Hybrid Pursuit-Evasion Game between UAVs and RF Emitters with Controllable Observations: A Hawk-Dove Game

Husheng Li[1], Vasu Chakravarthy[2], Sintayehu Dehnie[3], Deborah Walter[4], and Zhiqiang Wu[5]

[1] The University of Tennessee, Knoxville, TN
husheng@eecs.utk.edu
[2] Air Force Research Lab, Dayton, OH
Vasu.Chakravarthy@wpafb.af.mil
[3] Booz Allen Hamilton, Dayton, OH
Sintayehu.Dehnie.ctr@wpafb.af.mil
[4] Rose-Hulman Institute of Technology
Deborah.Walter@rose-hulman.edu
[5] Wright State University, Dayton, OH
zhiqiang.wu@wright.edu

Abstract. Unmanned aerial vehicles (UAVs) can be used to chase radio frequency (RF) emitters by sensing the signal sent out by the RF emitters. Meanwhile, the RF emitter can evade from the UAVs, thus forming a pursuit-evasion game. In contrast to traditional pursuit-evasion games, in which the players can always observe each other, the RF emitter can stop transmitting such that the UAVs lose the target. However, stopping the transmission also incurs cost to the RF emitter since it can no longer convey information to destinations. Hence, the RF emitter can take both continuous actions, i.e., the moving direction, and discrete actions, i.e., whether to stop transmission. Meanwhile, there are both discrete states, i.e., whether the RF transmitter is transmitting, and continuous states, i.e., the locations of UAVs and RF emitter, thus forming a hybrid system. We will study the game theoretic properties of this novel game and derive the optimal strategies for both parties under certain assumptions.

Keywords: UAV, pursuit-evasion game.

1 Introduction

Unmanned aerial vehicle (UAV) is a remotely piloted aircraft, which is widely used in military. It can be used for many tasks, particularly in surveillance or renaissance. In recent years, people have studied how to use UAVs as a flying sensor network to monitor various activities, such as radio activities [2][4][5][9][11]. This is particularly useful in military due to the inexpensive cost and efficient deployment.

In this paper, we study how UAVs can be used to chase RF emitters. When a UAV is equipped with directional antenna, it can determine where the RF

V. Krishnamurthy et al. (Eds.): GameNets 2012, LNICST 105, pp. 99–114, 2012.

emitter is and then pursue it, either to continue the surveillance or destroy the RF emitter. We assume that the RF emitter is also mobile, but with a slower speed than the UAV. The RF emitter can move to evade the pursuit of the UAV. Then, it forms a pursuit-evasion game which was originally studied by R. Isaacs [6]. Since such a game is played in a continuous space and continuous time, it belongs to the category of differential games. In contrast to traditional game theory, in which randomness is a key factor of the game, the pursuit-evasion game is deterministic, which can be described by a partial differential equation (called Isaacs Equation). It has been widely applied in the study of warfares, such as the doggame of gameers and the bunker hill battle [6]. The value functions and the optimal strategies at the equilibrium have been obtained for many applications.

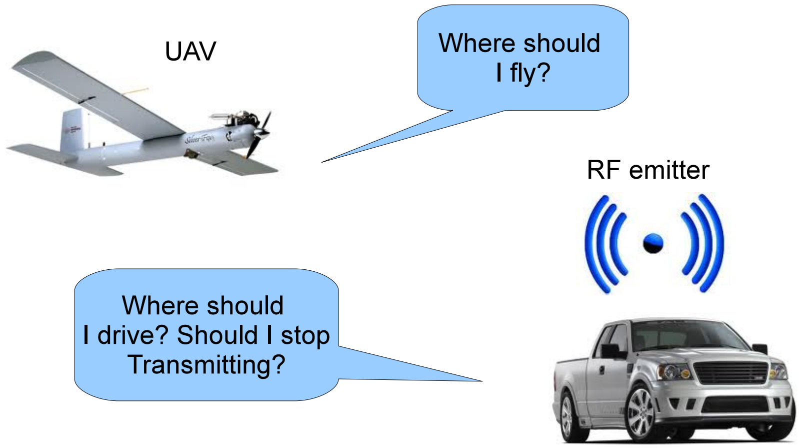

Fig. 1. An illustration of the pursuit-evasion game of UAV and RF emitter

In contrast to the traditional pursuit-evasion games, the game studied in this paper is characterized by its hybrid action space and state space. Besides choosing the moving direction, the RF emitter can choose to stop the RF transmission. Since the UAV's geolocationing capability is completely dependent on the RF signal, the UAV will lose the observability of the RF emitter. During this 'blind' period, the RF emitter can try to evade the UAV. However, it incurs penalty to the RF emitter when it ceases the RF transmission. Hence, the RF emitter must find a good tradeoff between the risk of being caught by the UAV and the penalty of ceasing transmitting. Note that the action of whether to transmit and the state of whether being transmitting are both discrete. Therefore, the game is actually played in a hybrid system in which both discrete and continuous states exist [7].

Note that hybrid systems have been intensively studied in recent years due to its wide applications in various areas such as smart grid and robotic network. However, there have been very few studies on the games in hybrid systems [8].

In particular, there have been no studies on the pursuit-evasion games with the observation controllable by the evader, to our best knowledge. In this paper, we will consider both cases of discounted and non-discounted rewards. The feedback Nash equilibrium will be obtained and described by a combination of Bellman's equation and Isaacs equation. Due to the prohibitive challenge of solving the equations, we will study heuristic steering strategies of the UAV and RF emitter and then use numerical simulations to explore the strategy of whether to stop transmitting.

The remainder of this paper is organized as follows. The system model for the UAV and RF emitter is introduced in Section 2. The case of single UAV and single RF emitter is studied in Section 3 and is then extended to the multiple-UAV-multiple-emitter case in Section 4. The numerical results and conclusions are obtained in Sections 5 and 6, respectively.

