Cooperative Control for Multi-Target Interception with Sensing and Communication Limitations: A Game Theoretic Approach

Mohammad Khosravi, Hossein Khodadadi, Hassan Rivaz and Amir G. Aghdam

Abstract—In this paper, the problem of multi-vehicle cooperative interception of moving objects with unknown arrival times, trajectories and dynamics is investigated. The vehicles are assumed to have limited sensing and communication ranges. Therefore centralized approaches are not efficient, especially for the case of large number of vehicles and targets. In this work, a game theoretic cooperative receding horizon controller is proposed to address this problem. The design is based on the prediction of the future positions of targets with limited information, as well as a reward allocation policy for accomplishing the target interception tasks. To learn the dynamics in the resulting potential game, state of the art methods such as spatial adaptive play and generalized regret monitoring are applied and their effectiveness is demonstrated.

I. INTRODUCTION

In the last two decades, cooperative control of multi-agent systems has attracted substantial research interest [1]–[3], with applications spanning a wide variety of fields such as biology, control and computer science. In particular, control design problems are investigated for a variety of applications such as surveillance [4], search and rescue missions [5] and reconnaissance [6], to name only a few. The main goal of multi-agent systems is to achieve a global objective with a set of simple and limited components and the proper use of information exchange between the agents.

In recent literature in the area of multi-target interception, more general models for target behavior are considered. In [7], the vehicles, based on dynamic Voronoi partitioning, are assigned to the moving objects with known kinematics. The method in [8] proposes a distributed cooperative strategy toward interception of a set of moving objects with similar unicycle model. In [9], a dynamic decision-making trend is proposed for multi-target interception by a single vehicle in an uncertain environment. In [10], a centralized cooperative receding horizon controller is design for intercepting a set of target moving in the mission space on priori unknown trajectories. In all of the aforementioned papers, there are some restrictive assumption: in [11], [12], the target objects are assumed stationary points, in [9], [13] a single vehicle is considered for mission accomplishment, in [13] a simple model for target movements are assumed, and in [7], [8], arrival times of the targets are assumed to be known.

In this paper, a Cooperative Receding Horizon Controller is designed for heading control of a set of vehicles toward intercepting targets, arriving the mission space in priori unknown times and priori unknown positions, and also moving with priori unknown dynamics. Here, similar to [10], a team of vehicles are supposed to capture a set of targets moving with unknown trajectories, and further generalize it by assuming that the arrival positions and times are unknown. Moreover, vehicles have limited ranges for sensing the targets and also limited ranges for communication, i.e. each vehicle can only sense the targets being in a region around it and also communicate only with vehicles with distance less than some ranges. Dealing with this level of uncertainties in the environment and vehicles limitations, a distributed online controller using receding horizon is required. Toward this goal, the method introduced in [10] has been extended by exploiting recent developments in games theory. In this approach, each of the targets is assigned a time decreasing reward, which is collectible only if the target is visited by some vehicles, and considering these rewards and problem constraints, a utility function is designed with respect to each vehicle. The resulting structure forms a potential game with total collectible reward as its potential functions. Using appropriate learning dynamics, vehicles decide upon their strategies and consequently on their headings.

II. NOTATIONS AND PRELIMINARIES

A. Notations

Throughout the paper, $\mathbb{N}$, $\mathbb{R}$, $\mathbb{R}_{\geq 0}$ respectively denote the set of natural numbers, real numbers, and non-negative real numbers. Also, the set of natural numbers less than or equal to $n$ is denoted by $\mathbb{N}_n$. For a given a set $A$, and subset of it like $B$, the indicator function of $B$ is denoted by $1_B$ which is a function from $A$ to $\{0, 1\}$:

$$1_B(x) := \begin{cases} 1, & \text{if } x \in B, \\ 0, & \text{if } x \notin B, \end{cases}$$

Let $I$ be a set of indices, and $(a_i)_{i \in I}$ represent a point in $A^I$ with entries $a_i$. Also, let $J$ be a non-empty set of indices such that $J \subseteq I$. Subsequently, for any point in $a \in A^I$, $a_{|J}$ represents a point in $A^J$, obtained by eliminating the entries with indices not listed in $J$.

The $d$ dimensional Euclidean space is denoted by $\mathbb{R}^d$. Also, all-zero and all-one vectors in $\mathbb{R}^d$ are respectively represented by $0_d$ and $1_d$. For any vectors $a$ and $b$ in $\mathbb{R}^d$ inequality $a \geq b$ indicates that all entries of $a - b$ are non-negative. For any point $x \in \mathbb{R}^d$ and any scalar $r \in \mathbb{R}_{\geq 0}$, the closed ball with radius $r$ centered at $x$ is denoted by $B(x, r)$, and is defined as

$$B(x, r) = \{ y \in \mathbb{R}^d \mid \|x - y\| \leq r \}. \quad (1)$$
Denote by $C^p_{\mathbb{R}_+}({\mathbb{R}}^d)$ as the set of piecewise continuous functions defined over $\mathbb{R}_+\mathbb{R}$ and taking values from $\mathbb{R}^d$.

A bipartite graph is a graph like $\mathcal{G} = (U \cup V, E)$, where the set of its vertices is union of two disjoint partitions $U$ and $V$, such that there is no pair of adjacent vertices in each of partitions. For any bipartite graph $\mathcal{G} = (U \cup V, E)$, its biadjacency matrix is defined as $[\mathcal{U}]$ by $[V]$ matrix $B = (b_{ij})$ with 0 and 1 entries, such that $b_{ij} = 1$ if and only if $i$th vertex in $U$ be adjacent with $j$th vertex in $V$.

B. Preliminaries

A game of $n$ players is defined as $(\mathbb{N}_n, \times \subset \subset \subset \mathbb{N}_n, A_i, \{U_i\} \in \mathbb{N}_n)$, where $\mathbb{N}_n$ is the set of indices for players, $\times \subset \subset \subset \mathbb{N}_n, A_i$ is the set of action profiles and, for any $i \in \mathbb{N}_n$, $A_i$ is the action set and $U_i : \times \subset \subset \subset \mathbb{N}_n, A_i \rightarrow \mathbb{R}$ is the utility function, for the $i$th player. For any $i \in \mathbb{N}_n$ and any action profile $(a_j)_{j \in \mathbb{N}_n} \in \times \subset \subset \subset \mathbb{N}_n, A_j$, let $a_{i-}$ and $(a_i, a_{i-})$ denote $(a_j)_{j \neq i}$ and $(a_i)_{\mathbb{N}_n \backslash \{i\}}$, respectively.

Definition 1 ([17]): The game $(\mathbb{N}_n, \times \subset \subset \subset \mathbb{N}_n, A_i, \{U_i\} \in \mathbb{N}_n)$ is a potential game, if there exists a function $\phi : \times \subset \subset \subset \mathbb{N}_n, A_i \rightarrow \mathbb{R}$, called potential function, such that for any $i \in \mathbb{N}_n$, any actions $a_i', a_i'' \in A_i$ and any $a_{i-} \in \times \subset \subset \subset \mathbb{N}_n, A_j$, the following relation holds

$$U_i(a_i', a_{i-}) - U_i(a_i'', a_{i-}) = \phi(a_i', a_{i-}) - \phi(a_i'', a_{i-}).$$

III. Problem Formulation

Define mission space as a closed convex subset of $\mathbb{R}^d$, and denote it by $M$. Consider a finite number of objects, referred here as targets, arriving in the mission space sequentially. One can index the targets with respect to their arrival order by indices in $\mathcal{I}_T = \mathbb{N}_{|\mathcal{I}_T|}$, where $|\mathcal{I}_T|$ is the number of targets. Without loss of generality, it can be assumed that the mission starts at time $t = 0$, where $n_0 \in \{0\} \cup \mathbb{N}_{|\mathcal{I}_T|}$ is the initial number of targets in the mission space. Let $\tilde{T}_i$ denote the arrival time of target $i$ and set the finite sequence of non-negative real scalars $\{T_i\}_{i=1}^{T_1}$ as the time sequence between consecutive targets arrivals, i.e. $T_{i+1}$ represents the interarrival time between $i$th and $(i+1)$th targets, for any $i \in \mathbb{N}_{|\mathcal{I}_T|}$. Considering the definition of $\{T_i\}_{i=1}^{T_1}$, arrival time of target $i$, for any $i \in \mathcal{I}_T$, is denoted by $\tilde{t}_i$ and defined as following

$$\tilde{t}_i = \begin{cases} 0, & i \leq n_0, \\ \sum_{j=0}^{i-n_0} T_j, & i > n_0. \end{cases}$$

(2)

Also, for any $t \in \mathbb{R}_{\geq 0}$, denote the set of indices of targets arrived up to time $t$ by $\tilde{I}_T(t)$, and define it as

$$\tilde{I}_T(t) := \{i \in \mathcal{I}_T : \tilde{t}_i \leq t\}.$$  

(3)

For any $i \in \mathcal{I}_T$, let $y_i \in M$ be the initial position of target $i$ as it arrives in the mission space. Thus, $\{y_i\}_{i \in \mathcal{I}_T}$ is the finite sequence of arrival positions of targets. Also, since the targets are assumed to be a set of moving objects in the mission space, by abuse of notation, one can set $y_i(\cdot)$ as the trajectory of $i$th target after its arrival, for any $i \in \mathcal{I}_T$. It can easily be seen that given any $i \in \mathcal{I}_T$, one has $y_i = y_i(\tilde{t}_i)$. More details of properties of $\{y_i(\cdot)\}_{i \in \mathcal{I}_T}$ are discussed later in this section.

