

A Hierarchical Approach To Multi-Player Pursuit-Evasion Differential Games

Dongxu Li, *Student Member, IEEE*, Jose B. Cruz, Jr., *Life Fellow, IEEE*,
Genshe Chen, Chiman Kwan, *Senior Member, IEEE*, and Mou-Hsiung Chang

Abstract—The increasing use of unmanned assets and robots in modern military operations renews an interest in the study of general pursuit-evasion games involving multiple pursuers and multiple evaders. Due to the difficulty in formulation and rigorous treatment, the literature in this field is very limited. This paper presents a hierarchical approach to this kind of problem. With an additional structure imposed on decision-making of pursuers, this approach provides conservative guidance to pursuers by finding certain engagement between pursuers and evaders, and the saddle-point strategies are utilized by each pursuer in chasing the engaged evaders. A combinatorial optimization problem is formulated and scenarios are created to demonstrate the feasibility of the algorithm. This is a preliminary study on multi-player pursuit-evasion games and future directions are suggested.

I. INTRODUCTION

In a Pursuit-Evasion (PE) game, the problem of one or a group of pursuers catching one or a group of moving evaders is studied. It has extensive applications such as missile guidance, military strategy, aircraft control and aerial tactics. Under the framework of game theory and optimal control, a number of formal solutions regarding optimal strategies in particular PE problems can be achieved [1]-[2]. In the literature, most studies on PE games have concentrated on two-player games with a single pursuer and a single evader. As the use of unmanned assets and robots increases in modern military operations, newly emergent scenarios usually involve multiple pursuers and evaders. The problem of formulation and computation of optimal pursuit strategies of multiple players in continuous time needs to be addressed.

A PE game is usually formulated as a zero-sum game. Since the 1950s, the deterministic PE game of a single pursuer and a single evader with perfect information and common knowledge has been extensively studied. Isaacs

solved a PE problem for a saddle-point equilibrium solution by the method of “tenet of transition” [1]. Although PE games with multiple pursuers and multiple evaders have been investigated recently, most of them deal with discrete time problems or are in an ad hoc manner [6], [8]. Little has been done for generic multi-player PE differential game problems in continuous time. In this paper, we focus on deterministic multi-player differential PE game problems [4]. We extend Isaacs’ approach on a two-player PE game to a game with multiple players. A conservative strategy is applied from the pursuer’s perspective to achieve an upper-bound of the performance index.

This paper is organized as follows. A generic problem is formulated in the next section. In section III, difficulty of application of conventional optimal control theory on multi-player PE games has been analyzed, and then a suboptimal hierarchical approach is proposed from the pursuers’ perspective. In addition, the issue of capturability is discussed and methods are suggested to address the problem. Simulation results are presented in section IV. Finally, the paper concludes with suggestions for future work.

II. PROBLEM FORMULATION

Consider a general PE differential game with N pursuers and M evaders in a n_0 dimensional space S , $S \subseteq \mathbb{R}^{n_0}$. Denote by x_p^i (x_e^j) the state variable associated with pursuer i , $i = 1, \dots, N$ (evader j , $j = 1, \dots, M$), where $x_p^i \in \mathbb{R}^{n_p^i}$ ($x_e^j \in \mathbb{R}^{n_e^j}$). Notice that n_p^i (n_e^j) $\geq n_0$ because of the specific dynamics of the pursuer (evader). Assume that the first n_0 elements in x_p^i (x_e^j) specify the physical position of pursuer i (evader j) in space S . In general, the dynamic equations for each pursuer i and evader j are

$$\begin{aligned} \dot{x}_p^i(t) &= f_p^i(x_p^i(t), u_i(t)), & x_p^i(0) &= x_{p0}^i \\ \dot{x}_e^j(t) &= f_e^j(x_e^j(t), v_j(t)), & x_e^j(0) &= x_{e0}^j \end{aligned} \quad (1)$$

In (1), $u_i(t) \in U_a^i$, $v_j(t) \in V_a^j$ are control variables for time $t \geq 0$, where $U_a^i \subseteq \mathbb{R}^{m_p^i}$ and $V_a^j \subseteq \mathbb{R}^{m_e^j}$ are the sets of corresponding admissible control actions; function $f_p^i(\cdot, \cdot)$ ($f_e^j(\cdot, \cdot)$) is a mapping from $\mathbb{R}^{n_p^i} \times U_a^i$ ($\mathbb{R}^{n_e^j} \times V_a^j$) to $\mathbb{R}^{n_p^i}$ ($\mathbb{R}^{n_e^j}$). In this paper, we consider the deterministic case, where the function f_p^i (f_e^j) does not depend on time t

Manuscript received September 14, 2005. This work was supported in part by the U.S. Army under Contract W911NF-05-C-0018.

