Optimal Navigation Policy for an Autonomous Agent operating in Adversarial Environments

Emmanuel Boidot1 Aude Marzuoli2 Eric Feron3 *

Abstract—We consider an autonomous vehicle navigation problem, whereby a traveler aims at traversing an environment in which an adversary tries to set an ambush. Optimal strategies are computed as random path distributions, a realization of which is the path chosen by the traveler. Theoretical optimal policies are derived under assumptions from the Minimal Cut-Maximal Flow literature. Numerical approaches to compute such optimal strategies are proposed. These numerical approaches, which borrow from randomized path planning techniques, can be implemented for high-dimensional configuration spaces. The methodology developed allows for the application of ambush games on complex environments for realistic applications regarding vehicle routing in adversarial settings.

I. INTRODUCTION

We tackle the management of ambushes in a continuous environment from a game theoretic standpoint. In the classical ambush model, an autonomous agent (BLUE) must cross an environment where traps have been placed. If the agent comes within a given distance R from the location of the trap, the agent is penalized or suffers a loss. This problem can be seen as a specific case of robotic path planning in the presence of threats, whose formulation in continuous configuration spaces is new to the best of the authors knowledge. Typical applications include the naval battleships versus mines problem [1], where a battleship travels through a canal where bombs have been set up. Other possible applications are: a setting in which an autonomous agent (such as a delivery drone, or a patrol drone) has radio-frequency communications towards a home base with high level goals, but an adversary sets up jamming honey pots in order to take over control of the agent; the routing of an Army Joint Light Tactical Vehicle (JLTV) traveling through a hostile environment. Given the topology of the environment, and the impact range of the trap, our goal is to address the following questions by means of game theory:

• What are the best policies for the agent and his/her adversary?
• What is the probability that the agent traverses the environment safely?

This work builds upon previous research on game theory by Ruckle [2], [3] who extended Isaac’s [1] classical battleship versus bomber duel by interpreting a two-dimensional environment as a rectangular array of lattice points [4] and later as a set of tiles[3]. Ruckle identified optimal policies for an agent, BLUE, and his opponent, RED, for different games played on this rectangular environment, with varying conditions on the type of ambush set that RED may occupy or the type of paths BLUE may follow. The main limitations of his work are that it is only applicable to 2D rectangular environments, where no obstacles are present.

In order to address these limitations, the concept of minimal cut of a continuous environment becomes important. The minimal cut of a continuous environment Ω containing a source and a sink is described in a variety of research endeavours, including the work of Strang [5], Mitchell [6], Gomory and Hu [7] showed that it is possible to approximate the continuous flow by constructing a discrete network representation of the environment, something that will be explored for our problem later in this paper. So as to compute the minimal cut, additional hypotheses regarding the environment are required, as detailed in Section II. Mitchell [6] and Polishchuk [8] addressed the problem of constructing minimal cuts and maximal flows on polygonal domains efficiently using an algorithm that runs in polynomial time in the dimension of the environment. For the remainder of this paper, the environment is supposed to be a polygonal environment, as defined by Strang [5], with holes, also modeled as polygons. This assumption, while strong, remains very realistic for the purpose of studying routing problems in unstructured 2D environments. For applications in a higher dimensional setting, the method proposed in Section III remains valid, and can be easily extended to aerial vehicles navigation problems.

Making use of the polygonal environment hypothesis, the contributions of this paper are as follows. First, we elaborate a new game theoretic setting for the problem of vehicle ambushing. In this framework, we demonstrate that the optimal value of the game, described in Section II, depends on the length of the minimal cut (minimal with respect to a certain metric dependent on R, the reaching capacity, or reach, of RED). Second, building upon the work of Mitchell [6] and Polishchuk [8], we show that an optimal strategy of the game, as measured in terms of probability of agent survival, can always be found in a polygonal environment Γ. Moreover, such a strategy can be computed in \(O(n \log n + nh)\), where h is the number of holes in the environment Ω, which is a \(n\)-gon. Third, for practical computation of realistic environments, where a rich description of the obstacles and boundaries of the environment, such as a polygonal description, is not readily available, we can approximate the optimal policies of both RED and BLUE by solving a linear program corresponding to a similar game played on a graph.
representing the environment.