Assumption 1: The arrival times are priori unknown and at any moment of time, there is no exact information regarding future of trajectories of targets. More precisely, there is no information available on $\tilde{t}_i$ and $y_i(\cdot)$, at any given time $t \in \mathbb{R}_{\geq 0}$ prior $\tilde{t}_i$.

Besides the targets, consider a finite number of vehicles inside $M$ with indices in $\mathcal{I}_V = \mathbb{N}_{|\mathcal{I}_V|}$, where $|\mathcal{I}_V|$ represents the number of vehicles. For any $j \in \mathcal{I}_V$, denote $x_j(t)$ as the position of vehicle $j$ in mission space at given time $t$. Also, let dynamics of $x_j(t)$ be described by

$$\dot{x}_j(t) = u_j(t),$$

(4)

where $u_j$ belongs to $\mathcal{U}_{\text{max}}$, the set of admissible controls, defined as

$$\mathcal{U}_{\text{max}} = \{ u : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^d ; u \in C^p_{\mathbb{R}_+}({\mathbb{R}}^d), \|u\|_{\sup} \leq u_{\text{max}} \},$$

(5)

where $C^p_{\mathbb{R}_+}({\mathbb{R}}^d)$ is the set of piecewise-continuous functions, defined over $\mathbb{R}_{\geq 0}$ and taking values in $\mathbb{R}^d$. Note that here $u_{\text{max}}$ is a non-negative real scalar independent of $j$. It can be seen that for some piecewise continuous functions $u_j$ and $d_j$, one has $u_j(t) = u_j(t)d_j(t)$, where $u_j(t) \in [0, u_{\text{max}}]$ is the control input for magnitude of velocity for vehicle $j$ and $d_j(t) \in S^{d-1} = \{d \in \mathbb{R}^d; \|d\| = 1\}$ is the control input for its direction, for any $j \in \mathcal{I}_V$ and at any given time $t \in \mathbb{R}_{\geq 0}$.

For any $j \in \mathcal{I}_V$, denote $\mathcal{S}_j(t)$ as sensing region of vehicle $j$ at given time $t \in \mathbb{R}_{\geq 0}$ which is defined as

$$\mathcal{S}_j(t) = \{ x \in \mathbb{R}^d ; \| x - x_j(t) \| \leq r_j \} = \mathcal{B}(x_j(t); r_j),$$

(6)

where $r_j \in \mathbb{R}_{\geq 0}$ is a scalar representing the radius of sensing for the vehicle. Similarly, for any $j \in \mathcal{I}_V$ and any $t \in \mathbb{R}_{\geq 0}$, define communication region of vehicle $j$, denoted by $\mathcal{C}_j(t)$, as following

$$\mathcal{C}_j(t) = \{ x \in \mathbb{R}^d ; \| x - x_j(t) \| \leq r_c \} = \mathcal{B}(x_j(t); r_c),$$

(7)

where $r_c \in \mathbb{R}_{\geq 0}$ is a scalar representing the radius of communication for the vehicle. Here, it is assumed that for any $i \in \tilde{I}_T$ and any $j \in \mathcal{I}_V$, vehicle $j$ is capable of sensing target $i$, at the given time $t \in \mathbb{R}_{\geq 0}$, if the target $i$ be in the sensing region of the vehicle $j$, i.e. the condition $y_i(t) \in \mathcal{S}_j(t)$ is satisfied. Also, for any pair of targets with indices $j_1, j_2 \in \mathcal{I}_V$ and any moment of time $t \in \mathbb{R}_{\geq 0}$, vehicle $j_2$ can receive information sent by vehicle $j_1$, if the vehicle $j_2$ be in the communication region of the vehicle $j_1$, i.e. $x_{j_2}(t) \in \mathcal{C}_{j_1}(t)$ is satisfied.

Remark 1: Note that for any $j_1, j_2 \in \mathcal{I}_V$ and for any $t \in \mathbb{R}_{\geq 0}$, one has that $x_{j_2}(t) \in \mathcal{C}_{j_1}(t)$ if and only if $x_{j_1}(t) \in \mathcal{C}_{j_2}(t)$.

Definition 2: For any $i \in \mathcal{I}_T$ and any $j \in \mathcal{I}_V$, the vehicle $j$ is said to visit the target $i$ at time $t$, if $|x_j(t) - y_i(t)| \leq d_{ij}$, where $d_{ij}$ is a given positive scalar in $\mathbb{R}_{\geq 0}$.

Remark 2: It is worth to note that, the scalar $d_{ij}$ is introduced in Definition 2 for considering practical issues such as the physical size of the targets and vehicles. In fact,
having that targets and vehicles expressed as some points of mass, the scalar $d_{ij}$ can be used to account for the size of the target $i$ and the size of the vehicle $j$. More precisely, if for any $i \in \mathcal{I}_T$ and $j \in \mathcal{I}_V$, the target $i$ and vehicle $j$ approximately spherical objects in $\mathbb{R}^d$, with radius $r_i$ and radius $s_j$, respectively, then one may set $d_{ij} = r_i + s_j$. This clarifies the reason behind the introdution of $d_{ij}$.

For any $i \in \mathcal{I}_T$, define the first visit time of target $i$, denoted by $\tilde{\tau}_i \in \mathbb{R}_{\geq 0}$, as the time where $i$ is visited by one of the vehicles for the first time, i.e.

$$\tilde{\tau}_i = \inf \{ t \in \mathbb{R}_{\geq 0} : \min (|x_j(t) - y_i(t)| - d_{ij}; j \in \mathcal{I}_V) \leq 0 \}. \quad (8)$$

Note that, for any $i \in \mathcal{I}_T$, being first visit time of target $i$ infinity, $\tau_i = \infty$, is equivalent to the situation where none of the vehicles visit target $i$. Based on the definition of first visit times, for any $t \in \mathbb{R}_{\geq 0}$, one can define the set of indices of targets visited up to time $t$ as following

$$\hat{\mathcal{I}}_T(t) = \{ i \in \mathcal{I}_T | \tilde{\tau}_i \leq t \}. \quad (9)$$

Similarly, denote $\mathcal{I}_T(t)$ as the set of indices of targets arrived and not visited up to time $t$:

$$\mathcal{I}_T(t) = \hat{\mathcal{I}}_T(t) \backslash \hat{\mathcal{I}}_T(t) = \{ i \in \mathcal{I}_T ; \tilde{\tau}_i \leq t < \tilde{\tau}_i \}. \quad (10)$$

For any $i \in \mathcal{I}_T$, the trajectory of target $i$ is assumed to be a function $y_i : [\tilde{\tau}_i, \tau_i] \rightarrow \mathbb{R}^d$ satisfying local and global geometric properties introduced as follows.

**Assumption 2:** (Geometric Conditions) Let $i \in \mathcal{I}_T$. Then:

1. (Global Geometric Condition) For any $\tau \in [\tilde{\tau}_i, \tau_i]$, one has $y_i(\tau) \in \mathcal{M}$, i.e. $y_i([\tilde{\tau}_i, \tau_i]) \subset \mathcal{M}$.

2. (Local Geometric Condition) $y_i : [\tilde{\tau}_i, \tau_i] \rightarrow \mathcal{M}$ is a continuously differentiable function. Moreover, there exist non-negative scalars $\hat{\alpha}_{\max}, \hat{\alpha}_{\max}$, such that for any $\tau \in [\tilde{\tau}_i, \tau_i]$, one has

$$\left\| \frac{d}{dt} y_i(\tau) \right\| \leq \hat{\alpha}_{\max}. \quad (11)$$

If $\alpha_i(t, \tau)$ be the function behaving the following equality

$$y_i(t) = y_i(\tau) + \frac{d}{dt} y_i(\tau)(t - \tau) + \frac{1}{2} \alpha_i(t, \tau)(t - \tau)^2, \quad (12)$$

where $t \in [\tau, \tilde{\tau}_i]$, then

$$\sup_{s \in (\tau, \tilde{\tau}_i)} \| \alpha_i(s, \tau) \| \leq \tilde{\alpha}_{\max}. \quad (13)$$

The global geometric condition assures that once a target arrives in the mission space, it will remain inside it. One should note that this property depends not only on trajectories of targets, but also on the geometry of the mission space.

**Remark 3:** For the particular case where mission space is $d$-dimensional Euclidean space, the global geometric condition is immediately satisfied.

