

# Model-Predictive Target Defense by Team of Unmanned Surface Vehicles Operating in Uncertain Environments

Eric Raboin<sup>1</sup>, Petr Švec<sup>2</sup>, *Member, IEEE*, Dana Nau<sup>3</sup>, and Satyandra K. Gupta<sup>4</sup>, *Member, IEEE*

**Abstract**—In this paper, we present a heuristic planning approach for guarding a valuable asset by a team of autonomous unmanned surface vehicles (USVs) operating in a continuous state-action space. The team’s objective is to maximize the amount of time it takes an intruder boat to reach the asset. The team must cooperatively deal with uncertainty about which boats are actual intruders, employ active blocking to slow down intruders’ movement towards the asset, and intelligently distribute themselves around the target to optimize future guarding opportunities. Our planner incorporates a market-based algorithm for allocating tasks to individual USVs by forward-simulating the mission and assigning estimated utilities to candidate task-allocation plans. The planner can be automatically adapted to a specific mission by optimizing the behaviors used to fulfil individual tasks. We present detailed simulation results that demonstrate the effectiveness of our approach.

## I. INTRODUCTION

Technological progress in the development of autonomy for unmanned surface vehicles (USVs) [1] is enabling unmanned boats to be used for guarding of sensitive areas in naval missions. The use of autonomous USVs for protecting an asset against intruder boats can lead to significant cost reduction while preserving the required level of security. This application, however, presents multiple challenges for the team of USVs from the planning perspective.

Guarding an asset requires the team of USVs to cooperatively patrol the area around the asset, observe passing boats, identify intruders, and delay their progress towards the target by active blocking (see Fig. 1). The vehicles have to make intelligent, balanced decisions about which tasks to perform in order to prevent intruders from attacking the target without being blocked. This presents a non-trivial challenge for the planning algorithm since the identity of the boats may not be known at the time they enter the visibility range of the USVs. Furthermore, the planning must be done efficiently despite the very large state-action space since multiple tasks can be assigned to multiple agents simultaneously. The planner has to also consider time dependencies since selecting the tasks requires knowledge of what future tasks are possible in order to maximize expected performance. Finally, the developed approach should be usable in a range of scenarios.

<sup>1</sup>E. Raboin is with the Department of Computer Science, University of Maryland, College Park, MD 20742, USA [eraboin@umd.edu](mailto:eraboin@umd.edu)

<sup>2</sup>P. Švec is with the Department of Mechanical Engineering, University of Maryland, College Park, MD 20742, USA [petrsvec@umd.edu](mailto:petrsvec@umd.edu)

<sup>3</sup>D. Nau is with the Department of Computer Science and Institute for Systems Research, University of Maryland, College Park, MD 20742, USA [nau@umd.edu](mailto:nau@umd.edu)

<sup>4</sup>S.K. Gupta is with the Department of Mechanical Engineering and Institute for Systems Research, University of Maryland, College Park, MD 20742, USA [skgupta@umd.edu](mailto:skgupta@umd.edu)

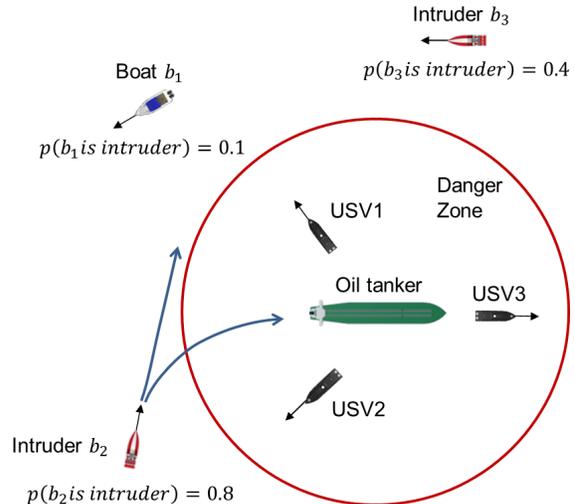


Fig. 1: A team of unmanned surface vehicles (USVs) guard an oil tanker against intruder boats. Each boat is assigned a probability of being an intruder based on observations made by the USV team.

The problem’s complexity merits solutions at multiple levels, including high-level task planning approaches, trajectory planning for collision-free guidance [2], [3], [4], machine learning for automated synthesis of behaviors for intruder interception [5], and generation of state transition models using GPU-accelerated simulation [4].