Section II describes the ambush game and presents the theoretical results regarding the optimal policies for both players. In Section III, a linear problem is formulated for the resolution of a similar ambush game on a graph and the optimal strategies are derived. We propose to obtain such a graph by sampling methods on the environment. Section IV shows that the results obtained on a sampled graph converge to the theoretical value of the game. Section V discusses the results obtained and concludes the paper possible extension of this work.

II. CONTINUOUS AMBUSH GAME

A. Ambush game and minimal cut in polygonal environments

A polygonal domain \( \Omega \) is a simple polygon with holes, represented as an ordered list of vertices comprising its outer boundary, and a list of holes, and also represented as a list of vertices.

Consider a polygonal environment \( \Omega \) with a source \( s \) and a sink \( t \), as displayed in Figure 1. The traveler (BLUE) is trying to go from \( s \) to \( t \) through \( \Omega \), while the adversary (RED) sets up an ambush at some point \( x_a \) in the free space \( C_{\text{free}}(\Omega) \). The only information available to both players is the description of the environment and the locations of the source and sink. No sensing ability is envisioned in the present definition of the game. The ambusher receives a reward of 1 if BLUE’s path comes within a distance \( R \) of \( x_a \), and receives 0 otherwise. This encounter is a zero sum game and is referred to as a Continuous Ambush Game in a polygonal environment. BLUE’s path belongs to the set \( B \) of all continuous and piecewise continuous functions \( f \) from \([0,1]\) into \( C_{\text{free}}(\Omega) \) such that \( f(0) \) belongs to \( s \) and \( f(1) \) belongs to \( t \).

The outcome of the game can be defined as

\[
\mathcal{V} = \langle f, x_a \rangle = \begin{cases} 
1, & \text{if } \min_{t \in [0,1]} \|f(t) - x_a\| \leq R \\
0, & \text{if } \min_{t \in [0,1]} \|f(t) - x_a\| > R 
\end{cases}
\]

Given this uniform reward function of 1, the outcome of the game can be understood as the probability that BLUE be ambushed by RED. For a fixed ambush radius \( R \), we are interested in understanding whether the game has an optimal value for both players, and how the structure of the environment and the optimal outcome of the game are related. Note that the optimal policies for RED and BLUE might be mixed strategies.

A cut is defined as a subset \( S \) of \( \Omega \) such that \( S \) contains the sources and no sinks. Let \( \partial S \) be the boundary of \( S \). We denote \( \gamma \) the part of \( \partial S \) that separates the source from the sink, and such that each point of \( \gamma \) has a neighborhood contained in \( \Omega \). A minimal cut is then a cut \( \hat{S} \) where \( \gamma \) has minimal length. Mitchell [6] provides several important results, summarized here. First, a minimal cut always exists, though it is not always unique. Second, a minimal cut can be computed quickly.

The bottom B (resp. top T) of the environment \( \Omega \) is defined to be the portion of the boundary of \( \Omega \) between \( t \) and \( s \) (resp. \( s \) and \( t \)), when following \( \partial \Omega \) clockwise, as displayed in Figure 1. The notion of the critical graph of a domain ([9], [6]) is central to finding minimal cuts and maximal flows in geometric domains. Figure 2 (left) shows the critical graph of the example environment displayed in Figure 1. The critical graph has a vertex for every hole \( H_i \) contained in the environment \( \Omega \) (with additional holes representing from Top and Bottom); the length of an edge \((i,j)\) is the distance between the holes \( H_i \) and \( H_j \). Mitchell [6] used the critical graph to formulate and prove the “Continuous MaxFlow-MinCut Theorem”. Using the critical graph of \( \Omega \), a cut can be described as a set \( \{s_i\}_{i=1}^P \) of edges of the critical graph. A minimal cut \( \gamma^\ast \) of the example environment is shown in Figure 2 (center), in red.

Polishchuk [8] introduces the concept of thick path and well separated path. A thick path \((\pi)^R\) of width \( R > 0 \) is defined as the Minkowski sum of a curve \( \pi \) in \( \mathbb{R}^2 \) (the reference path) and a disk of radius \( R \):

\[
(\pi)^R = \pi \oplus C_R = \{x + y | x \in \pi, y \in C_R\}.
\]

A set of \( R \) well separated (thin) paths (or \( R \)-WSP) is a set of paths separated by a distance greater than or equal to \( R \) at any point in the domain. Polishchuk defines a thresholded version of the critical graph of the domain, where the length of each edge \( s_i \) is reduced to the closest smaller multiple factor of \( 2R \), i.e. \([s_i]_{2R} \). This thresholded critical graph is used in the “Discrete Continuous MaxFlow-MinCut Theorem”, which states that the maximum number of thick non-crossing paths in \( \Omega \) equals to the length of the shortest top to bottom (T-B) path in the thresholded critical graph.