Consider the case $y_i(\cdot)$ is a $C^2$ function, for any $i \in \mathcal{I}_T$, and there exist non-negative scalars $\tilde{\alpha}_{\max}, \tilde{\alpha}_{\max}$ such that

$$\left\| \frac{d}{dt} y_i(\tau) \right\| \leq \tilde{\alpha}_{\max}, \quad \left\| \frac{d^2}{dt^2} y_i(\tau) \right\| \leq \tilde{\alpha}_{\max}, \quad (14)$$

for any $i \in \mathcal{I}_T$ and $\tau \in [\tilde{\tau}_i, \tau_i]$. One can see from Taylor’s theorem with mean-value form of the remainder [14], that $y_i(\tau)$ satisfies Local Geometric Condition in Assumption 2.

**Assumption 3:** For any $\tau \in \mathbb{R}_{\geq 0}$ and $i \in \mathcal{I}_T(\tau)$, the position and velocity vectors of target $i$ are available at the beginning of each time horizon, i.e. at time instant $\tau$ in [12].

For any $\tau \in \mathbb{R}_{\geq 0}$ and $i \in \mathcal{I}_T(\tau)$, using the Assumption 3 one can estimate the positions of the target $i$ for any future instant $t \in [\tau, \tilde{\tau}_i]$ as

$$\hat{y}_i(t) = y_i(\tau) + (t - \tau) \frac{d}{dt} y_i(\tau). \quad (15)$$

**Remark 4:** It can be easily seen that from (13) that

$$\|\hat{y}_i(t) - y_i(t)\| \leq \frac{1}{2^2} (t - \tau)^2 \sup_{t \in (\tau, \tilde{\tau}_i)} \|\alpha_i(t, \tau)\| = \frac{1}{2^2} (t - \tau)^2 \tilde{\alpha}_{\max} \quad (16)$$

for any $t \in [\tau, \tilde{\tau}_i]$. Thus, choosing $t$ such that $t - \tau$ be small enough, one can obtain desired precision in estimation [15].

One can define a task with respect to each target, which can be accomplished only if one of the vehicles visits the respective target in finite time. By abuse of notation, denote $\hat{\mathcal{I}}_T$ as the set of all the tasks, $\hat{\mathcal{I}}_T(t)$ as the set of tasks started by time $t$, $\hat{\mathcal{I}}_T(t)$ as the set of tasks accomplished by time $t$, and $\mathcal{I}_T(t)$ as the set of tasks which are in process at time $t$. Finally, one may define the mission as accomplishment of all the tasks. Respecting the possible uncertainties and the limitations on information in the discussed paradigm, it is desired to obtain a near-optimal cooperative control policy for mission accomplishment, which is discussed in subsequent sections.

### IV. Game Theoretic Cooperative Receding Horizon Scheme

In order to encourage the vehicles to visit the targets, with respect to each task, a time-decreasing reward is considered for every target that can be collected only if the target is visited and hence the task is accomplished. The vehicles are expected to dynamically make their decisions towards maximizing the total collected rewards. The decision-makings process of every vehicle, which is iterative, consists of planning their paths and also deciding upon their strategies regarding vehicles and targets assignments. At the beginning of each iteration, vehicles update their information by checking their sensing regions, communicating with their neighbors, and then calculating the headings, in order to update the strategies.

#### A. Reward Allocations

For any $i \in \mathcal{I}_T$, let $\mathcal{R}_i$ be the initial reward considered for accomplishment of task $i$. Define $\mathcal{D}_i : \mathbb{R}_{\geq 0} \rightarrow [0, 1]$ as the time discount function, which is a decreasing function capturing the rate of reward loss over time. The reward function of task $i$ is obtained as $\mathcal{R}_0 \mathcal{D}_i$, for any $i \in \mathcal{I}_T$. There exist various suitable candidates for the time discount function for modeling different aspects of timing and scheduling, such...
as deadlines and priorities. As a simple example, one can consider the following function
\[ q_i(t) = e^{-\gamma_i t}, \quad \forall i \in \mathcal{I}_T \]  
(17)
where \( \gamma_i \in \mathbb{R}_{>0} \) is a parameter reflecting the degree of importance of target \( i \).

**B. Cooperative Structure**

In each step of decision making, vehicles are expected to plan task assignments based on the available information. For any \( i \in \mathcal{I}_T \) and \( j \in \mathcal{I}_V \), an assignment of vehicle \( j \) to task \( i \) is characterized as a real scalar in \([0, 1] \), denoted by \( a_{ij} \), which reflects the amount of interest of vehicle \( j \) in being assigned to target \( i \), and depends on the rewards, positions of the vehicles and positions of targets. Note that
despite the infinite sensing and communication capability of vehicles, there may be several constraints on these assignments. Therefore, here the case with unlimited sensing and communication range is investigated first and then, the discussion is generalized to the case with limited sensing and communication ranges.

1) **Unlimited Sensing and Communication Ranges:** In this case, each vehicle is capable of sensing all the targets and is also capable of communicating with any other vehicle. Thus, at any point in time, each of the assignments depends on the positions of all vehicles and targets. More precisely, for any \( i \in \mathcal{I}_T(t) \) and \( j \in \mathcal{I}_V \), the assignment \( a_{ij} \) can be a function of the following form:
\[ a_{ij} : \mathcal{M}^{[\mathcal{I}_V]} \times \mathcal{M}^{[\mathcal{I}_T(t)]} \rightarrow [0, 1]. \]  
(18)
Various methods can be used for designing the assignment functions, such as Voronoi-based assignment policy [15] and competition-based assignment [11], discussed in [10]. The desired assignments are expected to have some structures which are discussed in the sequel.

Let \( t \) be a time instant at which \( \mathcal{I}_T(t) \neq \emptyset \). In order for the current tasks to be considered by every vehicle, it is required to have the sum of task assignments for each vehicle be equal to one, i.e.
\[ \sum_{j \in \mathcal{I}_V} a_{ij}(x, y) = 1, \quad \forall j \in \mathcal{I}_V, \]  
(19)
where \( x = (x_j)_{j \in \mathcal{I}_V} \) and \( y = (y_i)_{i \in \mathcal{I}_T(t)} \) are vectors of vehicle positions and targets positions, respectively. Also, if the number of vehicles is at least equal to the number of current tasks, i.e. \( |\mathcal{I}_V| \geq |\mathcal{I}_T(t)| \), it is reasonable to increase the chances of accomplishing tasks by generously over-assigning the targets to the vehicles, i.e.
\[ \sum_{j \in \mathcal{I}_V} a_{ij}(x, y) \geq 1, \quad \forall i \in \mathcal{I}_T(t). \]  
(20)
Similarly, if the number of current tasks is at least equal to the number of vehicles, i.e. \( |\mathcal{I}_T(t)| \geq |\mathcal{I}_V| \), in order to accomplish as many tasks as possible, it is preferred to manage the resources efficiently by acting cautiously and under-assigning the targets to the vehicles, i.e.
\[ \sum_{j \in \mathcal{I}_V} a_{ij}(x, y) \leq 1, \quad \forall i \in \mathcal{I}_T(t). \]  
(21)
It is worth noting that in equations (20) and (21), if \( |\mathcal{I}_V| = |\mathcal{I}_T(t)| \), then the inequalities will turn to equalities.

From the above discussion, any desired assignment such as \( \mathbf{A}(x, y) = (a_{ij}(x, y))_{i \in \mathcal{I}_T(t), j \in \mathcal{I}_V} \) should satisfy the set of constraints introduced in (19), (20) and (21). In other words, defining set \( \mathcal{A}^{n \times m} \) for any \( n, m \in \mathbb{N} \), as
\[ \mathcal{A}^{n \times m} = \{ \mathbf{A} \in [0,1]^{n \times m} ; \mathbf{A}^{\top} \mathbf{1}_n = \mathbf{1}_m, \]  
\[ m \leq n \Rightarrow \mathbf{A} \mathbf{1}_m \leq \mathbf{1}_n \}, \]  
(22)
then \( \mathbf{A}(x, y) \) belongs to \( \mathcal{A}^{\mathcal{I}_T(t),\mathcal{I}_V} \) which is defined as follows:
\[ \mathcal{A}^{\mathcal{I}_T(t),\mathcal{I}_V} = \{ \mathbf{A} : \mathcal{M}^{[\mathcal{I}_V]} \times \mathcal{M}^{[\mathcal{I}_T(t)]} \rightarrow \mathcal{A}^{[\mathcal{I}_T(t)] \times [\mathcal{I}_V]} \}. \]  
(23)

2) **Limited Sensing and Communication Ranges:** Here, unlike the previous case, some of the vehicles may sense only a subset of targets, (with anything between 0 to \( |\mathcal{I}_T(t)| \) at time \( t \)), and not all of the targets. Moreover, some pairs of vehicles may not be able to communicate with each other directly. These constraints need to be addressed in the assignment functions. To this end, the notion of virtual targets is introduced, and the definitions of target set and vehicle set are modified accordingly to take the communication and sensing limitations into account.