D. Li is with the Department of Electrical and Computer Engineering, The Ohio State University, Columbus, OH 43210 USA. (phone: 614-404-5494; email: li.447@osu.edu)

J.B. Cruz is with the Department of Electrical and Computer Engineering, The Ohio State University, Columbus, OH 43210 USA. (email: jbcruz@ieee.org)

G. Chen is with Intelligent Automation, Inc, 15400 Calhoun Dr, Suite 400, Rockville, MD 20855 USA. (email: gchen@i-a-i.com)

C. Kwan is with Intelligent Automation, Inc, 15400 Calhoun Dr, Suite 400, Rockville, MD 20855 USA. (email: ckwan@i-a-i.com)

M.H. Chang is with the Mathematics Division, U.S. Army Research Office, (email: mouhsiung.chang@us.army.mil).

explicitly. For simplicity of notation, let

$$x_p = [x_p^{1T}, \dots, x_p^{NT}]^T, \quad x_e = [x_e^{1T}, \dots, x_e^{MT}]^T, \\ u = [u_1^T, \dots, u_N^T]^T \text{ and } v = [v_1^T, \dots, v_M^T]^T,$$

and accordingly, function $f_p = [f_p^{1T}, \dots, f_p^{NT}]^T$ and $f_e = [f_e^{1T}, \dots, f_e^{MT}]^T$. Then, the dynamic equation can be rewritten in a compact form as

$$\dot{x}_p(t) = f_p(x_p(t), u(t)), \quad \text{with } x_p(0) = x_{p0}, \\ \dot{x}_e(t) = f_e(x_e(t), v(t)), \quad x_e(0) = x_{e0}. \quad (2)$$

Let $X_p \triangleq \times_{i=1}^N \mathbb{R}^{n_p^i}$, $X_e \triangleq \times_{j=1}^M \mathbb{R}^{n_e^j}$, $U_a \triangleq \times_{i=1}^N U_a^i$ and $V_a \triangleq \times_{j=1}^M V_a^j$.

For pursuer i , define the projection $P: \mathbb{R}^{n_p^i} \rightarrow S \subseteq \mathbb{R}^{n_0}$ as

$$P(x_p^i) = [x_{p1}^i, \dots, x_{pn_0}^i]^T. \quad (3)$$

A similar operator can be defined for each pursuer and each evader. We use the notation P to denote the projection for all the pursuers and the evaders. Clearly, for any $x \in \mathbb{R}^{n_p^i}$, $i = 1, \dots, N$ (or $x \in \mathbb{R}^{n_e^j}$, $j = 1, \dots, M$), $P(x) \in S$.

In a PE game with multiple evaders, evaders are generally not captured simultaneously. The terminal time of the game can be defined based on the capture of all evaders, i.e., for any j , $j = 1, \dots, M$, there exists i , $0 \leq i \leq N$ such that $d(P(x_p^i(t)), P(x_e^j(t))) \leq \varepsilon$ for some $t \geq 0$. Here $d(\cdot, \cdot)$ is a metric in \mathbb{R}^{n_0} , e.g., $d(x, y) = \|x - y\|_2$ for $\forall x, y \in \mathbb{R}^{n_0}$; ε is a predetermined small positive real number. The capture time of evader j , T_j , can be defined as

$$T_j = \inf \left\{ t \geq 0, \exists i, 1 \leq i \leq N, \text{ s.t. } d(P(x_p^i(t)), P(x_e^j(t))) \leq \varepsilon \right\}.$$

Then, the terminal time of the PE game, T , is defined as

$$T = \max_{1 \leq j \leq M} \{T_j\} \quad (4)$$

Clearly, by (4), $T \in \mathbb{R}^+ \cup \{0, \infty\}$, where, \mathbb{R}^+ stands for the set of all positive real numbers. Consider that the objective function has a general form as follows, where subscript t denotes the time.

$$J(x_{p0}, x_{e0}, u, v) = \int_0^T G(x_{pt}, x_{et}, u_t, v_t) dt + Q(x_T) \quad (5)$$

subject to (2).

In (5), $u \in \{u(t) | 0 \leq t \leq T, u(t) \in U_a\}$ and similar for v .