B. Optimal policies of ambush games

We now leverage min-cut problems in continuous environments to derive the solution to the Continuous Ambush Game in polygonal environments. We introduce the ambush-cardinality \( l_R(\gamma) \) of a set of segments \( \gamma = \{s_i\}_{i=1}^P \) as the sum of the rounded up ratio of the length of each segment \( s_i \) to \( 2R \):

\[
l_R(\gamma) = \sum_{i=1}^P \left\lceil \frac{|s_i|}{2R} \right\rceil.
\]
A different thresholded critical graph corresponding to the environment $\Omega$ is defined, where the length of each edge $s_i$ is replaced by $l_R(s_i)$. For the remainder of this section, an ambush-minimal cut $\gamma^*_R$ is supposed to be a cut corresponding to a minimal T-B path in the modified critical graph (upper thresholded critical graph). Let $l_R(\gamma^*_R)$ be the capacity of this cut. An example of such a cut is displayed in Figure 2 (center and right) and can be compared with a “traditional” continuous minimal cut displayed in Figure 2 (left). It may be observed that the minimal cut and the ambush-minimal cut or their capacity are not directly related, except in the case where $R$ tends towards infinity and that the ambush minimal cut depends on the value of $R$.

**Theorem 1.** The maximum number of $2R$ well separated paths in the domain is equal to the length of the shortest T-B path in the ambush, i.e. $l_R(\Omega)^* = l_R(\gamma^*)$, where $\gamma^*$ is a minimal cut of $\Omega$ with respect to measure $l_R$. Moreover, such a set of WSP can be constructed in $O(n \log n + nh)$, where $\Omega$ is a $n$-gon with $h$ holes.

**Proof.** The existence of such a set of well separated paths for BLUE is discussed by Polishchuk [8], although no formal proof is provided. He states that the number of well separated paths can be significantly larger than the number of thick paths. The number of thick paths can be identified using the uppermost path algorithm described in [8], which runs in $O(n \log n + nh)$. This is the Discrete Continuous MaxFlow-MinCut theorem (DCMFMC).

Using a reasoning similar to the proof of the Theorem 6.21 [8], we establish Theorem 1 using the “grass fire” analogy first introduced by Mitchell [6], illustrated in Figure 4. Suppose that the free space $\Omega \setminus H$ is grass over which fire travels at speed 1. Suppose also that the holes are composed of a highly flammable material (the fire travels through a hole at infinite speed) so that as soon as the fire hits a hole, its entire boundary immediately ignites. Ignite the top $T$ at time 0 and initiate the set of $2R$-WSP with the path adjacent to $T$. The wavefront at time $\tau$ is the boundary of the grass that has burnt by $\tau$. The wavefronts make up the streamlines of the flow. The algorithm fills up the free space with the streamlines as the fire propagates until reaching the bottom $B$ so that no more streamlines can be found. The significant
The wavefronts make up the streamlines of the flow. After hitting an obstacle, the streamlines start going over, treating the wavefront at \( R \) as new \( T \).

The ambushes should be set up on the segments where \( \rho \) is continuous ambush game is well separated paths needed for BLUE’s strategy. Each path is used with probability \( 1/7 \).

Although the existence of BLUE’s well separated paths is proven only for polygonal environments, the authors believe that Theorem 2 remains true for any bounded environment \( \Omega \), where the traveler goes from a source \( s \) to a sink \( t \), placed on the boundary of \( \Omega \) or not. However, the computation of the minimal cut and well separated paths that follow on from this graph, become much more strenuous.

The authors would like to emphasize that a theorem similar to Theorem 2 can be proven for games played on graphs \((|N|,|E|)\) as an ambush where BLUE’s strategy space is the set of all mappings from \( E \) to \([-1,1]\] satisfing the flow constraints (3) and RED’s strategy is the set of all mappings from \( N \) to \([-1,1]\] such that \( \sum_{j \in N} q_j = 1 \), meaning that ambushes happen only at nodes. The optimal outcome of this game is \( c^* \), where \( c^* \) is the maximum number of well separated paths from \( s \) to \( t \), ans is equal to the capacity of the minimal cut of the graph.