2 System Model

Consider one UAV and one RF emitter. We denote by $\mathbf{x}_u = (x_{u1}, x_{u2})$ and $\mathbf{x}_e = (x_{e1}, x_{e2})$ the locations of the UAV and the RF emitter. We adopt a simple model for the motions of the UAV and RF emitter, using the following ordinary differential equations [3]:

$$\begin{cases} \dot{x}_{u1} = v_u \sin\theta_u \\ \dot{x}_{u2} = v_u \cos\theta_u \\ \dot{\theta}_u \;\;= w_u f_u \end{cases}, \tag{1}$$

and

$$\begin{cases} \dot{x}_{e1} = v_e \sin\theta_e \\ \dot{x}_{e2} = v_e \cos\theta_e \\ \dot{\theta}_e \;\;= w_e f_e \end{cases}, \tag{2}$$

where v_u and v_e are the velocities; f_u and f_e are the forces to make the direction change; w_u and w_e are the inertia. It is reasonable to assume that $v_u > v_e$. We assume that the forces are limited; i.e., $|f_u| < F_u$ and $|f_e| < F_e$, where F_u and F_e are the maximum absolute values of the forces. Note that the above model is very simple but more mathematically tractable than more complicated motion models.

3 Single-UAV-Single-Emitter Game

In this section, we assume that there is only one UAV and it can perfectly determine the location of the RF emitter when the emitter keeps transmitting. This is reasonable if the UAV employs a powerful sensor which can determine both distance (e.g., using the signal strength) and the angle (e.g., using an antenna array). However, when the emitter stops transmitting at a certain cost, the UAV loses the target; hence we say that the observation is controllable (by the emitter). In contrast to traditional continuous time pursuit-evasion game, the challenge of this game is the hybrid system state, which consists of both continuous one (the locations and directions) and discrete one (the emitter's transmission state).

3.1 Game Formulation

Obviously, there are two players in the game, namely the UAV and the RF emitter. Essentially, the UAV wants to pursue the RF using its sensor while the emitter wants to evade by moving or stopping emitting. For simplicity, we assume that the pursuit and evasion occur in a plane. The elements are itemized as follows.

State. We denote by $\mathbf{s}$ the state of the whole system, which consists of the following components:

- For the UAV side, its state includes its current location $\mathbf{x}_u = (x_{u1}, x_{u2})$ and the direction θ_u.
- For the emitter side, its state includes the current location $\mathbf{x}_e = (x_{e1}, x_{e2})$, the moving direction θ_e and its transmission state s_e: $s_e = 1$ when the the emitter transmits and $s_e = 0$ otherwise.

Since the game only concerns the the relative location $\mathbf{x} = \mathbf{x}_u - \mathbf{x}_e$, we can define the system state as $\mathbf{s} = (\mathbf{x}, \theta_u, \theta_e)$.

Actions. Both the UAV and emitter can move and change direction. Moreover, the emitter can choose to stop transmitting and then make the UAV lose track of the target. Hence, the actions in the game are defined as follows.

- UAV: The action is f_u which is visible to the emitter.
- Emitter: Its action includes f_e, which is also visible to the UAV when $s_e = 1$, and the decision on whether to stop the transmission, which is denoted by a_e.

For simplicity, we assume that, when the UAV loses the targets, it follows a certain predetermined track; e.g., keeping the original direction ($f_u = 0$). Moreover, we assume that the transmission state has a minimal dwelling time τ_0; i.e., each transmission state, namely on or off, must last for at least τ_0 units. To simplify the analysis, we assume that the decision on transmission can be made at only discrete times, namely 0, τ_0, $2\tau_0$, ... For the case in which the decision can be made at continuous time under the constraint of minimum dwelling time, the analysis is much more complicated and will be left to our future study.

Rewards. The purpose of the UAV is to catch the emitter or force the emitter to keep silent. When the distance between UAV and emitter is small, the game is ended. This stopping time is defined as

$$T^* = \inf \left\{ t | \|\mathbf{x}(t)\| \leq \gamma_d \right\}, \tag{3}$$

where γ_d is a predetermined threshold for the distance. It is possible that T^* is infinite if the UAV is unable to catch the emitter, e.g., if the RF emitter keeps silent forever.

Hence, the total reward of the UAV can be modeled using the non-discounted one or discounted one:

- Discounted reward: When the reward is discounted (i.e., the future is less important than now; hence, the requirement of time is incorporated into the decisions), we have (both $\alpha > 0$ and $0 < \beta < 1$ are parameters of discounting)

$$R = \int_{t=0}^{T^*} \left(R_0 e^{-\alpha t} \delta(\|\mathbf{x}(t)\| \leq \gamma_d) - c\beta^n a_e(t)\delta(t = n\tau_0) \right) dt, \tag{4}$$

  where R_0 is the reward for locating the emitter and c is the penalty on the UAV when the emitter transmits in one time slot. The reward at time t is given by

$$r(t) = R_0 e^{-\alpha t} \delta(\|\mathbf{x}(t)\| \leq \gamma_d) - c\beta^n a_e(t)\delta(t = n\tau_0). \tag{5}$$

- Non-discounted reward: When the reward is not discounted (i.e., the future is the same important as now) within a time window $[0, T_f]$, we have

$$R = \int_{t=0}^{\min\{T^*, T_f\}} \left(R_0 \delta(\|\mathbf{x}(t)\| \leq \gamma_d) - c a_e(t)\delta(t = n\tau_0) \right) dt. \tag{6}$$

  The reward at time t is given by

$$r(t) = R_0 \delta(\|\mathbf{x}(t)\| \leq \gamma_d) - c a_e(t)\delta(t = n\tau_0). \tag{7}$$

  For simplicity, we assume that $t_f = T_f/\tau_0$ is an integer.

