas, A., and Žilinskas, J. (2017). *Non-convex multi-objective optimization*. Springer.

[70] Parker, L. E. (1995). On the design of behavior-based multi-robot teams. *Advanced Robotics*, 10(6):547–578.

[71] Parker, L. E. (2007). Distributed intelligence: Overview of the field and its application in multi-robot systems. In *AAAI fall symposium: regarding the intelligence in distributed intelligent systems*, pages 1–6.

[72] Prorok, A., Malencia, M., Carlone, L., Sukhatme, G. S., Sadler, B. M., and Kumar, V. (2021). Beyond robustness: A taxonomy of approaches towards resilient multi-robot systems. *arXiv preprint arXiv:2109.12343*.

[73] Queralta, J. P., Taipalmaa, J., Pullinen, B. C., Sarker, V. K., Gia, T. N., Tenhunen, H., Gabbouj, M., Raitoharju, J., and Westerlund, T. (2020). Collaborative multi-robot search and rescue: Planning, coordination, perception, and active vision. *IEEE Access*, 8:191617–191643.

[74] Robin, C. and Lacroix, S. (2016). Multi-robot target detection and tracking: taxonomy and survey. *Autonomous Robots*, 40:729–760.

[75] Robinson, E. J., Jackson, D. E., Holcombe, M., and Ratnieks, F. L. (2005). 'no entry' signal in ant foraging. *Nature*, 438(7067):442–442.

[76] Rouček, T., Pecka, M., Čížek, P., Petříček, T., Bayer, J., Šalanskỳ, V., Heřt, D., Petrlík, M., Báča, T., Spurnỳ, V., et al. (2020). DARPA subterranean challenge: Multi-robotic exploration of underground environments. In *Modelling and Simulation for Autonomous Systems: 6th International Conference, MESAS 2019, Palermo, Italy, October 29–31, 2019, Revised Selected Papers 6*, pages 274–290. Springer.

[77] Sánchez-Ibáñez, J. R., Pérez-del Pulgar, C. J., and García-Cerezo, A. (2021). Path planning for autonomous mobile robots: A review. *Sensors*, 21(23):7898.

[78] Santini, A. and Archetti, C. (2023). The hazardous orienteering problem. *Networks*, 81(2):235–252.

[79] Santos, L. C., Santos, F. N., Pires, E. S., Valente, A., Costa, P., and Magalhães, S. (2020). Path planning for ground robots in agriculture: A short review. In *2020 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)*, pages 61–66. IEEE.

[80] Schranz, M., Umlauft, M., Sende, M., and Elmenreich, W. (2020). Swarm robotic behaviors and current applications. *Frontiers in Robotics and AI*, page 36.

[81] Schwager, M., Dames, P., Rus, D., and Kumar, V. (2017). A multi-robot control policy for information gathering in the presence of unknown hazards. In *Robotics Research: The 15th International Symposium ISRR*, pages 455–472. Springer.

[82] Shapira, Y. and Agmon, N. (2015). Path planning for optimizing survivability of multi-robot formation in adversarial environments. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4544–4549. IEEE.

[83] Sherali, H. D., Brizendine, L. D., Glickman, T. S., and Subramanian, S. (1997). Low probability—high consequence considerations in routing hazardous material shipments. *Transportation Science*, 31(3):237–251.

[84] Shi, G., Zhou, L., and Tokekar, P. (2023). Robust multiple-path orienteering problem: Securing against adversarial attacks. *IEEE Transactions on Robotics*.

[85] Shiraishi, M., Takeuchi, R., Nakagawa, H., Nishimura, S. I., Awazu, A., and Nishimori, H. (2019). Diverse stochasticity leads a colony of ants to optimal foraging. *Journal of Theoretical Biology*, 465:7–16.

[86] Sohrabi, S., Ziarati, K., and Keshtkaran, M. (2021). ACS-OPHS: Ant colony system for the orienteering problem with hotel selection. *EURO Journal on Transportation and Logistics*, 10:100036.

[87] Soldan, S., Welle, J., Barz, T., Kroll, A., and Schulz, D. (2014). Towards autonomous robotic systems for remote gas leak detection and localization in industrial environments. In *Field and Service Robotics: Results of the 8th International Conference*, pages 233–247. Springer.

[88] Tang, W. and Tang, F. (2023). The Poisson binomial distribution—old & new. *Statistical Science*, 38(1):108–119.

[89] Trevelyan, J., Hamel, W. R., and Kang, S.-C. (2016). Robotics in hazardous applications. *Springer Handbook of Robotics*, pages 1521–1548.

[90] Van Oudenhove, L., Billoir, E., Boulay, R., Bernstein, C., and Cerdá, X. (2011). Temperature limits trail following behaviour through pheromone decay in ants. *Naturwissenschaften*, 98:1009–1017.

[91] Vansteenwegen, P., Souffriau, W., and Van Oudheusden, D. (2011). The orienteering problem: A survey. *European Journal of Operational Research*, 209(1):1–10.

[92] Verbeeck, C., Vansteenwegen, P., and Aghezzaf, E.-H. (2017). The time-dependent orienteering problem with time windows: a fast ant colony system. *Annals of Operations Research*, 254:481–505.

[93] Watmough, J. and Edelstein-Keshet, L. (1995). Modelling the formation of trail networks by foraging ants. *Journal of Theoretical Biology*, 176(3):357–371.

[94] Wendt, S., Kleinhoelting, N., and Czaczkes, T. J. (2020). Negative feedback: ants choose unoccupied over occupied food sources and lay more pheromone to them. *Journal of The Royal Society Interface*, 17(163):20190661.

[95] Wilson, R. J. (1979). *Introduction to graph theory*. Pearson Education India.

[96] Wurman, P. R., D'Andrea, R., and Mountz, M. (2008). Coordinating hundreds of cooperative, autonomous vehicles in warehouses. *AI Magazine*, 29(1):9–9.

[97] Yadav, S. K. (2023). *Advanced Graph Theory*. Springer.

[98] Yang, L., Li, P., Qian, S., Quan, H., Miao, J., Liu, M., Hu, Y., and Memetimin, E. (2023). Path planning technique for mobile robots: A review. *Machines*, 11(10):980.

[99] Yehoshua, R., Agmon, N., and Kaminka, G. A. (2016). Robotic adversarial coverage of known environments. *The International Journal of Robotics Research*, 35(12):1419–1444.

[100] Yu, J., Schwager, M., and Rus, D. (2014). Correlated orienteering problem and its application to informative path planning for persistent monitoring tasks. In *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 342–349. IEEE.

[101] Yu, K., O'Kane, J. M., and Tokekar, P. (2019). Coverage of an environment using energy-constrained unmanned aerial vehicles. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 3259–3265. IEEE.

[102] Yuan, H., Xiao, C., Wang, Y., Peng, X., Wen, Y., and Li, Q. (2020). Maritime vessel emission monitoring by an UAV gas sensor system. *Ocean Engineering*, 218:108206.

[103] Zhou, L. and Tokekar, P. (2021). Multi-robot coordination and planning in uncertain and adversarial environments. *Current Robotics Reports*, 2:147–157.

[104] Zhou, L., Tzoumas, V., Pappas, G. J., and Tokekar, P. (2018). Resilient active target tracking with multiple robots. *IEEE Robotics and Automation Letters*, 4(1):129–136.

[105] Zhou, L., Tzoumas, V., Pappas, G. J., and Tokekar, P. (2022). Distributed attack-robust submodular maximization for multirobot planning. *IEEE Transactions on Robotics*, 38(5):3097–3112.