For any \( j \in \mathcal{I}_V \), denote \( \mathcal{O}_j \) a virtual target which is defined as an imaginary target that can be sensed only by vehicle \( j \), and let \( \mathcal{I}_T \) represent the set of all virtual targets, i.e. \( \mathcal{I}_T = \{ \mathcal{O}_j \}_{j \in \mathcal{I}_V} \). Then, at any time \( t \in \mathbb{R}_{\geq 0} \), the set of targets in sensing range for vehicle \( j \), denoted by \( \mathcal{I}_T,j(t) \), and the set of communicating vehicles with vehicle \( j \), denoted by \( \mathcal{I}_{V,j}(t) \), are defined respectively as
\[ \mathcal{I}_T,j(t) = \{ i \in \mathcal{I}_T(t) ; \quad y_i(t) \in \mathcal{O}_j \} \cup \{ \mathcal{O}_j \}, \]  
(24)
and
\[ \mathcal{I}_{V,j}(t) = \{ j \in \mathcal{I}_V ; \quad x_j(t) \in \mathcal{O}_j \}. \]  
(25)
Similarly, for any \( i \in \mathcal{I}_T \) or \( \mathcal{I}_{V,i}(t) \), the set of sensing vehicles for target \( i \) is defined below
\[ \mathcal{I}_{V,i}(t) = \{ j \in \mathcal{I}_V ; \quad x_j(t) \in \mathcal{O}_i \}. \]  
(26)
One can similarly define the set of targets in all sensing ranges as the targets with respect to which there exists a vehicle whose sensing region covers those targets. Denoting this set by \( \mathcal{I}_T(t) \), it can be expressed as
\[ \mathcal{I}_T(t) = \bigcup_{j \in \mathcal{I}_V} \mathcal{I}_T,j(t). \]  
(27)

**Remark 5:** Since, for any \( j \in \mathcal{I}_V \), one has \( \mathcal{O}_j \in \mathcal{I}_{T,j}(t) \), it yields that
\[ \{ \mathcal{O}_j ; \quad j \in \mathcal{I}_V \} \subseteq \bigcup_{j \in \mathcal{I}_V} \mathcal{I}_{T,j}(t) = \mathcal{I}_T(t), \]  
(28)
which implies \( |\mathcal{I}_V| \leq |\mathcal{I}_T(t)| \).

**Remark 6:** The virtual targets are introduced for the cases where there are no real target in the sensing region of a vehicle. In fact, a vehicle can only be assigned to a virtual
target if it is not assigned to any real target, and is supposed to perform a task other than tracking (e.g. search or returning to a defined depot). Note that such assignment is required if there is no real target in the sensing region of the vehicle. Thus, one needs to set the initial reward and discount factor of the underlying virtual target appropriately.

**Assumption 4:** For any \( j_1, j_2 \in \mathcal{I}_V \), it is assumed that
\[
r_{j_1} + r_{j_2} < r_t.
\] (29)

This is a realistic assumption as the sensing range of a vehicle is typically less than its communication range.

Assumption 4 implies that if a target can be sensed by a pair of vehicles, the two vehicles are able to communicate with each other. More precisely, for any \( i \in \mathcal{I}_T(t) \) and any \( j_1, j_2 \in \mathcal{I}_V \), the following condition holds
\[
j_1, j_2 \in \mathcal{I}_V,i(t) \implies j_1 \in \mathcal{I}_V,j_2 \land j_2 \in \mathcal{I}_V,j_1(t). \] (30)

Analogously to the case of unlimited sensing and communication ranges, for any \( j \in \mathcal{I}_V \) and \( i \in \mathcal{I}_T(t) \), an assignment function of the following form can be considered
\[
a_{ij} : \mathcal{M}[\mathcal{I}_V,i(t)] \times \mathcal{M}[\mathcal{I}_T(t)] \rightarrow [0, 1]. \] (31)

Note that the assignment \( a_{ij} \) depends on the vehicles that are sensing target \( i \) as well as the targets which are sensed by vehicle \( j \), at any time instant \( t \geq 0 \).

Similar to the case of unlimited sensing and communication ranges, the desired assignment need to satisfy certain conditions. For example, each vehicle should normally consider all the targets inside its sensing region. Therefore, for each vehicle, the sum of target assignments in its sensing region at any time instant \( t \), should be equal to one, i.e.
\[
\sum_{i \in \mathcal{I}_T(t)} a_{ij}(x_i, y_j) = 1, \quad \forall j \in \mathcal{I}_V, \] (32)

where \( x_i = (x_j)_{j \in \mathcal{I}_V,t}, y_j = (y_i)_{i \in \mathcal{I}_T,t} \).

**Remark 7:** For the definition of virtual targets, one can verify that for any \( j \in \mathcal{I}_V \), the set of targets in the sensing range of vehicle \( j \) is non-empty, i.e. \( \mathcal{I}_T(t) \neq \emptyset \). Therefore, the summation in (32) is well-defined.

On the other hand, since it is desired to accomplish as many tasks as possible, and also the number of targets in the sensing regions at any point in the time is more than or equal to the number of vehicles, it is important to be conservative and under-assign the vehicles to targets in their sensing regions, i.e.
\[
\sum_{j \in \mathcal{I}_V} a_{ij}(x_i, y_j) \leq 1, \quad \forall i \in \mathcal{I}_T \cup \mathcal{I}_T. \] (33)

**Definition 3:** Define sensing bipartite graph, denoted by \( \mathcal{G}_t = (\mathcal{T}_t \cup \mathcal{V}_t, \mathcal{E}_t) \), as a bipartite graph with vertex partitions \( \mathcal{T}_t = \mathcal{I}_T(t) \) and \( \mathcal{V}_t = \mathcal{I}_V \), and the edge set defined as following
\[
\mathcal{E}_t = \{(i, j) \in \mathcal{T}_t \times \mathcal{V}_t \mid i \in \mathcal{I}_T,j(t)\}. \] (34)

Let \( \mathbf{B}_t \) be the biadjacency matrix of \( \mathcal{G}_t \).

Given a sensing graph \( \mathcal{G}_t \), note that equations (32) and (33) introduce a set of constraints that any desired assignment \( \mathbf{A}(x, y) = (a_{ij}(x, y))_{i \in \mathcal{I}_T(t)} \) should satisfy them, i.e. \( \mathbf{A}(x, y) \) belongs to the set \( \mathcal{A}_{\mathcal{T}_T(t),\mathcal{I}_V} \) defined as
\[
\mathcal{A}_{\mathcal{T}_T(t),\mathcal{I}_V} = \{ \mathbf{A} : \mathcal{M}[\mathcal{I}_V,t] \times \mathcal{M}[\mathcal{I}_T(t)] \rightarrow \mathcal{A}[\mathcal{I}_T(t) \times |\mathcal{I}_V|, \mathbf{A} \leq \mathbf{B}_t] \}. \] (35)

**Remark 8:** In the case of unlimited sensing and unlimited communication ranges, the following relation hold
\[
\mathcal{I}_{\mathcal{T},j}(t) = \mathcal{I}_{\mathcal{T},j}(t) \cup \{ \emptyset_j \}, \quad \forall j \in \mathcal{I}_V \] (36a)
\[
\mathcal{I}_{\mathcal{V},i}(t) = \mathcal{I}_V, \quad \forall i \in \mathcal{I}_T \] (36b)
\[
\mathcal{I}_{\mathcal{V},j}(t) = \mathcal{I}_V, \quad \forall j \in \mathcal{I}_V \] (36c)
\[
\mathcal{I}_{\mathcal{T}}(t) = \mathcal{I}_{\mathcal{T}}(t) \cup \{ \emptyset_j : j \in \mathcal{I}_V \}. \] (36d)

However, in the case of unlimited sensing and communication ranges, each vehicle is allowed to choose a virtual target which is equivalent to not choosing any real target. It can be shown that by a proper choice of initial rewards and discount factors, this will not happen.

**Remark 9:** One can set all discount factors to zero and choose some arbitrary positive real numbers as initial rewards for virtual targets. Note that, the reward of any virtual is not discounted during the mission, and thus, a vehicle has no incentive to choose a virtual target unless it does not have any other option.

**C. Cooperative Receding Horizon Trajectory Construction**

It is desired now to develop a cooperative receding horizon controller (CRHC), which iteratively generates a set of headings, step sizes and optimal assignments for each vehicle such that the final collected rewards are maximized. The controller is applied at time instants denoted by \( \{ t_k \}_{k=0}^{\infty} \in \mathbb{R}_{\geq 0} \), where an optimization problem, estimating the collectible rewards in the future, is solved at each time instant. The solution of the optimization problem formulation is based on the currently available information, which are the current positions of the targets and vehicles, along with the predicted future positions of the targets. The solution of the optimization problem gives the optimal control input \( \mathbf{u}_k = (u_j(t_k))_{j \in \mathcal{I}_V} \) as well as the optimal assignments \( \{ a_{ij}(x(t_k+1), y(t_k+1)) \}_{i \in \mathcal{I}_T(t_k), j \in \mathcal{I}_V} \).