Since we model the game as a zero-sum one, the reward of the emitter at time slot t is simply given by $-r(t)$. Note that, in practice, the reward could be more complicated, e.g., taking the fuel consumptions into account. This requires much more complicated models and will be studied in the future.

System Dynamics. The dynamics of the game can be written as

$$\dot{\mathbf{s}}(t) = f_{a_e(t)}(\mathbf{s}(t), f_u(t), f_e(t)), \tag{8}$$

where $a_e(t)$ is the transmission state of the emitter. We denote by π_u and π_e the strategies of the continuous actions of the UAV and emitter, respectively; i.e., $f_u(t) = \pi_u(a_e(t), \mathbf{s}(t))$ and $f_e(t) = \pi_e(a_e(t), \mathbf{x}(t))$. As we have assumed in the game formulation, when $a_e(t) = 0$ (the emitter stops transmitting), π_u is independent of $\mathbf{s}(t)$; i.e., the UAV follows a predetermined track. In this paper, we assume $\pi_u = 0$ when $a_e = 0$; i.e., the UAV keeps the original direction when it loses track of UAV.

3.2 Feedback Nash Equilibrium

When $a_e(t)$ is always 1 (i.e., the emitter keeps transmitting all the time and is thus always visible to the UAV), the game degenerates to a traditional pursuit-evasion game. A brief introduction to the feedback Nash equilibrium[1] of the traditional pursuit-evasion game is provided in the Appendix for self-containedness. When $a_e(t)$ is not always 1, the challenge is that there are both discrete and continuous system states in the dynamics, thus eliminating the possibility of straightforwardly applying the traditional theories of stochastic games (discrete system state) and differential games (continuous system state).

Equilibrium for Discounted Reward. First, we define reward-to-go function $R_s(t)$; i.e., the reward from time t to the game termination. We have the following two observations:

- We notice that the decision actually depends on only the relative locations directions of UAV and emitter, not on the current transmission status of emitter.
- There are two types of reward-to-go functions; namely the ones at the times of deciding the transmission status and the ones in other times. We assume that the decision on whether to shut down the transmission is made at time slightly before $n\tau_0$; i.e., $(n\tau_0)^-$. Then, we have reward-to-go functions $\{R_{\mathbf{s}}((n\tau_0)^-)\}_{n=0,1,\ldots}$ and $R_{\mathbf{x}}(t)$, $t \neq n\tau_0$.

Then, the following proposition provides the reward-to-go functions at the feedback Nash equilibrium of the game with non-discounted reward:

Proposition 1. *The reward-to-go functions for the non-discounted reward are determined by*

$$R_{\mathbf{s}}((\tau_0)^-) = \min_{a_e}\left[-cI(a_e = 1) + R_{\mathbf{s}'}(0, a_e)\right], \tag{9}$$

and

$$R_{\mathbf{s}}(t, 1) = \max_{f_u}\min_{f_e}\left[\int_t^{\min(\tau, T^*)} R_0\delta(\|\mathbf{x}(t)\| < \gamma_d)dt + R_{\mathbf{s}'}(\tau_0^-)\right], \tag{10}$$

where $\mathbf{s}'$ *is the system state at time* τ_0*, respectively, and*

$$R_{\mathbf{s}}(t, 0) = \min_{f_e}\left[\int_t^{\min(\tau, T^*)} R_0\delta(\|\mathbf{x}(t)\| < \gamma_d)dt + R_{\mathbf{s}'}(\tau_0^-)\right]\Bigg|_{f_u=0}. \tag{11}$$

And (31) and (33) can be further written as

$$\begin{cases} -\frac{\partial R_{\mathbf{s}}(t,1)}{\partial t} &= \max_{\mathbf{f}_u}\min_{\mathbf{f}_e}\left[\frac{\partial R_{\mathbf{s}}(t,1)}{\partial \mathbf{s}} f(t, \mathbf{s}, \mathbf{f}_u, \mathbf{f}_e) + R_0\delta(\|\mathbf{x}(t)\| < \gamma_d)\right], \\ R_{\mathbf{s}}(\tau, 1) &= R_{\mathbf{s}}((\tau_0)^-), \end{cases} \tag{12}$$

[1] The definition of feedback Nash equilibrium can be found in [1].

and

$$\begin{cases} -\frac{\partial R_{\mathbf{s}}(t,0)}{\partial t} &= \min_{\mathbf{f}_e} \left[\frac{\partial R_{\mathbf{s}}(t,1)}{\partial \mathbf{s}} f(t,\mathbf{s},\mathbf{f}_u,\mathbf{f}_e) + R_0\delta(\|\mathbf{x}(t)\| < \gamma_d)\right], \\ R_{\mathbf{s}}(\tau,0) &= R_{\mathbf{s}}((\tau_0)^-), \end{cases} \tag{13}$$

Then, we can obtain the optimal strategies of the UAV and emitter, which are given in the following corollary

Corollary 1. *The strategies at the feedback Nash equilibrium are given by*

– *The strategy of UAV is given by*

$$\mathbf{u}_f^* = \arg\max_{\mathbf{f}_u}\min_{\mathbf{f}_e}\left[\frac{\partial R_{\mathbf{s}}(t,1)}{\partial \mathbf{s}} f(t,\mathbf{s},\mathbf{f}_u,\mathbf{f}_e) + R_0\delta(\|\mathbf{x}(t)\| < \gamma_d)\right]. \tag{14}$$

– *The strategy of the emitter is given by*

$$\mathbf{u}_e^* = \arg\min_{\mathbf{f}_e}\min_{\mathbf{f}_u}\left[\frac{\partial R_{\mathbf{s}}(t,1)}{\partial \mathbf{s}} f(t,\mathbf{s},\mathbf{f}_u,\mathbf{f}_e) + R_0\delta(\|\mathbf{x}(t)\| < \gamma_d)\right]. \tag{15}$$

and

$$R_{\mathbf{x}}((\tau_0)^-) = \min_{a_e}\left[-cI(a_e=1) + R_{\mathbf{x}}(0,a_e)\right], \tag{16}$$