This paper focuses on high-level task planning and behavior optimization for a team of USVs to guard a valuable asset. The developed heuristic planning approach is able to deal with uncertainty about which boats are actual intruders. It computes an approximate solution to an instance of the MT-MR-TA (i.e., multi-task robots, multi-robot tasks, time-extended allocation) variant of the task allocation problems [6] in real-time. The individual tasks are assigned to USVs incrementally using market-based exchanges [7] between the vehicles. The task allocation is evaluated using model-predictive simulation, i.e. by looking-ahead and estimating the utility of the allocation in order to optimize the assignment of future tasks based on the current state of the boats in the scene. Each task is executed by a corresponding parameterized behavior that is optimized for specific properties of the mission (i.e., the number of available USVs, an estimated number of intruders, spatial distribution of the coming boats in respect to the target, etc.). The behaviors are optimized

for all vehicles concurrently to account for their individual contributions to the guarding strategy.

## II. RELATED WORK

We review representative approaches for task allocation, multi-robot patrolling, and learning of cooperative behaviors.

From the task planning perspective, a mission is either manually or automatically divided into tasks, hierarchical task tree structures, or roles defined for robots. The tasks are allocated to robots based on a number of factors according to the multirobot task allocation (MRTA) taxonomy in [6]. The particular factors include the number of tasks that can be performed by a single robot (i.e., ST as single-task robots, and MT as multi-task robots), number of robots that are required to fulfil a task (i.e., SR as single-robot tasks, and MR as multi-robot tasks), and whether the current assignment of tasks is optimized for future tasks or not (i.e., IA as instantaneous task allocation, and TA as time-extended allocation). Mostly related to our work is MT-MR-TA variant that is known to be  $\mathcal{NP}$ -hard making real-time computation of the optimal task allocation infeasible.

The core techniques developed for solving MRTA problems can be categorized (1) into market-based and behavior-based approaches, (2) based on their ability to allocate simple or complex tasks, and (3) whether they decompose first and then allocate or vice versa, or approaches that do not separate the two phases. Our task allocation algorithm is most closely related to the market-based group of approaches that have low computational requirements compared to the centralized approaches [8] and compute more close-to-optimal solutions than distributed approaches [9]. A thorough survey on the current state of the art in market-based techniques for multirobot task allocation is given in [10]. These include MURDOCH [11], TraderBots [12], and Hoplites [13].

A survey of the current state-of-the-art patrolling algorithms is provided in [14]. The representative approaches are evaluated in detail in [15] in terms of the average idleness of a patrolling graph and scalability to the number of agents metrics. In our approach, the patrolling strategy is computed indirectly through the market-based exchange of guard tasks commanding the vehicles to computed waypoints or predefined patrolling locations.

In the USV domain, Simetti et al. [16] developed a heuristic approach for a team of USVs to intercept an intruder. The approach selects the best USV to intercept the detected intruder while considering obstacles in the scene. The positions of the USVs are optimized using a combination of Monte Carlo and gradient descent algorithms. Zhang and Meng [17] developed a distributed, STAGS heuristic approach for multi-USV target defense. The approach utilizes a motivational, behavior-based algorithm for task allocation. Deployment of USVs around the target is handled by a heuristic, self-deployment algorithm based on dynamically created gaps. The parameters of the approach are optimized using multi-objective optimization to minimize the average response time (i.e., the time to investigation since the detection of the intruders by static sensors) and missing rate.

In our approach, we have developed two-side share and offer types of contracts as the tools for marginal cost based contracting [18], allowing decentralized task negotiation among agents. Other contributions of our work include explicit consideration of uncertainty in recognizing intruder boats (i.e., due to limited sensing). Moreover, the outlined task planning and allocation problem does not allow us to explicitly compute utilities of individual agent-task assignment pairs because of the dependency of these utilities on the task assignments of other agents. Hence, our market-based algorithm is driven by a user-defined objective function that evaluates the utility of a particular, collective task assignment. We use model-predictive simulation to evaluate candidate task allocation plans, which allows robots to explicitly consider the tasks of its surrounding robots when evaluating candidate plans, and to perform time-extended allocation of tasks. Finally, due to the market-based nature of the developed algorithm, task negotiation can be terminated at any time and still provide a reasonable solution [19].

## III. PROBLEM FORMULATION

We define a multi-agent planning problem where a team of USVs must defend a stationary target against a team of hostile intruders. The USV team's objective is to delay the hostile boats' arrival at the target. More formally, given,