Function G is the cost rate,

$$G(\cdot, \cdot, \cdot, \cdot): X_p \times X_e \times U_a \times V_a \rightarrow \mathbb{R}^+ \cup \{0\},$$

and function Q quantifies the terminal cost as a function of the states at the terminal time of a game, $Q(\cdot): X \rightarrow \mathbb{R}^+ \cup \{0\}$. In a game where the capture time is the objective, $G=1$ and $Q=0$. In this paper, PE games involving N pursuers and M evaders are modeled as zero-sum games, where pursuers try to find an optimal

strategy u^* by minimizing the objective (5) subject to (2); while evaders try to maximize it. We use the notation $U_a[s, t]$ to stand for the following set.

$$U_a[s, t] \triangleq \{u(\tau) | u(\tau) \in U_a, 0 \leq s \leq \tau \leq t\}$$

Use the notation $u_{s:t}$, $u_{s:t} \in U_a[s, t]$.

By (5), for $\forall x_p \in X_p$ and $x_e \in X_e$, define $\bar{V}(x_p, x_e)$ as

$$\bar{V}(x_p, x_e) = \min_{u_{0:T} \in U_a[0, T]} \max_{v_{0:T} \in V_a[0, T]} \left\{ \int_0^T G(x_{pt}, x_{et}, u_t, v_t) dt + Q(x_T) \right\}.^1 \quad (6)$$

Similarly for $\bar{V}(x_p, x_e)$,

$$\underline{V}(x_p, x_e) = \max_{v_{0:T} \in V_a[0, T]} \min_{u_{0:T} \in U_a[0, T]} \left\{ \int_0^T G(x_{pt}, x_{et}, u_t, v_t) dt + Q(x_T) \right\}. \quad (7)$$

$$\text{If } \bar{V}(x_p, x_e) = \underline{V}(x_p, x_e) = V(x_p, x_e), \quad (8)$$

$V(x_p, x_e)$ is called the **Value function**, and this condition is called the **Isaacs condition** [2]. This Value is the so-called saddle-point equilibrium defined for a zero-sum game [4]. In this paper, the capitalized ‘‘Value’’ stands for the value function defined in (8), avoiding the confusion with the ordinary meaning of the word.

III. CONSERVATIVE HIERARCHICAL APPROACH

A. Dilemma in Backward Analysis

A PE game has been formulated as a zero-sum game and the saddle-point equilibrium concept is adopted under the Isaacs condition. This fact makes the mathematical tools that are available for solving optimal control problems also useful in solving conventional two-player PE differential games. The kernel of modern optimal control theory includes Pontryagin’s minimum principle and Bellman’s dynamic programming. Both of them specify a set of conditions on optimal dynamic controls. The former involves a set of adjoint Ordinary Differential Equations (ODE) while the latter is associated with a partial differential equation called the Hamilton-Jacobi-Bellman² (HJB) equation [15]. In both approaches, boundary conditions on the states at the terminal time are needed. Furthermore, Isaacs’ method of ‘‘tenet of transition’’ in treating two-player PE differential games is closely related to dynamic programming, which is based on the underlying idea of state rollback. Starting from the terminal, an optimal trajectory of the states is traced backwards and with a formulation of the Hamilton-Jacobi-Isaacs (HJI) equation the Value function can be determined, such that saddle-point equilibrium strategies of state feedback can be obtained accordingly.

In the problem of a PE game involving multiple players, the backward analysis encounters tremendous difficulty in

¹ We assume the optimum can be attained by some strategy pair (u^*, v^*)

the process of tracing back. The main challenge lies in identifying the terminal states of both pursuers and evaders. Starting from the end, under a saddle-point strategy, each evader has been captured by some specific pursuer. If pursuer i catches evader j in the game, we say that both players are engaged. (Note that it is possible that one evader may be engaged with more than one pursuer and vice versa.) Clearly, the number of possible engagements between pursuers and evaders increases at least exponentially with N and M . This explosion of the number of possible engagements makes the terminal state in a multi-player PE game extremely difficult to specify in contrast to that in a two-player game. If not impossible, the backward analysis should be repeated so many times starting from different engagements even when N and M are not very large. Furthermore, evaders are generally not captured at the same time, which makes determining the terminal state further intractable. Therefore, the backward approach used by Isaacs and Bellman cannot be practically applied in multi-player PE games and the same situation holds for the minimum principle as well. In summary, a multi-player PE differential game cannot be characterized by either a set of adjoint ODEs or a HJI equation.