Let the action horizon of CRHC be denoted by \( \mathcal{H}_k \). For any \( j \in \mathcal{I}_V \), apply the control input \( u_j(t_k) \) to the vehicle \( j \), in the time interval \( (t_k, t_k+H_k) \). Then, it follows from equation (4) that at time \( t_k + H_k \), the position of vehicle \( j \) is
\[
x_j(t_k + H_k) = x_j(t_k) + u_j(t_k)H_k, \quad j \in \mathcal{I}_V. \] (37)

Similarly, based on the available information at time instant \( t_k \) and using equation (15), for any \( i \in \mathcal{I}_T(t_k) \), the position of target \( i \), can be estimated at time instant \( t_k + H_k \) as
\[
y_i(t_k + H_k) = y_i(t_k) + H_k \frac{d}{dt} y_i(t_k). \] (38)

**Remark 10:** It can be concluded from Remark 4 that the estimation error is bounded by \( \frac{1}{2} H_k^2 \alpha_{\max} \). Thus, for acceptable estimation accuracy, it is desired that \( \frac{1}{2} H_k^2 \alpha_{\max} \).
be much smaller than $H_k \hat{v}_{\text{max}}$, i.e. $H_k \ll 2\hat{v}_{\text{max}} \hat{\alpha}_{\text{max}}^{-1}$. Intuitively, this means that $H_k$ should be small enough to have a reliable estimation.

Denote by $\tau_{\text{min}, k}$ the earliest time that the next immediate target can be visited, using the estimates obtained based on the information available at time instant $t_k$, i.e.

$$\tau_{\text{min}, k} = t_k + \min_{i \in I_T(t_k), j \in I_V} \left\| x_j(t_k) - y_i(t_k) \right\| u_{\text{max}} + \hat{v}_{\text{max}}. \tag{39}$$

Since the CRHC path planning continues until either the next immediate target is visited, or a new target is arrived and then updates the information, the following relation holds

$$H_k \leq \eta_k (\tau_{\text{min}, k} - t_k) = \eta_k \min_{i \in I_T(t_k), j \in I_V} \left\| x_j(t_k) - y_i(t_k) \right\| u_{\text{max}} + \hat{v}_{\text{max}}, \tag{40}$$

where $\eta_k \in (0, 1)$ is a coefficient reflecting uncertainties.

**Remark 11**: For the implementation of CRHC, the length of the action horizon is chosen between $H_{k, \text{min}}$ and $H_{k, \text{max}}$, where $H_{k, \text{min}}$ depends on the physical characteristic of the problem, while $H_{k, \text{max}}$ is related to some upper bounds like the one given in (40). This bound depends on all the mutual distances between the vehicles and targets at any point in time. Therefore, its calculation requires a great deal of communication and computations, resulting in significant delays. Alternatively, one can use $\eta_k \Delta$ as an upper bound for $H_k$, where

$$\Delta = (u_{\text{max}} + \hat{v}_{\text{max}})^{-1} \min_{i \in I_T, j \in I_V} d_{ij}. \tag{41}$$

Note that $\Delta$ can be computed efficiently. Note also that, this new upper bound is more conservative than the one in (40).

**Remark 12**: For simplicity, let the action horizon be chosen equal to the planning horizon. Therefore, $t_{k+1} = t_k + H_k$, which by substituting in equation (40) yields

$$\tau_{\text{min}, k} - t_{k+1} = (\tau_{\text{min}, k} - t_k) - H_k > 0. \tag{42}$$

The position of target $i$ at any time instant $t \geq \tau_{\text{min}, k}$ is

$$y_i(t) = y_i(t_k) + \frac{d}{dt} y_i(t_k)(t - t_k) + \frac{1}{2} \alpha_i(t, t_k)(t - t_k)^2, \tag{43}$$

for any $i \in I_T(t_k)$. Now, it follows from equation (38), that

$$y_i(t) = \dot{y}_i(t_k) + \frac{d}{dt} y_i(t_k)(t - t_k) + \frac{1}{2} \alpha_i(t, t_k)(t - t_k)^2 = \dot{y}_i(t_k) + \frac{1}{2} \hat{\alpha}_k(t - t_k)^2 \hat{\alpha}_{\text{max}}, \tag{44}$$

where $\hat{\alpha}_k$ is defined as

$$\hat{\alpha}_k = \alpha_i(t, t_k) + \frac{2(t - t_k + 1) \frac{d}{dt} y_i(t_k)}{(t - k)^2 \hat{\alpha}_{\text{max}}}. \tag{45}$$

Since $t - t_k \leq t - t_{k+1}$, it is concluded that

$$\left\| \hat{\alpha}_k \right\| \leq 1 + \frac{1}{2} \frac{d}{dt} y_i(t_k)(t - t_k) \hat{\alpha}_{\text{max}}. \tag{46}$$

The second term in the right hand side of equation (46) satisfies the following inequality

$$\left\| \frac{2 \frac{d}{dt} y_i(t_k)}{(t - t_{k+1}) \hat{\alpha}_{\text{max}}} \right\| \leq \frac{1}{2} \frac{d}{dt} y_i(t_k)(t - t_k) \hat{\alpha}_{\text{max}}. \tag{47}$$

Note that the denominator in the right side of (47) is sufficiently large if the targets and vehicles are very far from each other, or $\hat{\alpha}_{\text{max}}$ is sufficiently large. In that case, the right hand side of (47) will be negligible, and hence

$$\left\| \frac{2 \frac{d}{dt} y_i(t_k)}{(t - t_{k+1}) \hat{\alpha}_{\text{max}}} \right\| \ll 1. \tag{48}$$

Define $\hat{\alpha}_k = \hat{\alpha}_{\text{max}} \alpha_i(t, t_k)$ and note that the uncertainty in the trajectory of target $i$ is unknown, but uniformly distributed bounded by $\hat{\alpha}_{\text{max}}$. Choose $\hat{\alpha}_k$ as a uniformly distributed random vector, taking different directions and magnitudes between 0 and $\hat{\alpha}_{\text{max}}$, such that (48) and (47) yield $\mathbb{E}[\hat{\alpha}_k] = 0$. Using this, one can estimate $y_i$ for large values of $t$ as $\hat{y}_i(t_{k+1})$. From this estimation as well as the current positions of targets and vehicles, and also the control input $u_{ij}$, for any $j \in I_V$, the time that vehicle $j$ visits target $i$ can be estimated to be

$$\hat{\tau}_{ij}(u_{ij}, t_k) = (t_k + H_k) + \frac{\left\| y_i(t_{k+1}) - y_j(t_{k+1}) \right\|}{u_{\text{max}}}, \tag{49}$$

assuming that the control input $u_{ij}$ remains unchanged after the time instant $t_{k+1}$ until the vehicle reaches the position of target $y_i(t_{k+1})$. Note that CRHC updates the estimates (including the ones given above) in each iteration to reduce the estimation error.

Similarly, if $a_{ij} (x(t_{k+1}), \hat{y}(t_{k+1}))$ is the optimal assignment, regardless of uncertainties, one can expect that this assignment remains unchanged until vehicle $j$ visits target $i$. Therefore

$$a_{ij} (x(\hat{\tau}_{ij}(u_{ij}, t_k)), \hat{y}(\hat{\tau}_{ij}(u_{ij}, t_k))) = a_{ij} (x(t_{k+1}), \hat{y}(t_{k+1})). \tag{50}$$

Consequently, one can estimate the maximum total reward which the vehicles expects at time $t_{k+1}$ to collect by the end of the mission. Denote by $\mathcal{R}^{k+1}$ this estimated expected reward. For simplicity of notations, let $\bar{a}_{ij}(u_{ij}, t_k) = a_{ij}(\hat{\tau}_{ij}(u_{ij}, t_k))$ and $\bar{a}_{ij}(u_{ij}, t_k) = a_{ij}(x(\hat{\tau}_{ij}(u_{ij}, t_k)), \hat{y}(\hat{\tau}_{ij}(u_{ij}, t_k)))$. Then

$$\mathcal{R}^{k+1}(u_{ij}, t_k) = \sum_{j \in I_V} \sum_{i \in I_T} R_{ij}(u_{ij}, t_k) \tilde{a}_{ij}(u_{ij}, t_k). \tag{51}$$

Note that the target index set in (50) is time-dependent, as expected. From (60), the $k^{\text{th}}$ iteration in CRHC, $P_k$, can be written as

$$P_k : \quad \max \mathcal{R}^{k+1}(u_{ij}, t_k), \quad \text{s.t.} \quad \bar{A}(u_{ij}, t_k) \in \mathcal{A}^{k}, \quad u_{ij} \in U^{k}, \tag{52}$$

where $\mathcal{A}^{k} = \mathcal{A}(t_k, x(t_k), y(t_k))$ and $U^{k} = \{u = (u_{ij})_{i \in I_T, j \in I_V} : u_{ij} \in \mathbb{R}^d, \|u_{ij}\| \leq u_{\text{max}}, \forall j \in I_V \}$ is the set of admissible heading control.

