1 INTRODUCTION

In recent years, small unmanned vehicles have become inexpensive and deadly weapons. It is easy to imagine a scenario where a small unmanned explosives-packed submarine is launched by terrorists from a freighter, at a safe distance from the entrance to a harbor, with the mission of destroying a large cruise ship carrying many thousands of passengers. The effect could be as devastating as the 9/11, 2001 attack on the World Trade Center in New York. Idealizing a real world situation, we consider a harbor that can be reached via a rectangular channel of width $W$ and assume that one or more intruders tries to reach the target ships anchored at the end of the harbor. The task of the defender craft is to prevent the intruders from reaching their target, as illustrated in Fig. 1. The defenders can achieve their goal either by destroying the intruders or causing them to flee.

We assume that the defending vehicles are manned or unmanned submarines, manned or unmanned hovercraft, or drones. The purpose of this paper is to construct a feedback receding horizon control (RHC) law (see (Mayne, Rawlings, Rao, & Scokaert 2000)) for the defenders, based on max-min optimal control problems which, we believe, capture the essence of the intruders’ goal of at least one of them getting within striking distance of their target, as well as the intruders’ perception of how they can be destroyed by the defenders.

The idea behind RHC is quite old, going back to the 1950’s, and is based on the following observation. Suppose that we have a dynamical system
that is modeled by a differential equation of the form

$$\dot{x}(t) = f(x(t), u(t)), \quad x(0) = x_0, t \geq 0$$ \hspace{1cm} (1)

where $x(t)$ is the state of the system and $u(t)$ is the control. Now suppose that one wants to optimize its behavior by solving an optimal control problem of the form

$$\min_{u(t) \in U, x(t) \in X} \int_0^\infty f^0(x(t), u(t))dt,$$ \hspace{1cm} (2)

subject to the differential equation constraint (1).

Let $\hat{u}(t)$ be solution of this optimal control problem and $\dot{x}(t) = x(t, \hat{u}(t)), t \geq 0$ the resulting trajectory. Now suppose that the model (1) is not perfect, so the actual trajectory resulting from the control $\hat{u}(t)$, $x^*(t, \hat{u}(t))$, is quite different from $\dot{x}(t)$. To remedy this situation, it was proposed that every $\Delta$ time units, the actual state be measured and the problem (1) re-solved, to obtain a corrected control, effectively creating a feedback mechanism. Experiments have shown that this is an excellent idea. For an excellent survey of RHC, see (Mayne, Rawlings, Rao, & Scokaert 2000). For previous attempts of using RHC in pursuit-evasion situations, see (Sprinkle, Eklund, Kim, & Sastry 2004, Walrand, Polak, & Chung 2011, S. Lee & Walrand 2013).

Our situation is more complicated than the one above. We have multiple defender dynamic systems and multiple intruder dynamic systems and the optimal control problem (2) must be replaced by a problem that reflect this fact as well as the fact that the defenders do not know the intruders strategies. Hence we propose a worst case approach and, since continuous time optimal control problems require a great deal of time for their solution, we will assume that the controls are constant during the sample times, which results in the replacement of dynamics described by differential equations by derived dynamics described by difference equations.

2 RECEDING HORIZON CONTROL LAW FORMULATION

2.1 ASSUMPTIONS

First, it does not seem to be possible to solve the type of problem we propose over an infinite horizon, as in (2). Hence we introduce a finite horizon $N\Delta > 0$, with $N$ a positive integer, the sample time $\Delta > 0$ has to be chosen taking into account the speed with which the craft are moving and the time it takes to solve the RHC determining min-max optimal control problem. For example, if, as in our case, with a 10 sample time horizon, it takes 0.2 seconds to solve the problem, the sample time could be 0.5-1.0 seconds, which is equivalent to adjusting the control every 0.01-0.02 miles for a torpedo travelling at 80 MPH. Since the longer the horizon the longer the computing time, the length of the horizon is largely determined by the computing power available for the defenders.