B. Hierarchical Approach

To attack this problem, first let us study whether or not the game can be decomposed into games between the pursuers and the evaders that are closely engaged. Assume that there exist continuous trajectories for the group of pursuers and evaders that comprise a saddle-point solution for the game. The engagement scheme among the pursuers and evaders can be obtained at the terminal time of the game. Under this engagement, pursuers do not switch evaders if evaders do not change their strategy, which is the basic rationale of the concept of equilibrium in game theory [2], [4]. In this sense, the problem of determining an engagement may be part of the pursuers' strategies if they try to maximize (5). This can be viewed as a natural hierarchical structure in decision-making from the pursuers' perspective. The upper level is to determine a proper engagement scheme among pursuers and evaders, while the lower level solves each engaged differential game. However, given an engagement scheme, the original game cannot be treated as a set of decoupled PE games between pursuers and their engaged evaders. Coordination fills this gap.

The major difficulty of multi-player PE differential games is to characterize the coordination among pursuers or evaders. It must be every difficult to solve it by Isaacs' approach for various terminal times and engagements if it is not impossible. In PE games, Singular Surfaces (SS) are almost always assumed to divide the state space into disjoint regions with continuous differentiability [1]-[2]. We think the

number of SS increases dramatically with the number of pursuers and evaders, which makes the rigorous treatment of this kind of problem extremely difficult. It is because conventional optimal control theory requires conditions on the smoothness of optimal functions.

Instead of determining the exact equilibrium solutions, we seek an upper-bound of the performance index from the pursuers' perspective. In Isaacs' approach, the "minmax" operator in (6) is to determine the best worst-case strategy. It is a conservative approach and the optimal objective acts as the least uniform (no dependence on evaders) upper-bound on the pursuers' performance. When this result coincides with that from (7), it becomes a well-defined solution for zero-sum games. In this paper, we determine a good upper bound instead of the least upper bound.

We focus on the capture time as our objective function. The objective function for a game with multiple players is $T = \max_j \{T_j\}$. Assume that each pursuer is engaged with at least one evader, and only captures one evader at a time. The original multi-player PE game can be converted into a hierarchical optimization problem. The upper level is to determine such an engagement that T is minimized. Given an engagement, the strategy of each pursuer is obtained by solving decoupled two-player PE games based on Isaacs' method at the lower level. This is a conservative approach because in many situations, the strategies of multiple pursuers are concealed so that the evaders cannot execute the "optimal" strategies against the engaged pursuers. The structure of the approach is illustrated in Fig. 1.

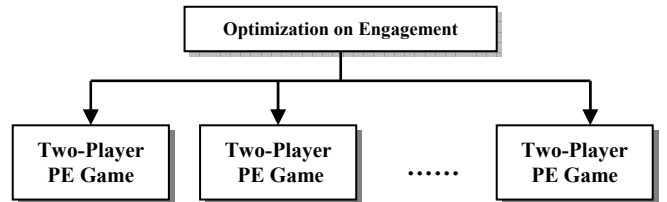


Fig. 1. Hierarchical Structure Approach

C. Optimization at the Upper Level

Let $V(x_p^i, x_e^j)$ denote the Value function if pursuer i is engaged with evader j . Notice that $V(x_p^i, x_e^j) = T_j$. Assume that $V(x_p^i, x_e^j)$ can be solved analytically and the optimization problem at the upper level can be formulated as

$$\min J = \min_{\{b_{ijk}\}} \left\{ \max_j \left\{ \sum_{i=1}^N \sum_{k=1}^K (V(x_p^i(k), x_e^j(k)) \cdot b_{ijk}) \right\} \right\} \quad (9)$$

subject to $b_{ijk} \in \{0, 1\}$, $\sum_{i=1}^N \sum_{k=1}^K b_{ijk} = 1$ and $\sum_{j=1}^M b_{ijk} \leq 1$.

In (9), the problem is formulated with multiple stages, taking into account the case when $N < M$. Here, k is the index for stages; b_{ijk} is a binary decision variable; $b_{ijk} = 1$ indicates that pursuer i is engaged with evader j at stage k ; $b_{ijk} = 0$

² In differential games, the corresponding equation is called the Hamilton-Jacobi-Isaacs (HJI) equation, where "minimax" replaces "min".

means the opposite. The maximum number of stages considered is $K = \lceil M/N \rceil$, which is the smallest integer greater than M/N . Solving problem (9) provides an upper bound of the objective for the original multi-player PE game.