- (i) a team of USVs  $U = \{u_1, u_2, \dots, u_m\}$  protecting a target positioned at location  $l_{target}$  where  $v_{u_i}$  is the maximum surge speed of  $u_i$
- (ii) a set of passing boats  $B = \{b_1, b_2, \dots, b_n\}$  including intruder boats  $X \subseteq B$  where  $v_{b_i}$  is the maximum surge speed of  $b_i$
- (iii) the state of the world  $s = \{l_{u_1}, \dots, l_{u_m}, l_{b_1}, \dots, l_{b_n}\}$  defining the location  $l_{u_i}$  of every USV  $u_i$  and location  $l_{b_i}$  of every boat  $b_i$ ,
- (iv) an observation history  $O = \{o_{t_1}, o_{t_2}, \dots, o_{t_k}\}$  where each  $o_{t_i} = \langle s_{t_i}, F_{t_i} \rangle$  represents the observations made by the USV team at time  $t_i$ , where  $s_{t_i}$  is the state of world and  $F_{t_i} = \{f_{b_1}, f_{b_2}, \dots, f_{b_n}\}$  is a set of observed features  $f_{b_j}$  (e.g. size, color) for each boat  $b_j$ ,
- (v) a set of *observe* tasks,  $H_o$ , where USVs are responsible for gathering information about the passing boats; a set of *guard* tasks,  $H_g$ , where USVs must position themselves in vulnerable areas around the target; and a set of *delay* tasks,  $H_d$ , where USVs must intercept and then block a hostile USV,
- (vi) an observation classification function  $P(b_i \in X|O)$  that returns the probability that boat  $b_i$  is an intruder given observation history  $O$
- (vii) a non-deterministic opponent model  $\pi_{b_i}(O)$  that returns a velocity vector  $\mathbf{v}$  for boat  $b_i$ , defining the behavior of boat  $b_i$  given observation history  $O$
- (viii) a blocking function,  $v_{block}(b_i, n) = v' \in [0, v_{b_i}]$  that returns the maximum achievable surge speed of boat  $b_i$  when it is blocked by  $n$  different USVs,
- (ix) a response team probability threshold  $p_{alert}$ , indicating at what probability  $P(b_i \in X|O)$  an alert should be triggered for boat  $b_i$

Compute,

- (i.) a joint task allocation  $\mathcal{A} = \{H_{u_1}, H_{u_2}, \dots, H_{u_m}\}$  for the USV team, where  $H_{u_i} \subseteq H_o \cup H_g \cup H_d$  is the set of tasks assigned to USV  $u_i$ . Each guard or observe task may only be assigned to one USV at a time, while delay tasks may be assigned to multiple USVs.
- (ii.) a policy  $\pi_{u_i}(O, \mathcal{A})$  returning a velocity vector  $\mathbf{v}$  that defines the behavior of USV  $u_i$  given the current observation history  $O$  and task allocation  $\mathcal{A}$ .

The USV team does not know *a priori* which boats are hostile, but can determine whether a boat is hostile through observation. The features  $f_{b_i} \in F_{t_j}$  in the USV team's observation history are used by the classification function  $P(b_i \in X|O)$  to determine the probability that boat  $b_i$  is an intruder. We assume that this function is given, and that its exact nature will vary depending on the scenario.

If at any time the probability  $P(b_i \in X|O)$  exceeds  $p_{alert}$  for any boat  $b_i$ , then an alert is triggered. The *time until arrival*,  $t_\delta$ , is the difference between the time  $t_{alert}$  that an alert was triggered, and the time  $t_{arrival}$  when an intruder arrives at the target. The objective of each USV is to find a task allocation  $\mathcal{A}^*$  and policy  $\pi_{u_i}^*$  that maximize the expected  $t_\delta$  for the first boat  $b_j$  that reaches the target,

$$\langle \mathcal{A}^*, \pi_{u_i}^* \rangle = \arg \max_{\mathcal{A}, \pi_{u_i}} E[t_\delta | \pi_{u_i}(O, \mathcal{A})]. \quad (1)$$

#### IV. APPROACH

The joint task allocation,  $\mathcal{A}$ , is computed online and updated during each planning time step. Our algorithm uses both heuristics and model-predictive simulation to determine which candidate task allocation is selected. We assume that the USVs have full communication during this process.

The actual policy for each USV,  $\pi_{u_i}(O, \mathcal{A})$ , takes the task allocation and observation history as input and generates an appropriate velocity vector. This policy is implemented using parameterized behaviors that have been tuned offline by a genetic algorithm. These low-level behaviors are defined below, followed by a description of our algorithm for high-level task allocation, a description of the model-predictive simulation process, and an overview of how we utilize of genetic optimization.