Second, We assume that the intruders cannot risk engaging the defenders in battle and we consider three possible scenarios. In the first, which is deterministic, the intruders assume that they are safe as long as they avoid coming within striking distance of the defenders (S. Lee & Walrand 2013). In the next two scenarios, which are probabilistic as well as more realistic, the intruders assume that the probability of their destruction is a function of their distances from the defenders.

In the first probabilistic formulation, the intruders attempt to survive over the entire horizon. In the second one, first the intruders attempt to destroy their target within the horizon time, and only if successful do they attempt to survive to the end of the horizon.

Third, we assume that the defenders are able to determine the dynamics of the intruders and that the floor of the harbor is seeded with sensors that enable the defending team to determine continuously the position, velocity, and direction of travel of the intruder. Just to be safe, we also assume that the intruders have access to similar information about the defenders.

Fourth, we assume that the actions of the defenders are coordinated and that the actions of the intruders are also coordinated. Coordination of the defenders (intruders) can be achieved either by using an offshore or mother ship computer to solve the RHC optimal control problem and then transmitting the required controls to each individual craft, or by each craft solving the same RHC
optimal control problem and using the appropriate resulting control.

2.2 DYNAMICS

Assuming that both the defenders and intruders are unmanned underwater vehicles (UUVs) and that they are confined to a rectangular channel of width $W$ at the end of which is the harbor with the high value target, their dynamics have the form

$$
\begin{align*}
\dot{x}^1(t) &= v(t) \cos \theta(t) \\
\dot{x}^2(t) &= v(t) \sin \theta(t) \\
\dot{\theta}(t) &= \sigma(t) \\
\dot{v}(t) &= \alpha(t),
\end{align*}
$$

(3)

where $x^1$ is the positional coordinate of the UUV along the channel, $x^2$ is the positional coordinate of the UUV perpendicular to the channel, and $\theta$ is the heading, i.e., the angle between the direction of motion of the UUV and the $x^1$ axis in our positional coordinate system. We assume that the channel is sufficiently shallow that a depth coordinate is not needed. We assume that there is a steering input $\sigma(t)$ and a propulsion input $\alpha(t)$, which are subject to constraints of the form:

$$
\begin{align*}
0 \leq v(t) \leq \bar{v} \\
|\sigma(t)| \leq \bar{\sigma} \\
|\alpha(t)| \leq \bar{\alpha} \\
\sigma(t)v(t) \leq k_f.
\end{align*}
$$

The constraint $\sigma(t)v(t) \leq k_f$ captures the relationship between centripetal force and velocity.

We assume that the receding horizon control law uses a sample time $\Delta$, so that for any integer $k \geq 0$ and $t \in [k\Delta, (k+1)\Delta)$ the controls are constant, i.e., for $t \in [k\Delta, (k+1)\Delta)$, $v(t) = v(k\Delta)$ and $\sigma(t) = \sigma(k\Delta)$. We can integrate the differential equation (2) for $t \in [k\Delta, (k+1)\Delta)$, to obtain the difference equations:

$$
\begin{align*}
x_{k+1}^1 &= x_k^1 + \Delta v_k \cos \theta_k + \Delta^2 \alpha_k \cos \theta_k \\
x_{k+1}^2 &= x_k^2 + \Delta v_k \sin \theta_k \Delta^2 \alpha_k \sin \theta_k \\
\theta_{k+1} &= \theta_k + \Delta \sigma_k \\
v_{k+1} &= v_k + \Delta \alpha_k,
\end{align*}
$$

(5)

where $x_k^1 = x^1(k\Delta)$, $x_k^2 = x^2(k\Delta)$, $\theta_k = \theta(k\Delta)$, $v_k = v(k\Delta)$, $\sigma_k = \sigma(k\Delta)$, and $\alpha_k = \alpha(k\Delta)$.