When $N \geq M$ ($K = 1$), problem (9) can be converted into a standard Mixed Integer Linear Programming (MILP) problem by introducing a slack variable w , such that commercial solvers such as CPLEX and LINDO can be utilized [10]. A proper formulation is shown in (10).

$$\min J = \min \{w\} \quad (10)$$

subject to $T_j \leq w$, for $1 \leq j \leq M$, $b_{ij} \in \{0,1\}$,

$$\sum_{i=1}^N b_{ij} = 1 \text{ and } \sum_{j=1}^M b_{ij} \leq 1, \text{ with } T_j = \sum_{i=1}^N (V(x_p^i, x_e^i) \cdot b_{ij}).$$

D. PE Differential Game of Two Players at the Lower Level

Problem (9) requires solving $V(x_p^i, x_e^i)$. A number of solutions have been achieved analytically [1]-[2]. The following theorem specifies a set of conditions for an equilibrium solution. Consider a general dynamics as

$$\dot{x}(t) = f(x, u, v) \text{ with } x(0) = x_0 \in \mathbb{R}^n. \quad (11)$$

The objective function is

$$J = \int_0^T G(x_t, u_t, v_t) dt + Q(x_T). \quad (12)$$

Theorem 1: For a two-player PE differential game described in (11) and (12), suppose that the control pair (u^*, v^*) is a saddle-point solution and $x^*(t)$ is the corresponding trajectory, there exists a costate function $p(t): [0, T] \rightarrow \mathbb{R}^n$ such that the following conditions are satisfied:

$$\dot{x}^*(t) = f(x_t^*, u_t^*, v_t^*)$$

$$\dot{p}(t) = -\nabla_x H(p_t, x_t^*, u_t^*, v_t^*)$$

$$H(p_t, x_t^*, u_t^*, v_t^*) \leq H(p_t, x_t^*, u_t^*, v_t^*) \leq H(p_t, x_t^*, u_t^*, v_t^*)$$

with $x^*(0) = x_0$ and $p(T) = \nabla_x Q(x^*(T))$ at the terminal time, where $H(p_t, x_t, u_t, v_t) = G(x_t, u_t, v_t) + p_t^T f(x_t, u_t, v_t)$.

Proof: See [2]. \square

A two-player PE problem can be difficult depending on the dynamics and the positions of the players and analytical solutions may be intractable. Here, we illustrate the idea by solving a PE game with simplified dynamics in order to reduce the complexity of determining $V(x_p^i, x_e^i)$. For practical problems, it is desired that the model reveal the main features of the players without involving too much of the details. Consider a PE game with two dimensions. The dynamics of the pursuer and the evader are

$$\begin{aligned} \dot{x}_p &= v_p \cos \theta_p, & \dot{x}_e &= v_e \cos \theta_e \\ \dot{y}_p &= v_p \sin \theta_p, & \dot{y}_e &= v_e \sin \theta_e \end{aligned} \quad (13)$$

Here θ is the control variable. Let the initial conditions be x_{p0}, y_{p0}, x_{e0} and y_{e0} . In this model, players are assumed to

perform at their maximum speeds and have complete maneuverability. Small ground vehicles can be approximately described by this dynamics. Define the new states as $x = x_p - x_e$ and $y = y_p - y_e$. It follows that

$$\dot{x} = v_p \cos \theta_p - v_e \cos \theta_e \quad (14)$$

$$\dot{y} = v_p \sin \theta_p - v_e \sin \theta_e$$

with $x_0 = x_p^0 - x_e^0$ and $y_0 = y_p^0 - y_e^0$.

The terminal set Γ is defined as $\Gamma = \{(x, y) \mid \|(x, y)\|_2 \leq \varepsilon\}$.

The objective function is the capture time, $J = \int dt$. The

Hamiltonian is

$$H = 1 + V_x (v_p \cos \theta_p - v_e \cos \theta_e) + V_y (v_p \sin \theta_p - v_e \sin \theta_e).$$

According to Theorem 1, the costate equations are $\dot{V}_x = 0$ and $\dot{V}_y = 0$. Thus, V_x and V_y are constant. By the Main Equation in [1], $\min_{\theta_p} \max_{\theta_e} \{H\} = 0$, and we obtain

$$\cos \theta_p^* = \cos \theta_e^* = -\frac{V_x}{\sqrt{V_x^2 + V_y^2}}, \quad \sin \theta_p^* = \sin \theta_e^* = -\frac{V_y}{\sqrt{V_x^2 + V_y^2}}$$

where $\sqrt{V_x^2 + V_y^2} = 1/(v_p - v_e)$.