##### A. Behaviors

The velocity vector returned by  $\pi_{u_i}(O, \mathcal{A})$  is computed by blending together the motion goals of lower-level policies for the guard, observe and delay tasks in  $u_i$ 's task assignment  $H_{u_i} \in \mathcal{A}$ . The motion goal for each individual task  $h_j \in H_{u_i}$  is given by  $M_{h_j}(O, \mathcal{A})$ , defined below,

$$M_{h_j}(O, \mathcal{A}) = \begin{cases} \text{a boat's location, } l_{b_j}, & \text{if } h_j \in H_o, \\ \text{a guard location, } l_{g_j}, & \text{if } h_j \in H_g, \\ l_{opt}(u_i, b_j, O, \mathcal{A}), & \text{if } h_j \in H_d, \end{cases} \quad (2)$$

where  $l_{opt}(u_i, b_j, O, \mathcal{A})$  is the optimal intercept point for USV  $u_i$  to intercept boat  $b_j$ , given the set of USVs that are assigned to delay  $b_j$  in  $\mathcal{A}$ . An example intercept path is shown in Figure 2b.

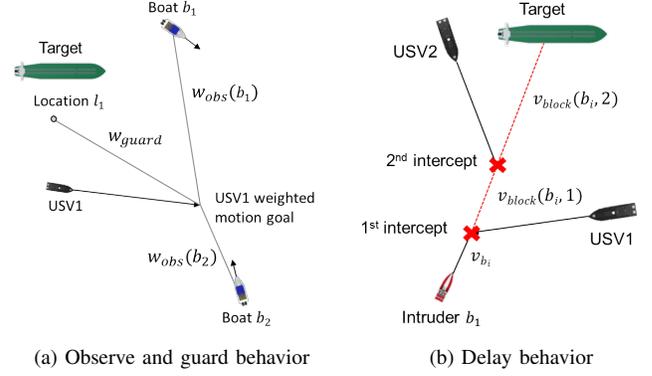


Fig. 2: Example behaviors: a) USV1 approaches a weighted motion goal corresponding to multiple observe and guard tasks, b) two USVs compute a joint intercept path for intruder  $b_1$  by estimating its reduction in velocity.

USV  $u_i$ 's motion goal is defined as,

$$M_{u_i}(O, \mathcal{A}) = \begin{cases} M_{h_j}(O, \mathcal{A}), & \text{if } \exists h_j \in H_{u_i} \cap H_d, \\ M_w(O, H_{u_i}), & \text{otherwise,} \end{cases} \quad (3)$$

which returns the result of Eqn. 2 if  $H_{u_i}$  contains a delay task. Otherwise, it returns a weighted motion goal based on USV  $u_i$ 's currently assigned guard and observe tasks,

$$M_w(O, H_{u_i}) = \frac{\sum_{h_j \in H_{u_i}} w_{h_j}(O) M_{h_j}(O, \mathcal{A})}{\sum_{h_j \in H_{u_i}} w_{h_j}(O)}, \quad (4)$$

where  $w_{h_j}(O)$  is equal to  $w_{guard}$  if  $h_j$  is a guard task, and equal to  $w_{obs}(b_j, O)$  if  $h_j$  is an observe task for boat  $b_j$ ,

$$w_{obs}(b_j, O) = w_{intr} P(b_j \in X|O) \left(1 + \frac{w_{dist}}{|l_{target} - l_{b_j}|}\right). \quad (5)$$

The weights  $w_{guard}$ ,  $w_{intr}$  and  $w_{dist}$  are tuned for each mission using the method described in Sec. IV-D. The resulting policy for USV  $u_i$  is defined as,

$$\pi_{u_i}(O, \mathcal{A}) = v_{u_i} \text{vec}(l_{u_i}, M_{u_i}(O, \mathcal{A})), \quad (6)$$

where  $\text{vec}(l_{u_i}, l_g) = (l_g - l_{u_i})/|l_g - l_{u_i}|$  is the unit vector in the direction of  $l_g$  from USV  $u_i$ 's current location and  $v_{u_i}$  is the surge speed of  $u_i$ .

##### B. Task Allocation

At the start of each scenario, an initial allocation  $\mathcal{A}_0$  assigns a set of guard locations to each USV, distributed uniformly at radius  $r_{guard}$  around the target. As new boats enter the scene, observation tasks for each boat are assigned to the nearest USV. At regular time intervals, a reallocation step occurs, in which each USV  $u_i$  computes a revised allocation  $\mathcal{A}'$ , defined as

$$\mathcal{A}' = \arg \max_{\mathcal{A}_j \in C} \text{eval}(O, \mathcal{A}_j) \quad (7)$$

where  $C$  is a set of candidate task allocations determined by Alg. 1, and  $\text{eval}(O, \mathcal{A}_j)$  computes an estimated *time until arrival* for candidate  $\mathcal{A}_j$  using model-predictive simulation, described in Sec. IV-C. Each candidate  $\mathcal{A}_j \in C$  differs