Theorem 3 is of interest to us because it relates the optimal outcome of the sampled game described in Section III to the theory behind minimal cut computation for graphs, which is a very thoroughly studied topic in Algorithmic. Later algorithms might be developed in order to solve ambush games using this relationship.

### III. Sampled Continuous Ambush Game

We now propose to solve the ambush avoidance problem numerically through uniform sampling over the environment. Prior research ([10], [11], [12]) relies on a single graph to describe both RED and BLUE’s strategy. BLUE’s strategy space is the set of all edges in the graph, while RED’s strategy space is the set of all nodes. In these setups, ambushes only take place at nodes, therefore RED is unable to intercept flows neighboring his ambush even if they are arbitrarily close to its ambush point. Another approach [13]
uses a fixed tiling of the environment in order to take RED’s reach radius $R$ into account. However the tiling is origin dependent and the optimal strategy for both RED and BLUE might vary largely depending on the location of the tiling origin. Moreover, the tile shape cannot be circular, resulting in an undesired non-uniform impact range.

In the formulation below, a parameter $\hat{R} > 0$ is introduced, as in [13], that defines the radius of the region that RED controls. The parameter $\hat{R}$ is the same as that in Section II. Working directly with circular reach zones removes the foregoing limitation and makes this discrete optimization problem closer to the continuous problem defined in Section II

Let $G = (N,E)$ be a graph and let $\mathcal{R} = \{a_j\}_{j \in \mathbb{Q}}$ be a finite set of ambush points spanning the environment. The nodes in $N$ and the ambush points are sampled according to a uniform probability distribution over the environment. The topic of graph representation of an environment is a thoroughly explored field and one can imagine using any of the numerous available techniques to create $G$, such as those developed for Probabilistic Roadmaps [14] or Rapidly-Exploring random Graphs [15].

We define the Sampled Continuous Ambush Game as follows. The traveler (BLUE) is trying to go from a set of nodes $s$ to a set of nodes $t$ through $G$. He can pick any node in $s$ as a starting node and any node in $t$ as a finishing node. The adversary (RED) sets up an ambush at a point $a_j$ in $\mathcal{R}$. The ambusher receives a reward of 1 if BLUE’s path comes through a node within a distance $R$ of $a_j$. A strategy for player BLUE is a mapping $p$ from $E$ to $[0,1]$ such that $0 \leq p(i,j) \leq 1$ and the flow constraints, defined later in this section, are satisfied, where $i$ and $j$ are node indices. A strategy for RED is a mapping $q$ from $\mathcal{R}$ to $[0,1]$ such that $0 \leq q(j) \leq 1$ and $\sum_{j \in \mathbb{Q}} q_j = 1$. Let $p$ be the vector representation of the image of the set of all edges $E$ by the mapping $p$ and $q$ is the image of $\mathcal{R}$ by the mapping $q$. The probability that BLUE use edge $e_{ij}$ is denoted by $p_{ij}$. The probability that RED set up an ambush at point $a_j$ is denoted by $q_j$.

Assume that the two players strategies are independent. At each ambush point $a_j$, the probability that BLUE be ambushed is equal to the probability that BLUE’s path enters the circle $C_R(a_j)$, circle of radius $\hat{R}$ centred at $a_j$, multiplied the probability that RED set an ambush at this point. The expected outcome of the game relative to this ambush point is

$$\mathcal{V} = \sum_{j \in \mathbb{Q}} \sum_{i \in G_j^R} p_{ij} q_j = q^T D p,$$

where $E_j^R$ is the set of all edges $e_{il}$ in $E$, directed between nodes $n_l$ and $n_i$, such that $n_l \notin C_R(a_j)$ and $n_i \in C_R(a_j)$. Therefore the strategic outcome of the game is

$$\mathcal{V} = \sum_{j \in \mathbb{Q}} \sum_{i \in G_j^R} p_{ij} q_j = q^T D p,$$

with $D_{jk} = 1$ if $e_{il} \in E_j^R$, and $D_{jk} = 0$ otherwise, when the $k^{th}$ line of $p$ represents the probability that BLUE uses edge $e_{il}$.