By the definition of S , the terminal state (x_T, y_T) satisfies $V_x \cdot (-y_T) + V_y \cdot x_T = 0$. It follows that

$$V_x / x_T = V_y / y_T = (v_p - v_e) / \varepsilon.$$

Considering the system dynamics, then

$$V(x, y) = (\sqrt{x^2 + y^2} - \varepsilon) / (v_p - v_e), \quad (15)$$

$$\cos \theta_p^* = \cos \theta_e^* = -\frac{x}{\sqrt{x^2 + y^2}}, \quad \sin \theta_p^* = \sin \theta_e^* = -\frac{y_0}{\sqrt{x^2 + y^2}}. \quad (16)$$

E. Region of Capture

In practice, the feasibility of the hierarchical approach depends on the capturability between any pair of a pursuer and an evader. In this section, we briefly discuss this issue.

In [1], Isaacs defined concepts of game of degree and game of kind to distinguish between the problem of solving optimal solutions and the problem of existence. For a two-player PE game, denote by C the capture region in the state space, where capturability is guaranteed. On its boundary ∂C , the following condition is satisfied.

$$\min_{u_t} \max_{v_t} \{\eta \cdot f(x_t, u_t, v_t)\} = 0 \quad (17)$$

Here, η denotes the normal direction of ∂C . Equation (17) specifies a necessary condition of the capture region. It is similar to HJI equation regardless of the specific objective function.

An alternative approach is based on a feedback control design method, which combines the concepts of function minimization and Lyapunov stability techniques. This method is called Lyapunov Optimizing Control (LOC) [13]-[14]. The method of LOC can be used to study the

capturability of PE games and to design a feasible pursuit strategy for pursuers [14]. It depends on a positive definite Lyapunov-type function $W(x)$. For any x in the state space, if the following condition is satisfied for any $t \geq 0$,

$$\inf_{u_t} \{\dot{W}(x)\} = \inf_{u_t} \{W_x(x) \cdot f(x, u_t, v_t)\} < 0 \text{ for } \forall v_t \in V_a,$$

then x belongs to the capture region. Clearly, this condition is sufficient and can be easily verified by stability analysis based on the function $W(x)$.

F. Summary and Discussion

The hierarchical approach introduced above is a suboptimal method. Instead of solving the problem formulated in (6) within the admissible set $U_a[0, T]$, we do optimization within a subset $U_a^S[0, T]$ of $U_a[0, T]$ by imposing a structure S on the strategy of the pursuers, dividing the decision-making of pursuers into two levels. The problem in (6) can be rewritten in the following form.

$$\tilde{V}(x_p, x_e) = \min_{u_{0T} \in U_a^S[0, T]} \max_{v_{0T} \in V_a[0, T]} \left\{ \int_0^T G(x_p, x_e, u_t, v_t) dt + Q(x_T) \right\} \quad (18)$$

Equation (18) is almost the same as (6) except that optimization of the pursuers is taken over a structured control set $U_a^S[0, T]$. In some sense, it is only a ‘‘local’’ optimum. Clearly, the solution of (18) is an upper-bound of the original objective function in (5), and $\tilde{V}(x_p, x_e) \geq V(x_p, x_e)$.

It is also worth noting that the hierarchical approach described in this section is conceptually identical to a two-level Stackelberg game problem [11] or bi-level programming [12], where the optimization at the upper level depends on that at the lower level. Note that in a two-level Stackelberg problem, there may be more than one player at the second level, where each player chooses a strategy according to its individual objective. One special case is that those objective functions are decoupled from each other, which fits in the hierarchical approach in this section. In general, for the more complicated coupled lower level, one has to formulate the second level as a game, and relevant solution concepts such as Nash, Pareto and etc may be adopted.

Finally, the quality of the solution from this suboptimal hierarchical approach depends on the combinatorial optimization problem in (9). This is a NP-hard problem [16], and the computational demand increases at least exponentially with the number of pursuer and of evaders, provided the fact the underlying $V(x_p^i, x_e^j)$ between pursuer i and evader j are solvable. Practically, for a PE game involving many pursuers and evaders, heuristic and approximation methods should be applied on (9). In that case, the upper-bound obtained is further degraded by loss of optimality. This result would still be acceptable since no better solution is available.

IV. SIMULATION AND DISCUSSION

As stated earlier, conventional optimal control theory is not applicable. Even for the simplest multi-player PE game with two pursuers and one evader, and if the evader is caught by one pursuer, the terminal state of the other pursuer cannot be specified. In this section, only the result from the hierarchical approach is presented. We create a pursuit evasion scenario involving 3 pursuers and 5 evaders in a two-dimensional space. The dynamics of players are given in (13). The capture time is considered as the objective function. The necessary parameters and the initial states are in Table I.