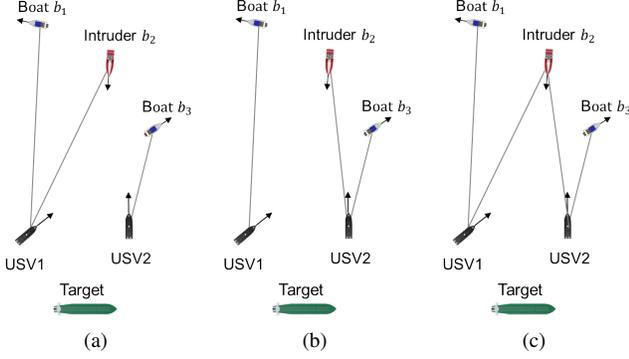


Fig. 3: Candidate task allocations a) the current task allocation  $\mathcal{A}$  without modification, b) modification of  $\mathcal{A}$  with a single delay task exchanged from  $USV1$  to  $USV2$ , c) modification of  $\mathcal{A}$  with a single delay task shared to  $USV2$

from  $\mathcal{A}$  by reassigning a single task from  $H_{u_i}$  to another USV, sharing a task from  $H_{u_i}$  with another USV, or both. These represent incremental changes, useful for gradually improving the joint task allocation without evaluating all possible allocations.

**Algorithm 1** GENERATECANDIDATES( $O, \mathcal{A}, u_i$ ): Generate a set of candidate task allocations.

- 1:  $C \leftarrow \{\mathcal{A}\}$
- 2: **if**  $\exists h_j \in H_{u_j} \cap H_d$  **then**
- 3:      $C \leftarrow C \cup \text{SHARETASKS}(O, \mathcal{A}, u_i)$
- 4: **for each**  $\mathcal{A}_j \in C$  **do**
- 5:      $C \leftarrow C \cup \text{EXCHANGETASKS}(O, \mathcal{A}_j, u_i)$
- 6: **return**  $C$

The function EXCHANGETASKS( $O, \mathcal{A}, u_i$ ) returns a set of task allocations  $C_E = \{\mathcal{A}_0, \mathcal{A}_1, \dots, \mathcal{A}_m\}$  where each  $\mathcal{A}_j \in C_E$  is the same as the input allocation  $\mathcal{A}$ , except that task  $h_f$  has been removed from  $H_{u_i}$  and given to another USV  $u_j \in U$  instead. Task  $h_f \in H_{u_i}$  is the task whose individual motion goal  $M_{h_f}(O, \mathcal{A})$  is furthest from  $u_i$ 's current motion goal. Intuitively, this means we are dropping a task that USV  $u_i$  is least able to fulfill, and creating candidate allocations for the other USVs who may be better suited for that task.

The definition of SHARETASKS( $O, \mathcal{A}, u_i$ ) is very similar to EXCHANGETASKS, only the original task  $h_f$  is never dropped and delay tasks are the only tasks considered. If a delay task already has  $d_{max}$  USVs assigned to it, the task will not be shared. If at any point  $P(b_i \in X|O)$  exceeds the alert threshold  $p_{alert}$ , the observe task for boat  $b_i$  will be automatically converted into a delay task for  $b_i$ .

### C. Predictive Simulation

Each USV generates a set  $C$  of candidate task allocations using the method in Alg. 1. To evaluate each  $\mathcal{A}_j \in C$ , we generate  $k$  scenarios  $W = \{w_i\}_{i=1}^k$  consistent with the current observation history  $O$ . Each scenario  $w_i$  is generated by selecting a set of possible intruder boats  $X_i \subseteq B$  based on the observations made by the USV team. The

number of possible sets  $X_i$  is determined by the minimum  $x_{min}$  and maximum  $x_{max}$  number of intruders, where  $k = \sum_{i=x_{min}}^{x_{max}} \binom{n}{i}$  given  $n$  total boats in the scene. For each set of possible intruders  $X_i$  we estimate the joint probability,

$$P_{X_i} = \left( \prod_{b_j \in X_i} P(b_j \in X|O) \right) \left( \prod_{b_j \in B \setminus X_i} 1 - P(b_j \in X|O) \right)$$

where  $P_{X_i}$  is an approximation of  $P(X = X_i|O)$  computed by assuming that the appearance of each intruder is statistically independent from the appearance of other intruders.

We use Monte Carlo simulation to estimate the expected utility  $E[\mathcal{U}(\Pi_u, O, \mathcal{A}_i)]$  of  $\mathcal{A}_i$ , where  $\Pi_u = \{\pi_{u_i} : u_i \in U\}$  is the joint policy of the entire USV team. As before, utility is defined as  $t_\delta = t_{arrival} - t_{alert}$ . During the simulation, the policies  $\pi_{u_i}$  or  $\pi_{b_j}$  of USVs, passing boats, and intruder boats, respectively, are integrated to produce new states.