Equilibrium for Non-discounted Reward. Similarly to the discounted reward case, the equilibrium for the non-discounted reward case is given in the following proposition:

Proposition 2. *The reward-to-go functions for the non-discounted reward are determined by*

$$R_{\mathbf{s}}^n((\tau_0)^-) = \min_{a_e}\left[-cI(a_e=1) + R_{\mathbf{s}}^{n+1}(0,a_e)\right], \tag{17}$$

and

$$R_{\mathbf{s}}^{n+1}(t,1) = \max_{f_u}\min_{f_e}\left[\int_t^{\min(\tau,T^*)} R_0\delta(\|\mathbf{x}(t)\| < \gamma_d)dt + R_{\mathbf{s}'}^n(\tau_0^-)\right], \tag{18}$$

where $\mathbf{s}'$ *is the state at time* τ *and*

$$R_{\mathbf{s}}^{n+1}(t,0) = \min_{f_e}\left[\int_t^{\min(\tau,T^*)} R_0\delta(\|\mathbf{x}(t)\| < \gamma_d)dt + R_{\mathbf{s}'}^n(\tau_0^-)\right]\Bigg|_{f_u=0}. \tag{19}$$

And (31) and (33) can be further written as

$$\begin{cases} -\frac{\partial R_{\mathbf{s}}^{n+1}(t,1)}{\partial t} &= \max_{\mathbf{f}_u}\min_{\mathbf{f}_e} \left[\frac{\partial R_{\mathbf{s}}^{n+1}(t,1)}{\partial \mathbf{s}} f(t,\mathbf{s},\mathbf{f}_u,\mathbf{f}_e) + R_0\delta(\|\mathbf{x}(t)\| < \gamma_d)\right], \\ R_{\mathbf{s}}^{n+1}(\tau,1) &= R_{\mathbf{s}}^{n+1}((\tau_0)^-), \end{cases} \tag{20}$$

and

$$\begin{cases} -\frac{\partial R_{\mathbf{x}}^{n+1}(t,0)}{\partial t} &= \min_{\mathbf{f}_e} \left[\frac{\partial R_{\mathbf{x}}^{n+1}(s,1)}{\partial \mathbf{s}} f(t,\mathbf{s},\mathbf{f}_u,\mathbf{f}_e) + R_0\delta(\|\mathbf{x}(t)\| < \gamma_d)\right], \\ R_{\mathbf{s}}^{n+1}(\tau,0) &= R_{\mathbf{s}}^{n+1}((\tau_0)^-), \end{cases}, \tag{21}$$

and

$$R_{\mathbf{s}}^{t_f}((\tau_0)^- = 0. \tag{22}$$

3.3 Computation of Strategy

Since we have both continuous and discrete actions, we address them separately and then integrate into one uniform procedure for computing the strategies at the feedback Nash equilibrium. For simplicity, we consider only the case of discounted rewards.

Discrete Action. For the discrete action, we consider only the emitter since there is no discrete action for the UAV.

- Case of Discounted Reward: We assume that, given $R_{\mathbf{s}}((\tau_0)^-)$, we know how to compute the strategies of the UAV and emitter in (20) and (21). Then, we can do the following value iteration for computing $R_{\mathbf{s}}((\tau_0)^-)$:

$$\begin{cases} R_{\mathbf{s}}^{k+1}((\tau_0)^-) = \min_{a_e} \left[-cI(a_e = 1) + R_{\mathbf{s}}^k(0, a_e)\right] \\ R_{\mathbf{s}}^0((\tau_0)^-) \quad = R_0(\mathbf{x}) \end{cases}, \tag{23}$$

  where R_0 is the initialization of the reward-to-go function, which is a function of the relative location, and $R_{\mathbf{s}}^k(0, a_e)$ is obtained from the the values of $R_{\mathbf{s}}^k((\tau_0)^-)$ in the k-th iteration. The difficulty of the value iteration is that $\mathbf{s}$ is a continuous state, thus requiring uncountable equations in the value iteration. One effective approach is that we can discretize the location, thus approximating the problem using a discrete one.
- Case of Non-discounted Reward: We assume that, given $R_{\mathbf{s}}((\tau_0)^-)$, we know how to compute the strategies of the UAV and emitter in (20) and (21). Then, we can do the following value iteration for computing $R_{\mathbf{s}}((\tau_0)^-)$:

$$\begin{cases} R_{\mathbf{s}}^{k+1}((\tau_0)^-) = \min_{a_e} \left[-cI(a_e = 1) + R_{\mathbf{s}}^k(0, a_e)\right] \\ R_{\mathbf{s}}^0((\tau_0)^-) \quad = R_0(\mathbf{s}) \end{cases}, \tag{24}$$

  where R_0 is the initialization of the reward-to-go function, which is a function of the relative location, and $R_{\mathbf{s}}^k(0, a_e)$ is obtained from the the values of $R_{\mathbf{s}}^k((\tau_0)^-)$ in the k-th iteration. The difficulty of the value iteration is that $\mathbf{x}$ is a continuous state, thus requiring uncountable equations in the value iteration. One effective approach is that we can discretize the location, thus approximating the problem using a discrete one.

Continuous Action. It is highly nontrivial to solve the partial differential equation, particularly when the cost function $R_{\mathbf{x}}^0((\tau_0)^-)$ is complicated. Unfortunately, we are still unable to solve it. Hence, we propose the following heuristic but reasonable strategy for both the UAV and the RF emitter, which is independent of whether the reward is discounted or not:

- UAV: When the RF emitter is transmitting, the UAV follows the direction towards the RF emitter using the full force.
- RF emitter: The RF emitter follows the direction perpendicular to the vector between the UAV and the RF emitter in full strength.