For convenience of notation, $x_j(t_k), y_i(t_k)$ and $\hat{y}_i(t_k)$ will hereafter be denoted by $x_j^k, y_i^k$ and $\hat{y}_i^k$, respectively, for any
such that for any

Theorem 1: Consider the performance index $J_{\gamma}^{k+1} : \mathcal{M}^{[\mathcal{T}] \times \mathbb{R}^{0_+}}$ for any $i \in \mathcal{I}$.

\begin{equation}
J_{\gamma}^{k+1}(x, A) = \sum_{j \in \mathcal{I}^V} \sum_{t \in \mathcal{I}^T} a_{ij} e^{-\gamma \|x_j - y_i^{k+1}\|},
\end{equation}

where $\gamma = \gamma_{\text{max}}$. Then, if $\mathcal{I}^T(t_k) \neq \emptyset$, the optimization problem $\text{P}^k$ presented in (52) is equivalent to

\begin{equation}
\begin{aligned}
& \max \ J_{\gamma}^{k+1}(x, A), \\
& \text{s.t.} \ A \in \mathbb{R}^{[\mathcal{I}^T]} \times \mathbb{R}^{[\mathcal{I}^V]}, \\
& \|x_j - y_i^{k+1}\| \leq \gamma_{\text{max}} H_k, \quad \forall j \in \mathcal{I}^V.
\end{aligned}
\end{equation}

Proof: See Appendix VII-A.

For any $j \in \mathcal{I}^V$ and $i \in \mathcal{I}^T(t_k)$, denote by $a_{ij}$ the $j^\text{th}$ column of $A$ and by $A^\text{i}_i$ its $i^\text{th}$ row. Define the penalty function $p(A) = \max(0, A^\text{i}_i \mathbb{I}^{[\mathcal{I}^V]} - 1)$, for any $A \in \mathbb{R}^{[\mathcal{I}^V]}$. You can show that as $\lambda \to \infty$, the solution of the following maximization problem

\begin{equation}
\begin{aligned}
& \max \ J_{\gamma}^{k+1}(x, A) - \lambda \sum_{i \in \mathcal{I}^T(t_k)} p(A_i), \\
& \text{s.t.} \ A \in \mathbb{R}^{[\mathcal{I}^T]} \times \mathbb{R}^{[\mathcal{I}^V]}, \\
& \|x_j - y_i^{k+1}\| \leq \gamma_{\text{max}} H_k, \quad \forall j \in \mathcal{I}^V, \\
& a_{ij} \mathbb{1}_{\mathcal{I}^T} = 1, \quad \forall j \in \mathcal{I}^V, \\
& \lambda \geq 0 \quad \forall i \in \mathcal{I}^T.
\end{aligned}
\end{equation}

converges to the solution of (54). Now, for any $j \in \mathcal{I}^V$ and $i \in \mathcal{I}^T(t_k)$, define $x_j^{k+1}$ as

\begin{equation}
x_j^{k+1} = x_j^k + \frac{\gamma_{\text{max}} H_k}{\|x_j^k - y_i^k\|} u_{\text{max}}^k,
\end{equation}

which is, in fact the future position of vehicle $j$ as it is aimed to move towards predicted position of target $i$, and let $d_{ij}^k = e^{-\gamma \|x_j^k - y_i^k\|}$. For any $t_k$, define the finite game $\mathcal{G}_k = (\mathcal{I}^V \times \mathcal{I}^T(t_k), \mathcal{U}, U)$, where

\begin{equation}
U_j((a_{ij})_{j \in \mathcal{I}^V}) = \sum_{i \in \mathcal{I}^T(t_k), j} e^{-\gamma \|x_j^k - y_i^k\|} - \lambda p((a_{ij})_{j \in \mathcal{I}^V}).
\end{equation}

Note that the set of action profiles here is the same as the set of integer assignments, i.e., the assignments with values 0 or 1. Similarly, it can be verified that each assignment is a strategy profile for the game $\mathcal{G}_k$.

Theorem 2: The game $\mathcal{G}_k$ is a potential game with the following potential function

\begin{equation}
P((a_{ij})_{j \in \mathcal{I}^V}) = \sum_{j \in \mathcal{I}^V} \sum_{i \in \mathcal{I}^T(t_k)} a_{ij} e^{-\gamma \|x_j^k - y_i^k\|} - \lambda \sum_{i \in \mathcal{I}^T(t_k)} p((a_{ij})_{j \in \mathcal{I}^V}).
\end{equation}

Proof: See Appendix VII-B.

Theorem 3: There exists a constant $\lambda$ such that for any $\lambda \geq \lambda$, the optimization problem (55) has a solution $(x_j^\lambda, A_j^\lambda)$ with the entries of $A_j^\lambda$ being 0 or 1 which is a solution of (54). Furthermore, this solution, $A_j^\lambda$, is a maximizer for the potential function (58), and hence a pure Nash equilibrium for $\mathcal{G}_k$.

Proof: See Appendix VII-C.

Remark 13: Since the vehicles are capable of communicating with each other, they can share with their neighbors their actions on the targets located in the intersection of their sensing regions. Based on these information exchanges, various learning methods such as spatial adaptive play (SAP) [16] can be applied to obtain a Nash equilibrium. One of the most efficient methods for this purpose is generalized regret monitoring (GRM) [17], which guarantees sufficiently fast convergence to a pure Nash equilibrium.

Remark 14: Even though any set of action profiles obtained from a solution of (55) with 0-1 assignments is a pure Nash equilibrium for (58), the converse is not generally true. However, one can obtain a set of headings from a pure Nash equilibrium, where the vehicles have no incentive to change their headings. Further investigation is required to verify that even with these headings, vehicles can accomplish the tasks and visit the targets in finite time.

V. SIMULATION RESULTS

In this section, the performance of the proposed method with two different multi-agent dynamic learning approaches are investigated through two simulations, each involving two vehicles and a set of five targets arriving sequentially in the mission space. Throughout this section, the sensing ranges for both vehicles are $r = 5$ and the mission space is a closed convex set in a flat plane $\mathcal{M} = [-20, 20] \times [-20, 20]$. Targets have unknown trajectories with the maximum velocity of $v_{\text{max}} = 1.5 m/s$ and the upper bound on the magnitude of vehicles’ velocity is $u_{\text{max}} = 2 m/s$. For generality, the targets’ trajectories are chosen randomly. Initially, along with the two vehicles, two targets are also present in the mission space, and the remaining three targets arrive sequentially at $\{T_i\}_{i=1}^3 = \{3, 4, 6\}$.

Case 1 (SAP): In this learning method [16], vehicles negotiate with each other to reach the pure Nash equilibrium by computing a utility function, where intercepting a target is rewarded while selection of one target by more than one vehicle is penalized by a negative term with a sufficiently large magnitude. Figure 1 shows the result of this learning mechanism, where it can be observed that other than target 2 that has been out of the sensing region of the vehicles, all other targets are intercepted by a vehicle in a cooperative manner.

Case 2 (GRM): In this learning method [17], a fading memory and inertia mechanism is also utilized to enable fast convergence to pure Nash equilibrium, where the forgetting factor is set to $\rho = 0.99$ and the inertia is $\alpha = 0.95$. The utility function is defined exactly the same as the
SAP algorithm in first case. Figure 2 depicts the result of this simulation, which demonstrates that targets 2 and 5 remain out of the sensing region of the vehicles but all other the targets are intercepted by the vehicles. As a result of negotiation in the game played by two vehicles in this example they switch their selected target 3 and 4 at some point in time. It is worth noting that in both simulations, the vehicles stop moving when there are no target in their sensing region. This is a result of the extra virtual target that was added for each vehicle which can be selected when no target is sensible.

VI. CONCLUSIONS

In this paper, the problem of multi-vehicle cooperative interception of moving objects is investigated where arrival times, trajectories and dynamics are unknown. The vehicles are assumed to have limited sensing and communication ranges. Accordingly, centralized trends are not efficient, especially in the case of having large number of vehicles and targets. In this work, a method is proposed based on a game theoretic cooperative receding horizon approach. The controller design is based on the limited information such as current position and velocity vectors of targets, as well as a reward allocation strategy for accomplishing the target interception tasks. To learn the dynamics in the resulting potential game, state of the art methods such as spatial adaptive play and generalized regret monitoring are applied and their effectiveness is demonstrated.