2.3 DEFENDER RECEDING HORIZON CONTROL LAW

Let $\bar{z}_{d,1}(k\Delta) = (\bar{z}_{d,1}(k\Delta), \ldots, \bar{z}_{d,N_d}(k\Delta))$ and $\bar{z}_{i,1}(k\Delta) = (\bar{z}_{i,1}(k\Delta), \ldots, \bar{z}_{i,N_i}(k\Delta))$, where $\bar{z}_{d,1}(k\Delta) = (x_{d,1}^1(k\Delta), x_{d,1}^2(k\Delta), \theta_{d,1}(k\Delta), v_{d,1}(k\Delta))$, and so forth.

Defender Receding Horizon Control Algorithm

Data: Sampling Time = $\Delta$, computing time $\delta < \Delta$, horizon = $N\Delta$, initial defender and intruder states $\bar{z}_d(0), \bar{z}_i(0)$,

Step 1: Set $k = 0$.

Step 2: Set $z(0) = \bar{z}_d(k\Delta)$ and $\bar{z}_i(0) = \bar{z}_i(k\Delta)$.

Step 3: Solve one of the defender min-max problems, below, for an optimal coordinated defender control $u^*_d$.

Step 4: Apply the control $u^*_d(0)$ to the defender for $\Delta$ units of time.

Step 5: At time $\delta + k\Delta$, measure the states $\bar{z}_d(\delta + k\Delta)$ and $\bar{z}_i(\delta + k\Delta)$.

Step 6: Estimate the states $\bar{z}_d((1 + k)\Delta)$ and $\bar{z}_i((1 + k)\Delta)$ using the differential equation (3).

Step 7: Replace $k$ by $k + 1$ and go to Step 2.

Note that the min-max problem in Step 4, above, can be changed at each sampling time, and so can the sample time $\Delta$. It makes sense to use a large $\Delta$ when the adversaries are far apart and decrease it as they get nearer to each other.

3 MIN-MAX PROBLEM FORMULATIONS

The next element that we must introduce is the min-max optimal control problem, reflecting a worst case scenario, that must be solved at each sample time within the RHC law. We will consider three possible scenarios: (a) where the intruders are risk averse, (b) where the intruders are willing to take risks, and (c) where the intruders are willing to sacrifice themselves to achieve their goal.

To distinguish between the intruder and defender, we will add a subscript $i, j$ to indicate the $j$-th intruder states, controls, and constraints, a subscript $d, k$ to indicate $k$-the defender states, controls, and and constraints.

Suppose that there are $N_d$ defenders and $N_i$ intruders and that the horizon length is $N\Delta$. For $k = 0, \ldots, N - 1$, in
assuming that the intruders are risk averse and hence will not venture withing torpedo striking distance \( \tau > 0 \) of the defenders, we propose the following max-min optimal control problem for the receding horizon control law:

\[
\begin{align*}
\max_{u_d \in \mathcal{U}_d} & \quad \min_{u_i \in \mathcal{U}_i(u_d), k \in \mathcal{N}} \min_{j \in \mathcal{N}_i} \{ x^1_{i,j}(k\Delta, u_i) \}, \\
\text{s.t.} & \quad 0 \leq v_{d,j}(k\Delta) \leq \bar{v}_{d,j}, \\
& \quad |\sigma_{d,j}(k\Delta)| \leq \bar{\sigma}_{d,j}, \\
& \quad |\alpha_{d,j}(k\Delta)| \leq \bar{\alpha}_{d,j}, \\
& \quad \sigma_{d,j}(k\Delta)v_{d,j}(k\Delta) \leq \kappa_{d,j}, \\
& \quad 0 \leq x^1_{d,j}(k\Delta, u_{d,j}), \\
& \quad 0 \leq x^2_{d,j}(k\Delta, u_{d,j}) \leq W, 
\end{align*}
\]

where \( k = \{0, \ldots, N - 1\} \), \( \bar{v}_{d,j} > 0 \), \( j = 1, \ldots, N_d \), are the speed limits for the defenders, \( \bar{\sigma}_{d,j}, j = 1, \ldots, N_d \), are the limits on the steering inputs for the defenders, and \( W \) is the width of the channel.