4 Multi-UAV-Multi-Emitter Game

In this section, we extend the study on the single-UAV-single-emitter game to the general case in which we consider multiple UAVs and multiple emitters.

4.1 Game Formulation

We assume that there are N_u UAVs and N_e RF emitters. We assume that both quantities N_u and N_e are known to all UAVs and emitters. This is reasonable since each emitter can count the number of UAVs due to the assumption of visibility. We also assume that the emitters are in the state of 'on' at the beginning such that the UAVs know the number of emitters. The elements of the game are then explained as follows.

Players: Since we do not consider any random factor, thus making the game a deterministic one, each UAV and each emitter know the future evolution of the game at the feedback Nash equilibrium. Hence, we can consider the the game as a two (virtual) player one; i.e., both the UAV side and the emitter side are controlled in centralized way. We assume that each emitter will be out of the game once it is caught by any UAV; e.g., it is destroyed by the UAV. Hence, the number of actual players may be changing during the game. We denote by $\mathcal{N}_e(t)$ the set of emitters still surviving at time t.

In practice, when there exists randomness in the observations or each UAV (emitter) has limited knowledge to the system state, the communications among the UAVs or the emitters need to be considered, which is concerned with the team formations due to limited communication range. This more complicated case will be studied in the future.

State Space. For each individual UAV and emitter, its state is the same as the single-UAV-single-emitter case. The system state space is the product of the individual ones; i.e., the state includes the locations and directions of all UAVs and emitters, denoted by $\{\mathbf{x}_n^u\}_{n=1,\ldots,N_u}$, $\{\theta_n^u\}_{n=1,\ldots,N_u}$, $\{\mathbf{x}_n^e\}_{n\in\mathcal{N}_e(t)}$. $\{\theta_n^e\}_{n=1,\ldots,N_e}$, as well as the emitters' transmission state. Note that, when an emitter is caught by a UAV, it is out of the game and the state space is reduced. Similarly to the single-UAV-single-emitter case, we still use $\mathbf{s}$ to denote the overall system state (excluding the discrete state of the transmission status of each emitter).

Action Space. For each individual UAV or emitter, its action space is the same as the single-UAV-single-emitter case in the previous section. We simply add superscript to distinguish the actions of different UAVs or emitters. For simplicity, we do not add more constraints like collision avoidance or formation maintenance.

Reward. Similarly to the single-UAV-single-emitter case, a reward is achieved by the UAVs when an emitter is caught. A cost is incurred to an emitter if it stops transmitting. Due to the limited space, we consider only the non-discounted case, in which the reward is given by

$$\begin{aligned} R = &\int_{t=0}^{T^*} e^{-\alpha t} R_0 \delta(\|\mathbf{x}_u^n - \mathbf{x}_e^m(t)\| \leq \gamma_d, \exists n, m \in \mathcal{N}_e(t)) \\ &- \sum_n \sum_{m \in \mathcal{N}_e(t)} c\beta^n a_e^m(t) \delta(t = n\tau_0) dt, \end{aligned} \tag{25}$$

where T^* is the earliest time that all emitters have been caught; i.e.,

$$T^* = \min\{t | |\mathcal{N}_e(t)| = 0\}. \tag{26}$$

Recall that R_0 is the reward for catching an emitter and c is the cost when an emitter transmits in one time slot. We can immediately obtain the instantaneous reward $r(t)$ of the UAVs.

4.2 Multi-UAV-Single-Emitter Game

To study the general case, we first study the special case in which there is only one emitter. Similarly to the single UAV case, we have the following conclusion for the multi-UAV-single-emitter game.

Proposition 3. *The reward-to-functions for the non-discounted reward are determined by*

$$R_{\mathbf{s}}((\tau_0)^-) = \min_{a_e} \left[-cI(a_e = 1) + R_{\mathbf{x}}(0, a_e) \right], \tag{27}$$

and

$$\begin{aligned} R_{\mathbf{s}}(t, 1) = &\max_{\mathbf{f}_u} \min_{\mathbf{f}_e} \\ &\left[\int_t^{\min(\tau, T^*)} R_0 \delta(\exists n, \|\mathbf{x}_n^u(t) - \mathbf{x}^e(t)\| < \gamma_d) dt + R_{\mathbf{s}'}(\tau^-) \right], \end{aligned} \tag{28}$$

and

$$\begin{aligned} R_{\mathbf{s}}(t, 0) = &\min_{\mathbf{f}_e} \\ &\left[\int_t^{\min(\tau, T^*)} R_0 \delta(\exists n, \|\mathbf{x}_n^u(t) - \mathbf{x}^e(t)\| < \gamma_d) dt + R_{\mathbf{s}'}(\tau^-) \right] \Bigg|_{\mathbf{f}_u = \mathbf{f}_0}. \end{aligned} \tag{29}$$

4.3 Multi-UAV-Multi-Emitter Game

Based on the discussion on the multi-UAV-single-emitter case, the general multi-UAV-multi-emitter case can be analyzed in a recursive manner: when an emitter is caught, the game is converted into a game with one less emitter.