VII. APPENDICES

A. Proof of Theorem 2

For any fixed $H_k \in \mathbb{R}_{>0}$, the equation (37) gives a one-to-one correspondence between $U^k$ and $\mathcal{B}(x^k, u_{\max} H_k)$. Considering (17) and (49), for any $i \in \mathcal{I}_T(t)$ and any $j \in \mathcal{I}_V$, one has

$$\delta_{ij}(u^k, t_k) = e^{-\gamma_i(t_k + H_k + \frac{||x^{k+1}_{i} - y^{k+1}_{j}||}{u_{\max}})} 
= e^{-\gamma_i(t_k + H_k)} e^{-\gamma_i u_{\max} ||x^{k+1}_{i} - y^{k+1}_{j}||}.$$  \hspace{1cm} (59)

Since for any $i \in \mathcal{I}_T$, it is assumed that $\gamma_i = \gamma$ and $\mathcal{R}_i = \mathcal{R}$, it yields

$$\mathcal{R}^{k+1}_i = \mathcal{R}_i e^{-\gamma_i(t_k + H_k)} \sum_{j \in \mathcal{I}_V \times \mathcal{I}_T(t_k)} \tilde{\delta}_{ij} e^{-\gamma_i ||x^{k+1}_{i} - y^{k+1}_{j}||},$$  \hspace{1cm} (60)

where the arguments $(u^k, t_k)$ are omitted for brevity. Hence, from (39), it concludes that

$$\mathcal{R}^{k+1}(u^k, t_k) = \mathcal{R}_i e^{-\gamma_i(t_k + H_k)} J_{\mathcal{I}_T(t_k)} (x^{k+1}, A^k).$$  \hspace{1cm} (61)

Therefore, as $\mathcal{R}_i e^{-\gamma_i(t_k + H_k)} > 0$, the optimization problem (52) is equivalent with

$$\max_{A^k} J_{\mathcal{I}_T(t_k)} (x^{k+1}, A^k),$$  \hspace{1cm} (62)

s.t. $A^k \in \mathcal{A}_{\mathcal{I}_T(t_k), \mathcal{I}_V},$ \hspace{1cm}

$||x^{k+1}_{i} - x^{k+1}_{j}|| \leq u_{\max} H_k,$ $\forall j \in \mathcal{I}_V.$

Changing names of the variables, the optimization problem (54) yields.

B. Proof of Theorem 2

Consider vehicle $j$. Let $j', j''$ be indices of two targets in $\mathcal{I}_{\mathcal{T}(t_k), j}$ and the standard vectors $e_{j'}, e_{j''} \in \mathbb{R}^{\mathcal{I}_{\mathcal{T}(t_k)}}$ be the respective action vectors, i.e. $a^j_{j'} = e_{j'}$ and $a^j_{j''} = e_{j''}$. Also, let $a_{j} \in \mathcal{A}_{\mathcal{T}(t_k), \mathcal{V}}$ represents actions for the vehicles with indices in $\mathcal{I}_{\mathcal{T}(\{j\})}$, i.e. $a_{j} = (a_{l})_{l \in \mathcal{I}_{\mathcal{V}} \setminus \{j\}}$ where $a_{l} = e_{l}$ is the action vector for vehicle $l$ and $i_{l} \in \mathcal{I}_{\mathcal{T}(t_k), l}$ is the target with respect to action vector $a_{l}$, for any $l \in \mathcal{I}_{\mathcal{V}} \setminus \{j\}$. Now, set $A' = (a'_{j'})$ as $(a'_{j'}, a_{j} - j')$ and $A'' = (a''_{j'})$ as $(a'_{j'}, a_{j} - j)$. In order to show that $\mathcal{S}_k$ is a potential game, one needs to verify that

$$P(a'_{j'}, a_{j} - j') - P(a''_{j'}, a_{j} - j) = U_j(a'_{j'}, a_{j} - j') - U_j(a''_{j'}, a_{j} - j).$$

First, let $j', j'' \notin \tilde{\mathcal{T}}$. From (58) one has

$$P(A') - P(A'') = \left( \sum_{l \in \mathcal{I}_{\mathcal{V}}} \sum_{i \in \mathcal{I}_{\mathcal{T}(t_k), l}, \tilde{\mathcal{T}}} (\alpha''_{i,l} p_{i,l} - \lambda \sum_{i \in \mathcal{I}_{\mathcal{T}(t_k), l}, \tilde{\mathcal{T}}} p_{i,l}) \right) - \left( \sum_{l \in \mathcal{I}_{\mathcal{V}}} \sum_{i \in \mathcal{I}_{\mathcal{T}(t_k), l}, \tilde{\mathcal{T}}} (\alpha''_{i,l} p_{i,l} - \lambda \sum_{i \in \mathcal{I}_{\mathcal{T}(t_k), l}, \tilde{\mathcal{T}}} p_{i,l}) \right).$$
Since, $A'$ and $A''$ differ only in $j^n$ column and in rows $i', i'' \in I_{T(t_k), j}$, it yields that

$$P(A') - P(A'') = d_{ij}^k - d_{ij}^l - \lambda \left( p((a'_{i1})_{le I_V}) - p((a''_{i1})_{le I_V}) \right).$$

Similarly, from (67) one has

$$U_j(A') - U_j(A'') = \left( \sum_{i \in I_{T(t_k), j}} a_{ij}J^k_i - \lambda \sum_{i \in I_{T(t_k), j}} p((a'_{i1})_{le I_V}) \right) - \left( \sum_{i \in I_{T(t_k), j}} a_{ij}J^k_i - \lambda \sum_{i \in I_{T(t_k), j}} p((a''_{i1})_{le I_V}) \right),$$

and also, as $A'$ and $A''$ differ only in $j^n$ column and in rows $i', i'' \in I_{T(t_k), j}$, one can see that

$$U_j(A') - U_j(A'') = \left( \sum_{i \in I_{T(t_k), j}} a_{ij}J^k_i - \lambda \sum_{i \in I_{T(t_k), j}} p((a'_{i1})_{le I_V}) \right) - \left( \sum_{i \in I_{T(t_k), j}} a_{ij}J^k_i - \lambda \sum_{i \in I_{T(t_k), j}} p((a''_{i1})_{le I_V}) \right),$$

it concludes that

$$P(a', a_{-j}) - P(a''_{i}, a_{-j}) = U_j(a', a_{-j}) - U_j(a''_{i}, a_{-j}).$$

In the case that $i' \in I_T$ or $i'' \in I_T$, with a similar discussion one can show that (66) holds. This proves that the game $\mathcal{G}$ is a potential game.  

\textbf{C. Proof of Theorem 4.4}

1) Preliminary Definitions and Theorems: Let $n, m \in \mathbb{N}$ and $\mathcal{G} = (\mathcal{U} \cup \mathcal{V}, \mathcal{E})$ be a bipartite graph with vertex partitions $\mathcal{U}$ and $\mathcal{V}$ where $|\mathcal{U}| = n$ and $|\mathcal{V}| = m$. Also, let $B_\mathcal{G}$ be the biadjacency matrix of the bipartite graph $\mathcal{G}$. Define

$$A_\mathcal{G} = \{A \in [0,1]^{n \times m} : A \preceq B_\mathcal{G}, A^T 1_n = 1_m\},$$

and

$$B_\mathcal{G} = \bigcap_{A \in A_\mathcal{G}} A \cap \{0,1\}^{n \times m}.$$  

Similarly, define

$$\tilde{A}_\mathcal{G} = \{A \in [0,1]^{n \times m} : A \preceq B_\mathcal{G}, A^T 1_n = 1_m\},$$

and

$$\tilde{B}_\mathcal{G} = \tilde{A}_\mathcal{G} \cap \{0,1\}^{n \times m}.$$  

Now, let $I$ be a set of natural numbers such that $I \subseteq \mathbb{N}$ and define

$$\mathcal{A}_{\mathcal{G}, I} = \{A \in [0,1]^{n \times m} : A \preceq B_\mathcal{G}, A^T 1_n = 1_m, A = (A_i)_{i \in I} \in \mathbb{N}, \forall i \in I, A_i^T 1_m \geq 1, \forall i \in I, A_i^T 1_m \leq 1\},$$

and

$$\mathcal{B}_{\mathcal{G}, I} = \tilde{A}_{\mathcal{G}, I} \cap \{0,1\}^{n \times m}.$$  

\textbf{Theorem 4:} For any bipartite graph $\mathcal{G}$, one has that

i) $A_{\mathcal{G}} = \text{conv } B_\mathcal{G}$,

ii) $\tilde{A}_{\mathcal{G}} = \text{conv } \tilde{B}_\mathcal{G}$,

iii) $A_{\mathcal{G}, I} = \text{conv } B_{\mathcal{G}, I}$.