For the intruder, the control constraint depends on the choice of a defender input \( u_d \) via the resulting defender trajectory \( x_d = (x_d(0, u_d), x_d(\Delta, u_d), \ldots, x_d(N\Delta, u_d)) \), determined by (5). Hence, the intruder's control problem is

\[
\begin{align*}
U_{i}(u_d) = \{ & u_i = (u_i(0), \ldots, u_i((N - 1)\Delta)) \\
\text{s.t.} & \quad 0 \leq v_{i,j}(k\Delta) \leq \bar{v}_{i,j}, \\
& \quad |\sigma_{i,j}(k\Delta)| \leq \bar{\sigma}_{i,j}, \\
& \quad \|x_{i,j}(k\Delta, u_i) - x_{d,l}(k\Delta, u_{d,l})\|^2 \geq \tau^2, \\
& \quad 0 \leq x^1_{i,j}(k\Delta, u_{i,j}), \\
& \quad 0 \leq x^2_{i,j}(k\Delta, u_{i,j}) \leq W, k = 0, \ldots, N - 1, \\
& \quad \sigma_{i,j}(k\Delta)v_{i,j}(k\Delta) \leq k_{i,j}, \}
\end{align*}
\]

where \( j = 1, \ldots, N_i, l = 1, \ldots, N_d, \) and \( \tau \) is a torpedo distance.

Note that the defenders’ actions do not affect the cost function. Defense is achieved by interference as expressed by the constraints imposed on the intruder.

There are three issues that must be dealt with in solving problem (6). The first two are obvious, the last one is subtle.

First, (6) is a type of generalized max-min problem (Polak & Royset 2005), because the constraint set of the intruders depend on the strategy \( u_d \) of the defenders and hence cannot be solved by standard max-min algorithms, such as outer approximations. In fact, it is a type of bilevel problem that can be converted to a “standard” max-min problem by adding the defender dependent constraints to the cost function using exact penalty functions, as was done in (Polak & Royset 2005). The exact penalty functions need not be used when evaluating the min part of the max-min problem for a given set of defender controls, but they must be used in the maximization process. The introduction of exact penalty functions transforms problem (6) into

\[
\begin{align*}
\max_{u_d \in \mathcal{U}_d} & \quad \min_{u_i \in \mathcal{U}_i, k \in \mathcal{N}} \min_{j \in \mathcal{N}_i} \{ x^1_{i,j}(k\Delta, u_i) + \pi \max\{ \\
0, -\|x_{i,j}(k\Delta, u_i) - x_{d,l}(k\Delta, u_{d,l})\|^2 + \tau^2 \} \}, \\
\text{s.t.} & \quad 0 \leq v_{d,j}(k\Delta) \leq \bar{v}_{d,j}, \\
& \quad |\sigma_{i,j}(k\Delta)| \leq \bar{\sigma}_{i,j}, \\
& \quad 0 \leq x^1_{i,j}(k\Delta, u_{i,j}), \\
& \quad 0 \leq x^2_{i,j}(k\Delta, u_{i,j}) \leq W, k = 0, \ldots, N - 1, \\
& \quad \sigma_{i,j}(k\Delta)v_{i,j}(k\Delta) \leq k_{i,j}. 
\end{align*}
\]

where \( \pi > 0 \) is the value of the exact penalty function and

\[
\begin{align*}
U'_i = \{ & u_i = (u_i(0), \ldots, u_i((N - 1)\Delta)) \\
\text{s.t.} & \quad 0 \leq v_{i,j}(k\Delta) \leq \bar{v}_{i,j}, \\
& \quad |\sigma_{i,j}(k\Delta)| \leq \bar{\sigma}_{i,j}, \\
& \quad 0 \leq x^1_{i,j}(k\Delta, u_{i,j}), \\
& \quad 0 \leq x^2_{i,j}(k\Delta, u_{i,j}) \leq W, k = 0, \ldots, N - 1, \\
& \quad \sigma_{i,j}(k\Delta)v_{i,j}(k\Delta) \leq k_{i,j}. \}
\end{align*}
\]