During the predictive simulation, the task re-allocation step (see Sec. IV-B) is performed at 1/6th the normal frequency, and a fast heuristic evaluation method is used to select the best USV. This is to prevent the predictive simulation from recursively calling itself. Each trial is also given a maximum duration, after which the arrival time of the intruders is estimated based on the current state.

When determining which USV should receive a guard or observe tasks, the heuristic selects the USV  $u_j$  that has the least distance between its current motion goal and the new task. If USV  $u_j$  is already assigned a delay task, the distance between its motion goal and the new task is multiplied by  $w_{occupied}$ . When assigning delay tasks, the USV is selected that minimizes the estimated arrival time of the intruder, given the trajectory provided by  $l_{opt}(u_i, b_j, O, \mathcal{A})$ .

In the worst case, if  $x_{min} = 0$  and  $x_{max} = n$ , the number of possible scenarios  $k$  evaluated by the predictive simulator is bounded by  $O(2^n)$ . For very large problems the running time may become prohibitive and require the pruning of scenarios based on their probability. However, in the experimental results section later in this paper, we show that the algorithm can be performed in real time without pruning for reasonably sized scenarios.

### D. Genetic Optimization

We used a genetic algorithm (GA) to optimize the underlying parameters  $w_{guard}$ ,  $w_{intr}$ ,  $w_{dist}$ ,  $d_{max}$ ,  $r_{guard}$  and  $w_{occupied}$  of the observe, guard, and delay behaviors to further improve the expected utility of the USV policy. The optimization of these behaviors allows the USVs to make balanced decisions between guarding a certain location, observing incoming boats, and intercepting and thus delaying the movement of identified intruders.

We used a population size of 100, with initial parameters for each chromosome assigned at random. We utilized roulette wheel selection to determine the breeding population, and applied genetic operators with a crossover rate of 0.35 and mutation rate of 0.08. Each chromosome's fitness was measured using the average *time until arrival* of 1000 random simulation runs.

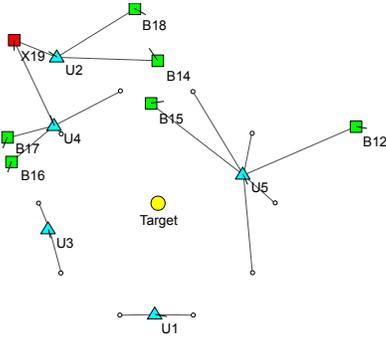


Fig. 4: Scenario 1, five USVs defend a target in an open ocean with several passing boats. Boat  $X19$ , identified as an intruder, is pursued by USVs  $U2$  and  $U4$ . The task assignments for each USV are shown as connecting lines.

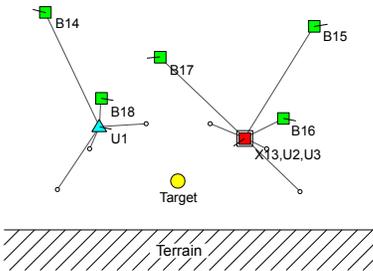


Fig. 5: Scenario 2, three USVs defend a target that is protected by terrain to the south. USVs  $U2$  and  $U3$  are actively blocking intruder  $X13$ , reducing the speed at which it approaches the target.

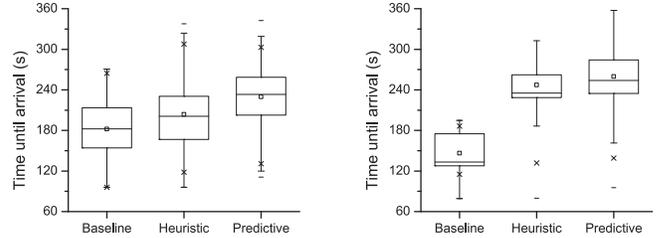
## V. RESULTS

### A. Experimental Setup

We have evaluated our planning approach using two distinct scenarios.<sup>1</sup> In scenario 1 (Fig. 4), the target is positioned within a circular region without any static obstacles. In scenario 2 (Fig. 5), the target is positioned above static terrain restricting the direction of incoming boats. In scenario 1 there are a total of 5 USVs and between 2 and 3 intruders, while in scenario 2 there are 3 USVs and between 1 and 2 intruders. Not counting USVs, both scenarios maintain a total of 8 boats in the scene at any given time.

At the beginning of each trial, boats are initialized at random locations around the target. During the run, new boats appear at the boundary of the operating space, which is defined as a ring in the scenario 1, (with an inner and outer radius of 1200 and 1500 m), or as two rectangles on the left and right sides of the target in the scenario 2 (with a distance of 1200 m from the target and a width of 300 m).