The objective of the proposed approach is to find a strategy for BLUE that minimizes the largest possible outcome for RED. Provided that $q_j \leq 1$ for all $j$, RED can always maximize $\mathcal{V}$ by choosing the point $a_j$ for which the probability of BLUE passing through $C_R(a_j)$ is maximal. Therefore BLUE’s optimal solution is to minimize this product across all nodes:

$$p^* = \arg\min_p \left\{ \max_{j \in \mathbb{Q}} \sum_{i \in G_j^R} p_{ij} q_j \right\}. \quad (2)$$

There are additional constraints to enforce the conservation of the flow of probabilities through the network. The probability that the agent arrives at node $n_j$ is equal to the probability that the agent leaves the same node. The probabilities of the agent leaves the set of origin nodes and arrives at the set of destination nodes are equal to 1. These conditions can be expressed as the linear constraints below.

$$\begin{align*}
\sum_{i \in G_j^R} p_{ij} &= \sum_{k \in \mathbb{Q}} p_{jk}, & \forall j \in N \setminus \{s,t\} \\
\sum_{k \in \mathbb{Q}} p_{kj} &= 1, & \forall k \in \mathbb{Q}
\end{align*} \quad (3)$$

This problem is solved as a linear program problem by introducing the slack variable $z$ satisfying:

$$z \geq \sum_{i \in G_j^R} p_{ij} q_j, \, \forall j \in N \quad (4)$$

Rewriting Equations (2), (3) and (4), the problem can be posed as the linear program

$$\begin{array}{ll}
\text{minimize} & z \\
\text{subject to} & Ap - b^T z \leq 0, \\
& Ap = b, \\
& p \geq 0.
\end{array} \quad (5)$$

where $A$ and $b$ are chosen to appropriately represent the flow conservation constraints. The probability of each edge being used is computed as to minimize expected losses. The computational cost of the resolution of this problem using the Interior Point Algorithm is polynomial in the number of edges of the network. Although possibly more efficient, Polischuk’s approach requires that the environment be polygonal and that a complete description be available. While developed on polygonal environments for benchmarking purposes, the approach in this section can be applied to more complex environments. Similarly to other sampling based planning algorithms, this method only requires the existence of an obstacle logical predicate.

A very interesting feature of the Sampled Continuous Ambush Game is that nothing prevents us from extending it to higher dimensional environments. Indeed, one could imagine creating a graph representation of a $n$-dimensional environment in the exact same way it would be done for a two dimensional environment.

The linear program formulation described above requires the existence of a network to optimize the routing policy. Any
Fig. 6: Optimal policy for RED and BLUE obtained through LP optimization on a sampled network with 25 nodes (left), 81 nodes (center) and 900 nodes (right) when $R = 1/2$. The theoretical optimal value of $V = \frac{1}{4} = \frac{1}{\Omega_{1/2}(N)}$ related to this environment is reached once the graph is dense enough to include 4 well separated paths.

graph representation of the environment might be suitable for this purpose. In Section IV, we are interested in incremental sampling approaches for graph creation. In particular, we are interested in observing the influence of the number of nodes in $G$ and the number of ambush points on the optimal value of the Sampled Continuous Ambush Game.

IV. RESULTS

This section compares the optimal outcome obtained through the approach developed in Section III with the theoretical optimal outcome in order to check the convergence of the sampling approach to the theoretical result. This comparison is done on polygonal environments in order to apply existing work on critical graph and minimal cut computation, as discussed in Section II. The Linear Program described in Equation 5 is implemented on incrementally larger networks in order to understand the influence of the graph density.

Note that, for the simulations discussed in this section, the network was created through a uniform sampling in the form of an eight-connected grid. Other network creation methods have been considered and used in previous work [13], such as random sampling with Delaunay triangulation [16] or grid sampling with Delaunay triangulation. Given the fact that all methods gave similar results, only the eight-connected grid will be discussed in this analysis.

We first look at the example environment presented in Figure 3. As shown above, the ambush-minimal cut capacity of this environment for $R = 1/2$ is $l_{1/2}(\gamma_{1/2}) = 4$. Therefore the optimal outcome of the Continuous Ambush Game for this environment is $1/4$. It can be observed in Figure 6 that for smaller, and therefore coarser, networks the full connectivity of the environment is not captured by the network, resulting in a number of 2-$R$-WSP lower than the theoretic maximum of 4. For a network with only twenty-five nodes, there can be more than two 1-WSP, therefore the optimal outcome of the Sampled Continuous Ambush Game is $1/2$, as seen in Fig. 6 (left). As the number of nodes in $N$ increase however, more well separated paths can be found in $G$, and the optimal outcome decreases to $1/3$ first (for $|N| = 81$ nodes, Fig. 6 (center)) and finally to $1/4$ when the number of nodes in $N$ becomes sufficiently large ($|N| = 900$ nodes, Fig. 6 (right)). Since the ambush-minimal cut capacity of the environment is 4, there cannot be more than four 1-

WSP in the environment. Therefore there cannot be more than four 1-WSP in the graph either and we can conclude that increasing the grid resolution will not change the optimal value of the Sampled Continuous Ambush Game.