Proposition 4. *Suppose that the feedback Nash equilibrium for $N_e - 1$ emitters has been obtained and we use a super script $N_e - 1$ in the reward-to-go function. The reward-to-functions for the non-discounted reward are determined by*

$$R_{\mathbf{s}}^{N_e}((\tau_0)^-) = \min_{a_e} \left[-cI(a_e = 1) + R^P N_{e\mathbf{x}}(0, a_e)\right], \tag{30}$$

and

$$R_{\mathbf{s}}^{N_e}(t,1) = \max_{\mathbf{f}_u} \min_{\mathbf{f}_e} \left[\int_t^{\min(\tau,T^*)} R_0 \delta(\exists n, \|\mathbf{x}_n^u(t) - \mathbf{x}^e(t)\| < \gamma_d) dt + R_{\mathbf{s}'}^{\tilde{N}_e}(\tau^-)\right], \tag{31}$$

where $\tilde{N}_e$ is the number of emitters after the time τ; i.e. ,

$$\tilde{N}_e = \begin{cases} N_e - 1, & \text{if } \exists n, t, \|\mathbf{x}_n^u(t) - \mathbf{x}^e(t)\| < \gamma_d) \\ N_e, & \text{otherwise} \end{cases}, \tag{32}$$

and

$$R_{\mathbf{s}}^{N_e}(t,0) = \min_{\mathbf{f}_e} \left[\int_t^{\min(\tau,T^*)} R_0 \delta(\exists n, \|\mathbf{x}_n^u(t) - \mathbf{x}^e(t)\| < \gamma_d) dt + R_{\mathbf{s}'}^{\tilde{N}_e}(\tau^-)\right]\Bigg|_{\mathbf{f}_u = \mathbf{f}_0}. \tag{33}$$

5 Numerical Results

In this section, we use numerical simulations to disclose some phenomena of the pursuit-evasion game. For simplicity, we consider only one UAV and one RF emitter.

5.1 Simulation Setup

We consider abstract length and time units. We assume $v_u = 0.1$, $v_e = 0.02$, $F_u = 0.05$ and $F_v = 0.1$. We assume $\gamma_d = 0.1$. Unless stated otherwise, the penalty of the RF emitter being caught by the UAV is 10, while the penalty of not transmitting is 3. For the case of discounted reward, the discounting factor is $\beta = 0.9$. We discretize d, $\delta\theta_1$ and $\delta\theta_2$ into $40 \times 20 \times 20$ grid. The value function is obtained from 50 iterations. For the case of non-discounted reward, we set $t_f = 5$, i.e., the RF only need to consider the game within 5 decision periods.

5.2 Case of Discounted Reward

Fig. 2 shows the value functions of different cases. We observe that the value function is high when $\delta\theta_1$ is close to zero. The reason is that both the UAV and RF emitter have similar initial direction; hence it is easier for the UAV to catch

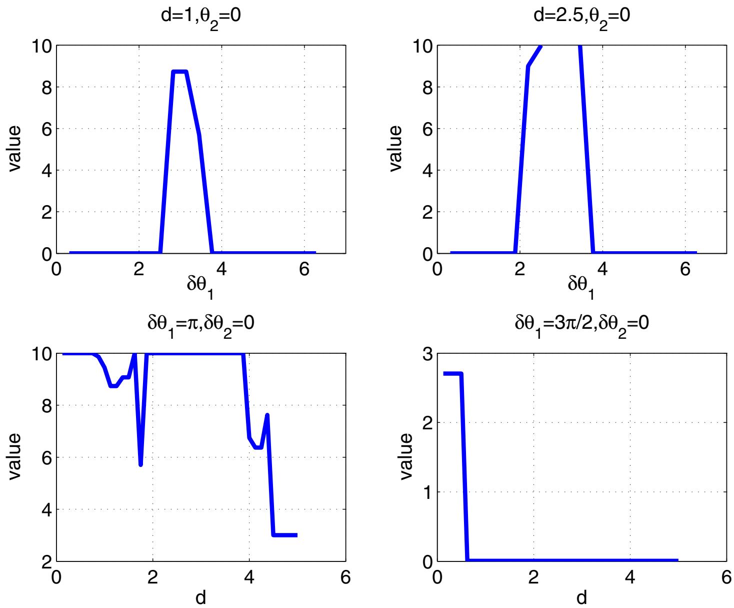

Fig. 2. Samples of value functions

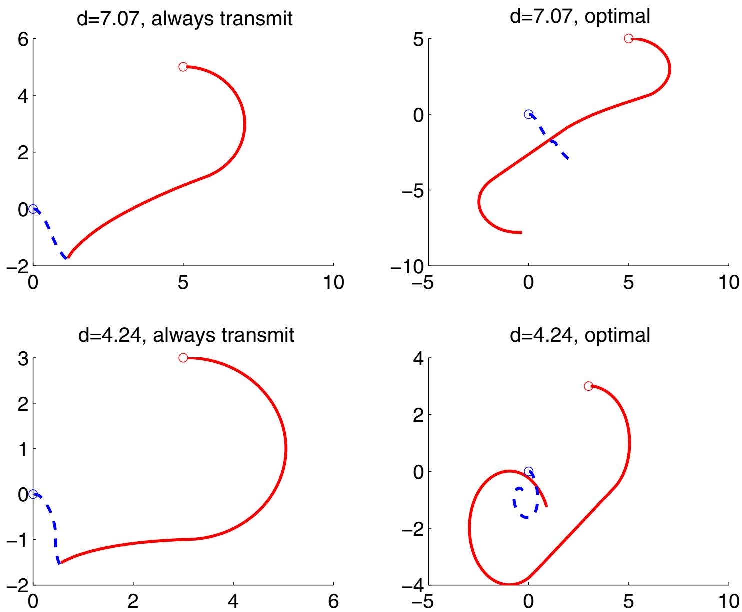

Fig. 3. Samples of tracks

the RF emitter. We also observe that the value usually decreases as the initial distance between UAV and RF is large (but there are some exceptions).

Fig. 3 shows the tracks of the UAV and RF emitter with different initial distances. In the left columns, the RF emitter always keeps transmitting; finally, it will be caught by the UAV. In the right column, the RF emitter adopts the optimized strategy. We observe that the RF emitter can escape from the pursuit of the UAV by stopping transmitting in certain times.

Then, we increase the penalty of stopping transmitting to 8. The tracks using the corresponding optimal strategy is shown in Fig. 4. We observe that, in both cases, the RF emitter is finally caught by the UAV, due to the large penalty of stopping transmitting.