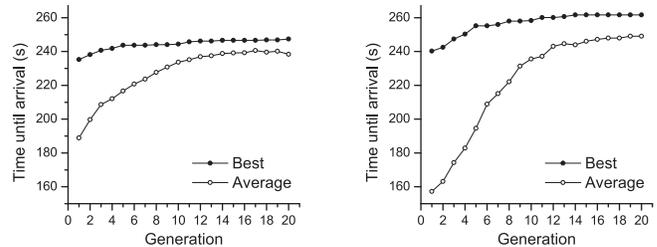
Each boat's initial trajectory is a path tangent to a randomly sized circle (or semi-circle in scenario 2) surrounding the target with minimum and maximum radius of 450 and 900 m. The intruders turn towards the target when they pass within 900 m, but if an intruder passes within 60 m



(a) Scenario 1

(b) Scenario 2

Fig. 6: Average *time until arrival*,  $t_\delta$ , across 1000 randomly seeded trials for USV teams using *predictive*, *heuristic* or *baseline* strategies. For both scenarios, the predictive strategy was better at delaying the intruders.



(a) Scenario 1

(b) Scenario 2

Fig. 7: Average *time until arrival*,  $t_\delta$ , for each generation of the genetic algorithm when optimizing the *predictive* strategy. Figures show the best performing chromosome and the average across all 100 chromosomes in the population.

of a USV, it will start approaching the target immediately. The maximum speed is 10  $m/s$  for USVs and 9  $m/s$  for all other boats, while the blocking function is defined as  $v_{block}(b_i, n) = v_{b_i}/(n + 1)$ . We do not consider differential constraints on a boat's movement, since the planning is executed on a large scale.

The observation classification function is simulated based on the USV team's distance from each boat; if no USV has ever come within 600 m of boat  $b_i$ , then  $P(b_i \in X|O)$  returns the prior probability 0.05. As USVs move within 600 m to 60 m of  $b_i$ , the probability converges to 1 or 0, depending on whether or not  $b_i$  is actually an intruder. If any USV comes within 60 m of  $b_i$ , then  $b_i$  is classified as an intruder or non-intruder immediately. Gaussian noise is added to the probability function so that the change is non-monotonic.

### B. Result of Policy Evaluation

The parameters for the observe, guard, and delay behaviors were optimized for both scenarios separately. The genetic algorithm discussed in Sec. IV-D was run for 20 generations, taking approximately six hours using Condor HTC with 100 cores @ 2.4 GHz. The results are shown in Fig. 7.

Fig. 6 a) and b) show the average *time until arrival* across 1000 different trials for three different strategies. The *predictive* strategy is the complete strategy described

<sup>1</sup>An accompanying video is available at <http://ieeexplore.ieee.org>

in Sec. IV. The *heuristic* strategy does not perform any predictive simulation and makes choices based only on the heuristic evaluation method described in Sec. IV-C. The *baseline* strategy does not assign observe tasks at all; each USV waits at its default guard location until an intruder is identified, at which point a delay task is assigned to the closest  $d_{max}$  USVs. Each strategy was optimized independently using the genetic algorithm.

As expected, the predictive strategy performed best, followed by the heuristic strategy, and the baseline strategy performed worst. Compared to the baseline strategy, the predictive strategy increased the *time until arrival* by 27% in scenario 1 and 78% in scenario 2. Compared to the heuristic strategy, the predictive strategy increased the time until arrival by 13% in scenario 1 and 5% in scenario 2. The difference in performance between the heuristic and predictive strategies is less apparent in scenario 2, likely due to the smaller number of choices during the task allocation step, increasing the chance of the heuristic making the right decision without any simulation.

In both scenarios, the predictive strategy was suitable for real-time computation. In scenario 1, the average running time for a single USV to complete one task allocation step was 30 ms, with a worst-case time of 1076 ms. For scenario 2, the average running time was 7 ms, with a worst-case time of 312 ms. The running time in the average case is significantly less than the worst-case, because there is only one possible world to evaluate once all the intruders have been identified. The allotted time between task allocation steps was fixed at 5000 ms, leaving room for more complex simulations to be used in the future.

## VI. CONCLUSIONS AND FUTURE WORK

We have developed a market-based planning approach for protecting an asset by a team of USVs operating in a continuous state-action space. The developed planner is able to deal with uncertainty about which boats are actual intruders, and can be optimized for a specific mission by using a genetic algorithm.

We have demonstrated the planner's performance in two simulation scenarios. Performance was defined in terms of the expected minimum arrival time to the target by an intruder boat after it is alerted to the presence of intruders. In both scenarios, the developed model-predictive planner had a significant performance advantage compared to a baseline strategy. Due to the simulation-based, model-predictive evaluation of task allocation plans, the planner also performed better than its hand-coded heuristic counterpart.