TABLE I
INITIAL STATES OF PURSUERS AND EVADERS

Pursuers					
	1	2	3		
(x_{p0}, y_{p0})	(0, 3)	(0, 5)	(0, 7)		
v_p (1/sec)	6	5	7		
ω_p (rad/sec)	0.8π	0.8π	0.8π		
θ_{p0}	$\pi/2$	$\pi/2$	$\pi/2$		
Evaders					
	1	2	3	4	5
(x_{e0}, y_{e0})	(1, 5)	(4, 5)	(6, 5)	(7, 5)	(9, 5)
v_e (1/sec)	5	4	3	3	5

Solving the combinatorial optimization problem in (9), we obtain the optimal engagement shown in Table II. The corresponding capture time is 8.8 seconds.

TABLE II
BEST ENGAGEMENT RESULT

Order	P1	P2	P3
1	E5	E3	E2
2	N/A	E4	E1

Next, we consider a little more complicated dynamics

$$\begin{aligned} \dot{x}_p^i &= v_p^i \cos \theta_p^i, & \dot{x}_e^j &= v_e^j \cos \theta_e^j, & x_{p0}^i &= x_{e0}^j \\ \dot{y}_p^i &= v_p^i \sin \theta_p^i, & \dot{y}_e^j &= v_e^j \sin \theta_e^j, & y_{p0}^i &= y_{e0}^j, \\ \dot{\theta}_p^i &= \omega_p^i u_p^i, & \dot{\theta}_e^j &= \omega_e^j u_e^j, & \theta_{p0}^i &= \theta_{e0}^j \end{aligned} \quad (19)$$

where, u_p and u_e are control variables, and assume $-1 \leq u_p, u_e \leq 1$; $v_p, v_e, \omega_p, \omega_e$ are constant. Consider the case $\omega_e \gg \omega_p$, i.e., evaders are viewed to directly control their orientations. This model was originally used by Isaacs in studying the homicidal chauffeur game [1]. The angular velocity ω and the initial orientation θ of each pursuer is given in Table I (shaded). It can be verified that each evader j is in the capture region of some pursuer i [2].

While in the capture region, the optimal strategy of evader j can be shown as that in (16) by maximizing the time derivative of the distance between pursuer i and evader j .

$$\dot{D} = \frac{x(v_e^j \cos \theta_e^j - v_p^i \cos \theta_p^i) + y(v_e^j \sin \theta_e^j - v_p^i \sin \theta_p^i)}{\sqrt{x^2 + y^2}}$$

Here $x = x_e^j - x_p^i$ and $y = y_e^j - y_p^i$. Given a constant strategy for evader j , the strategy of pursuer i may be solved by taking the second order derivative of D , which yields

$$u_p^i = -\text{sign}(x \sin \theta_p^i - y \cos \theta_p^i). \quad (20)$$

With the hierarchy imposed on the decision-making of the pursuers, given a possible engagement scheme, we simulate every decoupled game, in which pursuers utilize the strategy in (20) and evaders play according to (16). All possible engagements are enumerated and the best engagement result is the same as that in Table II, but the capture time increases to 12.5 seconds. The corresponding pursuit-evasion trajectories are shown in Fig. 2, in which the trajectory of each pursuer with its engaged evaders is plotted separately. Snapshots at the 1st and the 5th second are illustrated.

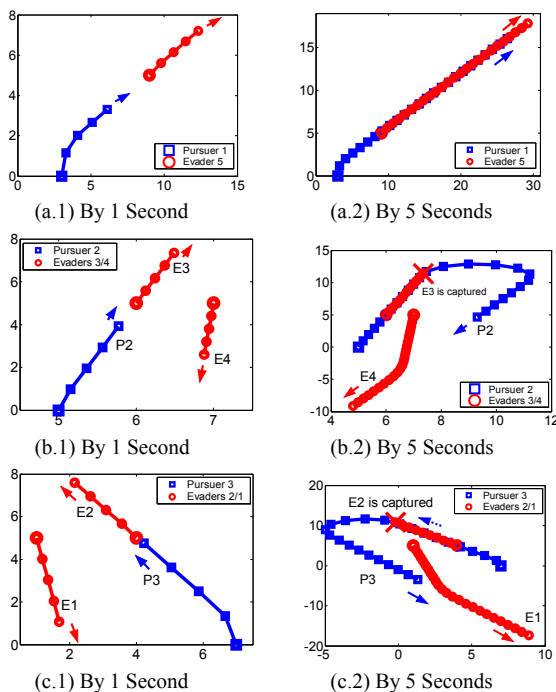


Fig. 2. Pursuit-evasion Trajectories under the Best Engagement.