Second, the cost function

\[
\min_{j \in \mathcal{N}_i} \{ x^1_{i,j}(k\Delta, u_i) \} 
\]

is not differentiable. This can be dealt with by smoothing (see GAMS), as was done in (Walrand, Polak, & Chung 2011), and found to cause
serious ill-conditioning, or, as we do now, by making use of the fact that the minimum over a set is equal to the minimum over its convex hull. This requires the addition of decision variables \( \mu_{j,k} \geq 0, j = 1, \ldots, N, k = 0, \ldots, N - 1 \) such that
\[
\sum_{j \in N, k \in \mathbb{N}} \mu_{j,k} = 1, \quad (12)
\]
i.e., we add \( N \times N \) variables, with positivity constraints and one equality constraint, which can be eliminated explicitly. The cost function now becomes
\[
\min_{j \in N} \{ \sum_{i \in N, k-1 \in \mathbb{N}} \mu_{j,k-1} x_{i,j}^{1}(k\Delta, u_i) \}. \quad (13)
\]

The third issue stems from the fact that when the separation between defenders and intruders is sufficiently large, the solution of the min part does not require that the constraints be active. Hence, at such situations, the value of the min function is independent of the value of the defender controls, and hence is a stationary point. At such points, solving the min-max problem (9) does not produce a meaningful result from the defenders’ point of view. Hence, we can replace the problem (9) with the problem
\[
\max_{u_d \in U_d} \min_{u_i \in U_d, k \in \mathbb{N}} \min_{j \in N} \{ x_{i,j}^{1}(k\Delta, u_i) + \}
\pi \{ -\| x_{i,j}(k\Delta, u_i) - x_{d,l}(k\Delta, u_d) \|^2 + \tau^2 \} \}, \quad (14)
\]
which results in the defenders always pursuing the intruders.

### 3.2 RISK TAKING INTRUDER

In this case, the defenders assume that the intruders are willing to take a chance of coming within striking distance of a defender, on the belief that the defender may miss him with a certain probability. For the case of a single defender and single intruder, this results in the following min-max optimal control problem that the defender must solve at each sample time.
\[
\min_{u_d \in U_d} \max_{u_i \in U_i(u_d), k \in \mathbb{N}} \phi_1(x_i(k\Delta, u_i)) \phi_2(u_i, u_d, k). \quad (15)
\]

where the probability of a successful strike by the intruder at time \( k\Delta \) is
\[
\phi_1(x_i(k\Delta, u_i)) = \frac{\exp(g_1(x_i(k\Delta, u_i)))}{1 + \exp(g_1(x_i(k\Delta, u_i)))} \quad (16)
\]
and the probability of the intruder surviving for horizon of \( N\Delta \) sample times is
\[
\phi_2(u_i, u_d, N) = \exp \left\{ -\sum_{k=1}^{N} \lambda \exp(g_2(u_i, u_d, k\Delta)) \right\} \Delta \quad (17)
\]
with
\[
g_1(x_i(k\Delta, u_i)) = -\alpha_1 (\| P(z_i(k\Delta, u_i)) - \tau \|^2 - s_1^2) \quad (18)
g_2(u_i, u_d, N) = -\alpha_2 (\| P(z_i(k\Delta, u_i) - z_d(k\Delta, u_d)) \|^2 - s_2^2) \quad (19)
\]
where \( \lambda, \alpha_1, \alpha_2 \) are parameters.

The expressions for multiple defenders and intruders are considerably more complicated and are omitted because of lack of space.

### 3.3 SUICIDAL INTRUDER

This case differs from the preceding one in that no intruder places any value on surviving after a successful attack, but, should his attack be successful will take evasive action. Hence, for the case of a single defender and single intruder, we get the following variant of (15)
\[
\min_{u_d \in U_d} \max_{u_i \in U_i(u_d), k \in \mathbb{N}} \phi_1(x_i(k\Delta, u_i)) \phi_2(u_i, u_d, k). \quad (20)
\]

where, \( \phi_1(x_i(k\Delta, u_i)) \) and \( \phi_2(u_i, u_d, k) \) are defined as in (16) and (17), respectively.