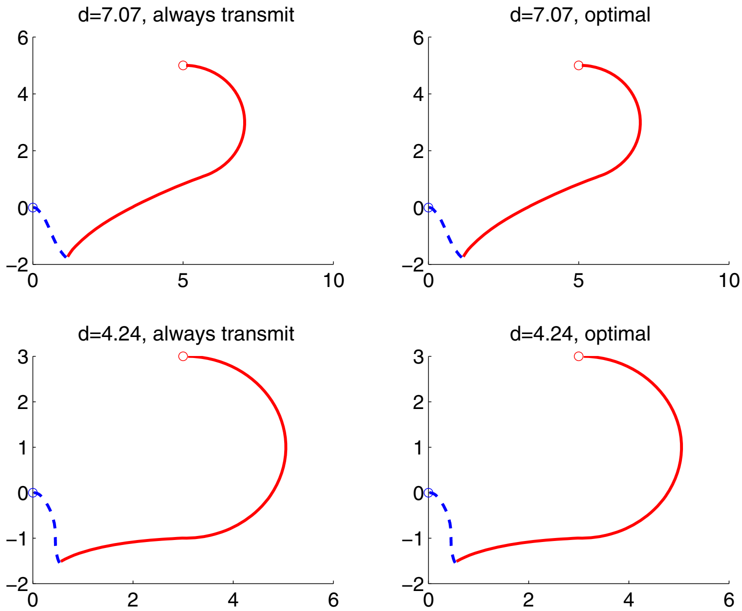

Fig. 4. Samples of tracks when the penalty of ceasing transmitting is increased

5.3 Case of Non-discounted Reward

For the case of no-discount reward, the value functions and the optimal actions in different stages are shown in Fig. 5. We observe that, in the 5-th stage, the RF emitter more intends to keep transmitting and take the risk of being caught by the UAV. The sample tracks are shown in Fig. 6. We observe that, in the first situation, the RF emitter stops transmitting to avoid the UAV at the beginning and finally gets caught by the UAV.

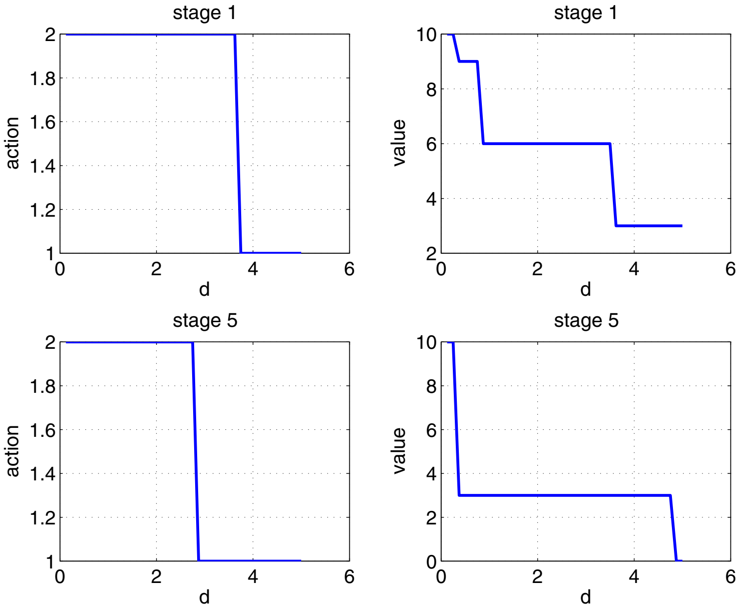

Fig. 5. Samples of value functions and optimal actions when the reward is not discounted

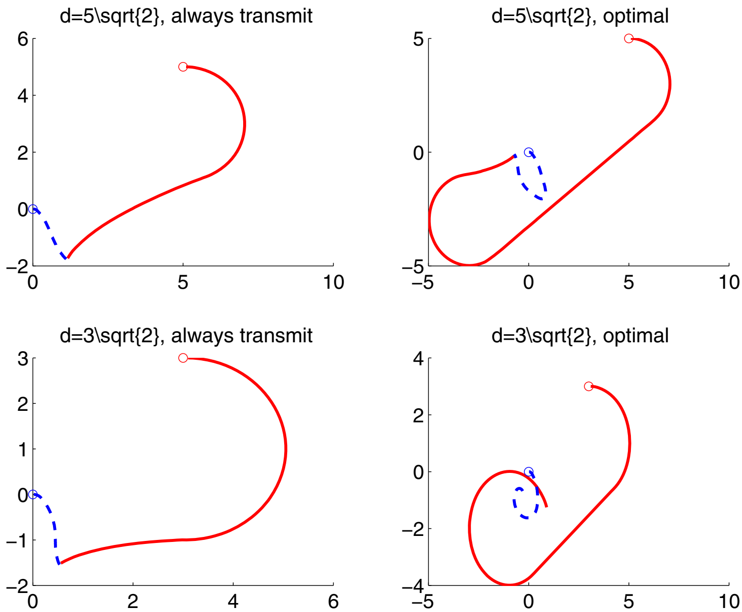

Fig. 6. Samples of tracks when the reward is not discounted

6 Conclusions

A The Isaccs Equation

We consider a differential game with N players over time period $[0, T]$, whose dynamics are given by (the system state $\mathbf{x}$ is in $\mathcal{R}^M$)

$$\dot{\mathbf{x}}(t) = f(t, \mathbf{x}(t), u_1(t), ..., u_N(t)), \tag{34}$$

and the cost functionals are given by

$$\begin{aligned} & L_n(u_1, ..., u_N) \\ = & \int_0^T g_n(t, \mathbf{x}(t), u_1(t), ..., u_N(t))dt + q_n(\mathbf{x}(\mathbf{T})). \end{aligned} \tag{35}$$

We assume that each player has perfect access to all dimensions of the system state; i.e., the closed-loop perfect state (CLPS). The following definition defines the feedback Nash equilibrium for the differential game.