The simulation results demonstrate the cooperation of pursuers when the hierarchical approach is applied, where pursuers cooperate by choosing appropriate evaders to go after. In the trajectories shown in Fig. 2, the “best” strategy is utilized by evaders against each pursuer when engagement is determined. This pursuit strategy is conservative from the pursuer’s perspective, since in practice evaders may not know the pursuer’s strategies perfectly.

V. CONCLUSION AND FUTURE WORK

In this paper, we deal with a multi-player PE differential game. A generic problem is formulated. Conventional optimal control theory is not applicable to this kind of problem due to the difficulty in specifying the final states. A suboptimal method is proposed to calculate a locally optimal strategy with a specified control structure imposed on pursuers. A combinatorial optimization problem is

formulated to determine an optimal engagement involving a single pursuer and a single evader. The underlying two-player games are solved by currently available differential game theory. The issue of capturability is discussed. Simulations show the feasibility of the approach. This suboptimal approach provides a theoretical upper-bound to a multi-player PE game when exact equilibrium solutions are unknown. In practice, solving the combinatorial optimization problem suffers from the difficulty of its NP-hardness.

Future work falls in the following directions: 1) the locally optimal strategy obtained by the hierarchical method may be improved and if the improvement can be implemented iteratively, a true equilibrium solution may be approached asymptotically; 2) approximation and heuristic methods may be designed to reduce the complexity of solving the combinatorial optimization problem in practical implementations; 3) PE games may be modeled in a stochastic environment for more realistic conditions.

REFERENCES

- [1] Isaacs, *Differential Games*, John Wiley & Sons, Inc., New York, 1965.
- [2] T. Basar and G.J. Olsder, *Dynamic Noncooperative Game Theory*, 2nd Ed, the Society for Industrial and Applied Mathematics, 1998.
- [3] Bertsekas, D.P. 2000. *Dynamic Programming and Optimal Control*: Volume 1, 2nd Edition. Athena Scientific, Belmont, Massachusetts.
- [4] M. J. Osborne and A. Rubinstein, *A course in game theory*, MIT press, Cambridge, Massachusetts, 1994, pp 73-89.
- [5] V. Turetsky and J. Shinar, J. “Missile Guidance Laws Based On Pursuit-Evasion Game Formulations,” *Automatica*, Vol. 39, No. 3, pp. 740-746, 2003.
- [6] Vidal, R., Shakernia, O., Kim, H.J., Shim, D.H. and Sastry, S, “Probabilistic Pursuit-evasion Games: Theory, Implementation, and Experimental Evaluation,” *IEEE Transactions on Robotics and Automation*, v. 18, pp. 662- 669, 2002.
- [7] S. M. LaValle, D. Lin, L. Guibas, J. C.Latombe, and R. Motwani. “Finding an unpredictable target in a workspace with obstacles,” *In IEEE Int. Conf. Robot. & Autom.*, 1997.
- [8] R. Vidal, S. Rashid, C. Sharp, O. Shakernia, J. Kim, and S. Sastry. “Pursuit-evasion games with unmanned ground and aerial vehicles,” *IEEE Int. Conf. Robot. & Autom.*, pages 2948-2955. 2001.
- [9] H. Yamaguchi, “A distributed motion coordination strategy for multiple non-holonomic mobile robots in cooperative hunting operations,” *in Proceedings of the 41st IEEE Conference on Decision and Control*, pp. 2984-2991, December 2002.
- [10] Tom Schouwenaars, Eric Feron, Bart de Moor, and Jonathan How, “Mixed Integer Programming for Multi-vehicle Path Planning,” *European Control Conference*, September 2001.
- [11] M. Simaan and J. B. Cruz Jr., On the Stackelberg strategy in nonzero-sum games, *Journal of Optimization Theory and Applications*, V. 11, 533 - 555, No. 5, 1973.
- [12] B. Colson, P. Marcotte and G. Savard, Bilevel programming: A survey, *A Quarterly Journal of Operations Research*, 3, 87– 107, 2005.
- [13] T.L. Vincent, Guidance Against Maneuvering Targets Using Lyapunov Optimizing Feedback Control, *Proceedings of the American Control Conference*, Anchorage, AK, 2002.
- [14] D.J. Sticht, T.L. Vincent and D.G. Schultz, Sufficiency Theorems for Target Capture, *Journal of Optimization Theory and Applications*, Vol.17, No.5/6, 1975.
- [15] Yong, J. and Zhou, X.Y., *Stochastic Controls: Hamiltonian Systems and HJB Equations*, Springer, 1999.
- [16] C.H. Papadimitriou and K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity*. Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1982.