When the solution time of (20) \( k\Delta < N\Delta \) i.e., the intruder may have succeeded in destroying his target, then the defender assumes that the intruder
switches cost functions at time $k^* \Delta$ and concentrates on escape. In that case, we get the following secondary problem for the defender

$$\min_{u_d \in U_d} \max_{u_i \in U_i, k \in \{k^*, \ldots, N\}} \phi_2(u_i, u_d, k). \quad (21)$$

Again, the expressions for multiple defenders and intruders are considerably more complicated and are omitted because of lack of space.

4 SIMULATION RESULTS

We only present results for the risk averse intruder, based on (9). We used a horizon of 10 sample times and used the method of outer approximations (MOA) (see (Polak 1997)) with the Polak-He unified method (PH) (see (Polak 1997)) as a subroutine.

The approximate solution of (9) required 3 iterations of MOA and a total of of 40 iterations of PH. Programmed in JAVA, the solution of (9) this required 0.18 seconds, while programmed in MATLAB with TOMLAB (K. Holmstrom & Edvall 2010), it took 1.8 seconds.

Although the Polak-He unified method is only a first-order method, it computes a good approximate solution to an inequality constrained optimization problem very rapidly. Given that we always had very good starting points for the MOA and the speed of the PH method, even using a laptop, we were able to compute controls at a rate that is compatible with real time implementation in craft moving up to 80 knots.

Since static figures do no convey the evolution in time of the defender and intruder trajectories, we have deposited videos of our experiments in https://sites.google.com/site/walrandberkeley/research/harbor.

4.1 SINGLE DEFENDER AND SINGLE INTRUDER

Fig. 2 illustrates an experimental setup for the case of a single intruder and single defender. The intruder and defender are located in the rectangular channel with a channel width $W$. The harbor is depicted as a thick line which is located behind the defender. Small red and blue circles indicate the locations of intruder and defender, respectively. The small bar attached to the circles indicates their orientation.

Fig. 3 (a) illustrates the case when the defender successfully defends the harbor. The dotted line are added as a trajectory guidance. Fig. 3 (b) depicts the case when the intruder successfully outmaneuver the defender, and reaches to the harbor.

A set of experiments was performed of a type that can be used to determine the maximum channel width that a single defender can protect, assuming that the parameters of the intruder are known. Initially, the intruder and the defender are facing each other: initial intruder and defender orientations are $\pi$ and 0, respectively. Their controls are bounded by identical limits.

4.2 TWO DEFENDERS AND TWO INTRUDERS

Fig. 4 shows the simulation result with two intruders and two defenders. Initially, they are facing each other. Fig. 4 (a) is the case when the defender team wins as the channel is narrow with $W = 20$. Fig. 4 (b) is the case when the intruder team wins as the channel is wide with $W = 25$. 

Figure 3: Example trajectories: single intruder and defender.
Since one of the intruder team member successfully reached the harbor, the intruder team wins the game.

4.3 **HUMAN-MACHINE INTERACTION**

Fig. 5 (a) is a screen capture of a real time 3D simulation. We assume the harbor is located behind of the defender (left side of the screen). The human controls the intruder. Fig. 5 (b) is a photo of a laboratory experiment involving a human intruder and an autonomous defender. Both in the 3D simulation and in the experiment, the human intruder uses a joystick to activate the intruder. In the experiment, HoTDeC (HOvercraft Testbed for DEcentralized Control) vehicles developed at the University of Illinois at Urbana-Champaign (UIUC) were used as players.

5 **CONCLUSION**

We have presented three alternatives to automating harbor defense vehicles using RHC and have shown that the computing times associated with solving the RHC min-max optimal control problems were sufficiently low for real time implementation in craft moving up to 80 knots.

6 **ACKNOWLEDGEMENT**

This work was supported in part by U.S. Air Force Office of Scientific Research (AFOSR) under grant number MURI FA9550-10-1-0573 and NSF under grant number 1016989. We wish to thank J. O. Royset, J. Foraker, and Vijay Kamble for providing us with the probabilistic model, as well as many useful comments.

**REFERENCES**