2) Proof of Theorem 4.4 Define function $\bar{J}_{\mathcal{G}, k+1}^* : M^{[I']_V} \times [0,1]^{\mathbb{I}_T(t_k) \times |\mathcal{V}|} \rightarrow \mathbb{R} \geq 0$ as

$$\bar{J}_{\mathcal{G}, k+1}^* (x, A) = \bar{J}_{\mathcal{G}, k+1}^* (x, A) - \lambda \sum_{i \in \mathbb{I}_T(t_k) \setminus I_T} p(A_i),$$

and note that

$$J_{\mathcal{G}, k+1}^* (x, A) = \sum_{i \in \mathbb{I}_T(t_k) \setminus I_T} \sum_{j \in \mathcal{V}} a_{ij} e^{-\gamma\|x_j - y^+_{k+1}\|}.\]$$

Denote by $\Omega_{t_k}$ as $\mathbb{T}_V \times \mathcal{G}(\mathcal{X}^T, H_{k, u_{\text{max}}})$. Take $(x^*, A^*)$ such that

$$(x^*, A^*) \in \arg\max_{(x, A) \in \Omega_{t_k}} \bar{J}_{\mathcal{G}, k+1}^* (x, A).$$

Define function $\bar{J}_{\mathcal{G}, k+1}^* : M^{[I']_V} \times [0,1]^{\mathbb{I}_T(t_k) \times |\mathcal{V}|} \rightarrow \mathbb{R} \geq 0$ as

$$\bar{J}_{\mathcal{G}, k+1}^* (x, A) = \bar{J}_{\mathcal{G}, k+1}^* (x, A) - \lambda \sum_{i \in \mathbb{I}_T(t_k) \setminus I_T} p(A_i),$$

where $\mathcal{A} = (A_i)_{i \in \mathbb{I}_T(t_k)}$. For any $(x, A) \in \Omega_{t_k} \times \mathcal{A}_{\mathcal{G}, I}$, one can simply verify that $\bar{J}_{\mathcal{G}, k+1}^* (x, A) = \bar{J}_{\mathcal{G}, k+1}^* (x, A)$. Therefore,

$$\arg\max_{(x, A) \in \Omega_{t_k} \times \mathcal{A}_{\mathcal{G}, I}} \bar{J}_{\mathcal{G}, k+1}^* (x, A) = \arg\max_{(x, A) \in \Omega_{t_k} \times \mathcal{A}_{\mathcal{G}, I}} \bar{J}_{\mathcal{G}, k+1}^* (x, A).$$

Since $\bar{J}_{\mathcal{G}, k+1}^*$ depends linearly on $A$ and $\mathcal{A}_{\mathcal{G}, I}$ is a polytope with etreme points belonging to $\mathcal{B}_{\mathcal{G}, I}$, there exists $A^* \in \mathcal{B}_{\mathcal{G}, I}$ such that

$$A^* \in \arg\max_{(x, A) \in \Omega_{t_k} \times \mathcal{A}_{\mathcal{G}, I}} \bar{J}_{\mathcal{G}, k+1}^* (x^*, A).$$

and therefore

$$A^* \in \arg\max_{(x, A) \in \Omega_{t_k} \times \mathcal{A}_{\mathcal{G}, I}} \bar{J}_{\mathcal{G}, k+1}^* (x^*, A).$$

Note that for any $x^* \in \Omega_{t_k}$ such that

$$x^* \in \arg\max_{x \in \Omega_{t_k}} \bar{J}_{\mathcal{G}, k+1}^* (x, A^*),$$

one has

$$x^* \in \arg\max_{x \in \Omega_{t_k}} \bar{J}_{\mathcal{G}, k+1}^* (x, A^*).$$

From (71) and (81), it yields that

$$\bar{J}_{\mathcal{G}, k+1}^* (x^*, A^*) \leq \bar{J}_{\mathcal{G}, k+1}^* (x^*, A^*).$$

(82)
Since \((x^{**}, A^{**}) \in \Omega_{t_k} \times \tilde{A}_{G_{tk}}\), it concludes from (76) and (82) that \(\tilde{J}_{y_{i,k}^{t-1}}(x^{**}, A^{*}) = \tilde{J}_{y_{i,k}^{t-1}}(x^{**}, A^{**})\) and
\[(x^{**}, A^{**}) \in \arg \max_{(x,A) \in \Omega_{t_k} \times \tilde{A}_{G_{tk}}} \tilde{J}_{y_{i,k}^{t-1}}(x,A) . \tag{83}\]
This shows that \((x^{**}, A^{**})\) is a solution of (55) with the entries of \(A^{**}\) being 0 or 1.

Let \(A = (a_{ij})\) be a \([|I_T|] \times |I_V|\) matrix with the property that \(a_{ij}\) is 1 only if \(i\) is the index according to \(D_j\), for any \(i \in \mathcal{I}_T(t_k)\) and any \(j \in \mathcal{I}_k\). Then, one can see that \((x^{k}, A) \in \Omega_{t_k} \times \tilde{A}_{G_{tk}}\) and \(\tilde{J}_{y_{i,k}^{t-1}}(x^{k}, A) = 0\), and consequently, \(\tilde{J}_{y_{i,k}^{t-1}}(x^{**}, A^{**}) \geq 0\). Set \(\lambda \in \mathbb{R}_{>0}\) as \[1 + \sum_{j \in \mathcal{I}_V} \sum_{i \in \mathcal{I}_T(t_k) \cap \mathcal{I}_V} d_{ij} - \lambda \], and let \(\lambda \geq \lambda\). Then, one has \(A^{**}\) as \(\lambda\) row of \(A^{**}\), for any \(i \in \mathcal{I}_T(t_k)\). Since, if there exists \(i \in \mathcal{I}_T(t_k)\) such that \(A^{**}_i \geq 1\), then \(\lambda\) does not belong to the set \(\{0, 1\}\), it follows that \(A^{**}_i \geq 1\), and consequently, \(p(A^{**}) \geq 1\). From this, it yields that \(\tilde{J}_{y_{i,k}^{t-1}}(x^{**}, A^{**}) \leq \sum_{j \in \mathcal{I}_T(t_k) \cap \mathcal{I}_V} \sum_{i \in \mathcal{I}_T(t_k) \cap \mathcal{I}_V} a_{ij} \tilde{e}^{-\gamma||x_j - y_{i,j}^{t-1}||} - \lambda \)
\[\leq \sum_{j \in \mathcal{I}_T(t_k) \cap \mathcal{I}_V} \sum_{i \in \mathcal{I}_T(t_k) \cap \mathcal{I}_V} d_{ij} - \lambda \]
\[< 0 ,\]
which contradicts \(\tilde{J}_{y_{i,k}^{t-1}}(x^{**}, A^{**}) \geq 0\). Thus, for any \(i \in \mathcal{I}_T(t_k)\), one has \(A^{**}_{i} \geq 1\), i.e.
\[(x^{**}, A^{**}) \in \Omega_{t_k} \times \tilde{A}_{G_{tk}} . \tag{84}\]
From (83), (84) and \(\Omega_{t_k} \times \tilde{A}_{G_{tk}} \subseteq \Omega_{t_k} \times \tilde{A}_{G_{tk}}\), it yields that
\[(x^{**}, A^{**}) \in \arg \max_{(x,A) \in \Omega_{t_k} \times \tilde{A}_{G_{tk}}} \tilde{J}_{y_{i,k}^{t-1}}(x,A) . \tag{85}\]
This shows that \((x^{**}, A^{**})\) is a solution of (54).

With a similar discussion, one can show that if \(\lambda \geq \lambda\) then there exists \(A^{*} \in \tilde{B}_{G_{tk}}\) such that
\[A^{**} \in \arg \max_{A \in \tilde{B}_{G_{tk}}} P(A) . \tag{86}\]

It can be easily seen that for any \((x,A) \in \Omega_{t_k} \times \tilde{A}_{G_{tk}}\), one has
\[J_{y_{i,k}^{t-1}}(x,A) \leq P(A) , \tag{87}\]
and hence
\[J_{y_{i,k}^{t-1}}(x^{**}, A^{**}) \leq P(A^{**}) . \tag{88}\]
Now, consider the map \(x : \tilde{B}_{G_{tk}} \rightarrow \Omega_{t_k}\) such that \(x(A) = (x_j(A))_{j \in \mathcal{I}_V}\) and for any \(j \in \mathcal{I}_V\), \(x_j(A)\) is defined as
\[x_j(A) = \begin{cases} x_j^k + \frac{y_{i,j}^{t-1} - x_j^k}{\tilde{e}^{y_{i,j}^{t-1} - x_j^k}} u_{\text{max}} H_k , & \text{if } i_j \notin \tilde{I} \\ x_j^k , & \text{if } i_j \in \tilde{I} \end{cases} \tag{89}\]
where \(i_j \in \mathcal{I}_T(t_k)\) is the index of the row that \(j^{th}\) column of \(A\) is 1 in that row. Hence, for any \(A \in \tilde{B}_{G_{tk}}\), it yields that \(J_{y_{i,k}^{t-1}}(x(A), A) = P(A)\). Thus, according to the definition of \((x^{**}, A^{**})\), it can be noticed that
\[J_{y_{i,k}^{t-1}}(x^{**}, A^{**}) = J_{y_{i,k}^{t-1}}(x(A^{**}), A^{**}) = P(A^{**}) , \tag{90}\]
and subsequently,
\[P(A^{**}) = J_{y_{i,k}^{t-1}}(x(A^{**}), A^{**}) \leq J_{y_{i,k}^{t-1}}(x^{**}, A^{**}) = P(A^{**}) . \tag{91}\]
Considering equations (88), (90) and (91), it follows that \(P(A^{**}) = P(A^{**})\), i.e.
\[A^{**} \in \arg \max_{A \in \tilde{B}_{G_{tk}}} P(A) . \tag{92}\]
This shows that \(A^{**}\) is a maximizer for the potential function (58) and hence a pure Nash equilibrium for \(\mathcal{G}_k\).

REFERENCES