Definition 1. *For the N-player game in (34) and (35), an N-tuple of strategies $\{\pi_n^*\}_{n=1,...,N}$ consists of a feedback Nash equilibrium solution if there exist functionals V_n over $[0, T] \times R^M$ such that*

$$V_n(T, \mathbf{x}) = q_n(\mathbf{x}), \tag{36}$$

$$\begin{aligned} V_n(t, \mathbf{x}) = & \int_t^T g_n(t, \mathbf{x}^*(s), \pi_1^*(\mathbf{x}^*), ..., \pi_N^*(\mathbf{x}^*))ds \\ & + q_n(\mathbf{x}^*(T)) \\ \le & \int_t^T g_n(t, \mathbf{x}(s), \pi_1^*(\mathbf{x}), ..., \pi_{n-1}^*(\mathbf{x}), \pi_n(\mathbf{x}), \\ & \pi_{n+1}^*(\mathbf{x}) ..., \pi_N^*(\mathbf{x}))ds + q_n(\mathbf{x}^*(T)), \qquad \forall \pi_n, \end{aligned} \tag{37}$$

where x^* is the trace of state when the actions are $\pi_1^*(s), ..., \pi_N^*(s)$ and x is the state trace when the action of player n is changed to π_n.

The following theorem provides a sufficient condition for the feedback Nash equilibrium for the general N-player case.

Theorem 1. *An N-tuple of strategies $\{\pi_n^*\}_{n=1,...N}$ provides a feedback Nash equilibrium if the functionals $\{V_n\}_{n=1,...,N}$ satisfy the following equations:*

$$\begin{aligned} -\frac{\partial V_n(t, \mathbf{x})}{dt} = \min_{u_n} \Bigg[& \frac{\partial V_n(t, \mathbf{x})}{\partial \mathbf{x}} f(t, \mathbf{x}, \{\pi_{-n}^*(t, \mathbf{x}), u_n\}) \\ & + g(t, \mathbf{x}, \{\pi_{-n}^*(t, \mathbf{x}), u_n\}) \Bigg], \end{aligned} \tag{38}$$

and

$$\begin{aligned} \pi_n^*(t, \mathbf{x}) = \arg\min_{u_n} \Bigg[& \frac{\partial V_n(t, \mathbf{x})}{\partial \mathbf{x}} f(t, \mathbf{x}, \{\pi_{-n}^*(t, \mathbf{x}), u_n\}) \\ & + g(t, \mathbf{x}, \{\pi_{-n}^*(t, \mathbf{x}), u_n\}) \Bigg], \end{aligned} \tag{39}$$

and

$$V_n(T, \mathbf{x}) = q_n(\mathbf{x}). \tag{40}$$

The following theorem provides a sufficient condition for two-player zero-sum game in which the cost for player 1 is given by

$$L(u_1, u_2) = \int_0^T g(t, x(t), u_1(t), u_2(t))dt + q(T, x(T)), \tag{41}$$

and the cost of player 2 is $-L(u_1, u_2)$.

Theorem 2. *The value function of the two-player zero-sum differential game satisfies the following Isaacs equation:*

$$\begin{aligned} -\frac{\partial V}{\partial t} &= \min_{u_1} \max_{u_2} \left[\frac{\partial V}{\partial \mathbf{x}} f(t, \mathbf{x}, u_1, u_2) + g(t, \mathbf{x}, u_1, u_2) \right] \\ &= \max_{u_2} \min_{u_1} \left[\frac{\partial V}{\partial \mathbf{x}} f(t, \mathbf{x}, u_1, u_2) + g(t, \mathbf{x}, u_1, u_2) \right] \end{aligned} \tag{42}$$

References

1. Başar, T., Olsder, G.J.: Dynamic Noncooperative Game Theory, 2nd edn. Society for Industrial and Applied Mathematics (1999)
2. Beard, R.W., McLain, T.W., Nelson, D.B., Kingston, D., Johanson, D.: Decentralized cooperative aerial surveillance using fixed-wing miniature UAVs. Proceedings of the IEEE 94(7), 1306–1324 (2006)
3. Bullo, F., Cortes, J., Martinez, S.: Distributed Control of Robotic Networks: A Mathematical Approach to Motion Coordination Algorithms. Princeton University Press (2009)
4. DeLima, P., York, G., Pack, D.: Localization of ground targets using a flying sensor network. In: Proc. of IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing, vol. 1, pp. 194–199 (2006)
5. Elsaesser, D.: Emitter geolocation using low-accuracy direction-finding sensors. In: IEEE Symposium on Computational Intelligence for Security and Defense Applications, CISDA, pp. 1–7 (2009)
6. Isaacs, R.: Differential Games. Wiley (1965)
7. Lunze, J., Lararrigue, F.L.: Handbook of Hybrid Systems Control: Theory, Tools and Applications. Cambridge Univ. Press (2009)
8. Nerode, A., Remmel, J.B., Yakhnis, A.: Hybrid system games: Extraction of control automata with small topologies. In: Handbook of Hybrid Systems Control: Theory, Tools and Applications. Cambridge Univ. Press (2009)
9. Scerri, P., Glinton, R., Owens, S., Sycara, K.: Locating RF Emitters with Large UAV Teams. In: Pardalos, P.M., Murphey, R., Grundel, D., Hirsch, M.J. (eds.) Adv. in Cooper. Ctrl. & Optimization. LNCIS, vol. 369, pp. 1–20. Springer, Heidelberg (2007)
10. Scerri, P., Glinton, R., Owens, S., Scerri, D., Sycara, K.: Geolocation of RF emitters by many UAVs. In: AIAA, Infotech@Aerospace 2007 Conference and Exhibit (2007)
11. Walter, D.J., Klein, J., Bullmaster, J.K., Chakravarthy, C.V.: Multiple UAV tomography based geolocation of RF emitters. In: Proc. of the SPIE Defense, Security, and Sensing 2010 Conference, Orlando, FL, April 5-9 (2010)