The convergence of the optimal value of the game has been illustrated on one example of polygonal environment. In order to verify that this convergence could be observed on any polygonal environments, a benchmark is created to compute the optimal outcome of the Sampled Ambush Game on a set of fifty randomly generated polygonal environments for different grid resolutions. For each environment, the linear program created from a graph of $n$ nodes outputs the value of the optimal outcome, while $n$ varies from nine to nine hundreds. The average outcome over fifty environments is computed as a function of $n$ and is displayed in Figure 7. These environments were generated by randomly varying the number of obstacles, their sizes and their locations. The environment displayed in Figures 3 and 6 is an instance of these environments.

Again, we observe that the optimal outcome of the Sampled Continuous Ambush Game is a decreasing function
of the number of nodes in the graph representation of the environment. Additionally, it converges to a value, which is only dependent on the structure of the environment, equal to the ambush-minimal cut capacity of the environment. We have empirically verified the following conjecture.

**Conjecture 1.** The optimal outcome of the Sampled Continuous Ambush Game converges to the optimal outcome of the Continuous Ambush Game as the number of sample points in the environment goes towards infinity.

Based on Theorem 2 and Theorem 3, Conjecture 1 suggests a relationship between the continuous minimal cut capacity of a polygonal environment and the asymptotic capacity of the cut of a network with increasingly many nodes. This Conjecture and its implications will be explored in more detail in future work.

**V. DISCUSSION**

This paper develops a theory of continuous ambush games for non-trivial environments. The problem of ambush games is of particular interest for developing navigation solutions for autonomous agents in adversarial environments, such as delivery vehicles or vehicles operating in warfare regions. Two complementary approaches are adopted.

First, a truly continuous approach, requiring the simplifying but realistic assumption that the environment is polygonal, allows us to compute theoretical optimal solutions. The optimal value of a continuous ambush game, as defined in Section II, is \( R(\Omega) \), where \( l_\Omega(\Omega) \) represents the maximum number of paths from \( s \) to \( t \) separated by at least \( 2R \) at any point, or \( 2R-WSP \). The optimal strategies for both agents depend solely on the geometry of the environment and on the reach \( R \) of the adversary. Second, the problem of ambush games in a continuous environment is tackled by sampling the environment and creating a linear problem based on the resulting graph, as presented in Section III. The optimal strategy returned by the linear program is shown to converge towards the theoretical optimal strategy as the number of node in the graph representation of the environment increases.

Both approaches exhibit interesting features yet have their own limitations. On the one hand, the continuous approach enables a truly continuous comprehension and modeling of the environment studied. Assuming some strong, yet realistic assumptions on the environment, which has to be polygonal, theoretical proofs regarding the optimality of the solution are derived. However, the corresponding algorithms are for now limited to two dimensional environments [8], and extensions to higher dimensions might prove difficult.

On the other hand, the sampling approach only requires an admissibility predicate [17], meaning whether a point of the environment is collision-free with respect to the obstacles and feasible with respect to the constraints. This approach allows for the creation of a graph on which a solution of the ambush game can be obtained by running optimization problems as described in Section III. The upside of this approach is its simple implementation, similar to sampling-based path planning approaches. For example, one could imagine building a lattice or a Rapidly-Exploring Dense Graph ([18],[19]) and running the optimization on the resulting network. Moreover, this approach can easily be generalized to higher dimensions, which could include velocity features for instance, allowing for non uniform reward functions depending on the velocity of the agent. An extension to complex state spaces with a dynamical system standpoint is envisioned as future work. The state space will take into account position and velocity. This work can also be adapted to include a vehicle model and its performance, to restrain the network to only a set of feasible maneuvers. The main limitation is to determine whether the solution of the continuous ambush game and the solution sampled continuous ambush game are the same, and if not, quantifying how large the optimality gap might be.

**REFERENCES**