In future work, we will study how blending multiple tasks with variable priorities will affect guarding performance, incorporate vehicle dynamics into the model to acquire a realistic planning solution for smaller scale environments, and learn an action-selection policy to produce state dependent, optimized parameters rather than utilizing a static set. We also will enhance the model to include sensor noise, communication failures, and static obstacles, and will apply the algorithm in ground- and aerial-vehicle domains.

## ACKNOWLEDGMENTS

This research has been supported by ONR grants N00014-10-1-0585 and N000141210430, ARO grants W911NF-12-1-0471 and W911NF1110344, and a UMIACS New Research Frontiers Award. The opinions expressed in this paper are those of the authors and do not necessarily reflect opinions of the sponsors.

## REFERENCES

- [1] S. Corfield and J. Young, "Unmanned surface vehicles—game changing technology for naval operations," *Advances in unmanned marine vehicles*, pp. 311–328, 2006.
- [2] P. Svec, M. Schwartz, A. Thakur, and S. K. Gupta, "Trajectory planning with look-ahead for unmanned sea surface vehicles to handle environmental disturbances," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'11)*, September 2011.
- [3] P. Svec, A. Thakur, and S. K. Gupta, "USV trajectory planning for time varying motion goal in an environment with obstacles," in *ASME 2012 International Design Engineering Technical Conferences (IDETC) & Computers and Information in Engineering Conference (CIE)*, 2012.
- [4] A. Thakur, P. Svec, and S. K. Gupta, "GPU based generation of state transition models using simulations for unmanned surface vehicle trajectory planning," *Robotics and Autonomous Systems*, 2012.
- [5] P. Svec and S. K. Gupta, "Automated synthesis of action selection policies for unmanned vehicles operating in adverse environments," *Autonomous Robots*, vol. 32, no. 2, pp. 149–164, 2012.
- [6] B. Gerkey and M. Mataric, "A formal analysis and taxonomy of task allocation in multi-robot systems," *The International Journal of Robotics Research*, vol. 23, no. 9, pp. 939–954, 2004.
- [7] Y. Shoham and K. Leyton-Brown, *Multiagent systems: Algorithmic, game-theoretic, and logical foundations*. Cambridge Univ Pr, 2009.
- [8] R. Simmons, D. Apfelbaum, D. Fox, R. Goldman, K. Haigh, D. Musliner, M. Pelican, and S. Thrun, "Coordinated deployment of multiple, heterogeneous robots," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2000)*, vol. 3. IEEE, 2000, pp. 2254–2260.
- [9] T. Huntsberger, P. Pirjanian, A. Trebi-Ollennu, H. Das Nayar, H. Ag-hazarian, A. Ganino, M. Garrett, S. Joshi, and P. Schenker, "Campout: A control architecture for tightly coupled coordination of multirobot systems for planetary surface exploration," *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, vol. 33, no. 5, pp. 550–559, 2003.
- [10] L. Parker, "Multiple mobile robot systems," *Springer Handbook of Robotics*, pp. 921–941, 2008.
- [11] B. Gerkey and M. Mataric, "Sold!: Auction methods for multirobot coordination," *IEEE Transactions on Robotics and Automation*, vol. 18, no. 5, pp. 758–768, 2002.
- [12] M. Dias, "Traderbots: A new paradigm for robust and efficient multirobot coordination in dynamic environments," Ph.D. dissertation, Carnegie Mellon University, 2004.
- [13] N. Kalra, D. Ferguson, and A. Stentz, "Hoplites: A market-based framework for planned tight coordination in multirobot teams," in *IEEE International Conference on Robotics and Automation (ICRA'05)*. IEEE, 2005, pp. 1170–1177.
- [14] D. Portugal and R. Rocha, "A survey on multi-robot patrolling algorithms," *Technological Innovation for Sustainability*, pp. 139–146, 2011.
- [15] —, "On the performance and scalability of multi-robot patrolling algorithms," in *IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*. IEEE, 2011, pp. 50–55.
- [16] E. Simetti, A. Turetta, G. Casalino, E. Storti, and M. Cresta, "Protecting assets within a civilian harbour through the use of a team of usvs: Interception of possible menaces." OCEANS, 2010.
- [17] Y. Zhang and Y. Meng, "A decentralized multi-robot system for intruder detection in security defense," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'10)*. IEEE, 2010, pp. 5563–5568.
- [18] T. Sandholm, "Contract types for satisficing task allocation," in *Proceedings of the AAAI spring symposium: Satisficing models*, 1998, pp. 23–25.
- [19] M. Dias, R. Zlot, N. Kalra, and A. Stentz, "Market-based multirobot coordination: A survey and analysis," *Proceedings of the IEEE*, vol. 94, no. 7, pp. 1257–1270, 2006.