Please select category below:
Normal Paper ☐
Student Paper ☒
Young Engineer Paper ☐

Orbital Pursuit-Evasion Games with Incomplete Information in the Hill Reference Frame

Hai Zhu [1,2], Ya-Zhong Luo [1], Zhen-Yu Li [1] and Zhen Yang [1]

[1] *College of Aerospace Science and Engineering, National University of Defense Technology, Changsha, 410073, China*
[2] *Department of Cognitive Robotics, Delft University of Technology, Delft, 2628 CD, Netherlands*

Abstract

This paper presents an online information estimation method and designs optimal control laws for orbital pursuit-evasion games with incomplete information. Pursuit-evasion of two spacecraft in the Hill reference frame is considered, in which each player only has knowledge on its own cost function and control gains in the game. From the perspective of the pursuer, an augmented game state estimation problem is formulated to estimate the evader’s cost function and control gains and is performed using an Unscented Kalman Filter (UKF). Furthermore, the optimal control law of the pursuer is designed based on the estimated game information. Numerical simulation results show that the proposed online information estimation method and control law design can overcome the performance loss of the pursuer due to incomplete information in orbital pursuit-evasion.

Keywords: orbital pursuit-evasion, differential games, incomplete information.

Introduction

The orbital pursuit-evasion problem of two spacecraft is typically modelled as a two-player zero-sum differential game [1]. The two players have contradictory objectives and are called the pursuer and evader respectively. Existing researches on orbital pursuit-evasion generally assume the game has complete information [2], i.e. each player has full knowledge of the control gains of the other player. Under the assumption, finding optimal strategies of both players in the game results in solving a two-sided optimization problem [3], which can be solved by bi-level programming [4], numerical shooting [5], semi-direct collocation [6] and the Newton’s iteration [7], etc. However, the assumption fails in more realistic pursuit-evasion scenarios where the players have no or partial knowledge of their opponents’ strategies. [8] proposed to apply behaviour learning [9] to directly estimate the control input of the evader from the perspective of the pursuer and designed a control policy based on the estimated information. However, only a one-dimensional scenario was considered in the method.

The main contribution of this paper is an online game information estimation method in the context of three-dimensional orbital pursuit-evasion, which can overcome the performance loss due to incomplete information. More precisely, pursuit-evasion of two spacecraft/players is considered in the Hill reference frame. Each player has a cost function to be minimized regarding to its control inputs and distance to the other player with corresponding weighting matrices, which are not known by the other player. Thus, the orbital pursuit-evasion game has incomplete information. To find the optimal control strategy in this case, an online game information estimation method is proposed. From the perspective of the pursuer, by assuming an initial guess of the game information (weighting matrices of the evader), it computes the control law of itself and the evader, thus obtaining a predicted state of the game. With the game processing, the pursuer measures a new game state and updates its guess of the information. The estimation process is performed online using an Unscented Kalman Filter (UKF).

Incomplete-Information Orbital Pursuit-Evasion

Pursuit-Evasion with Complete Information

Orbital pursuit-evasion in the Hill reference frame is considered, as illustrated in Fig. 1. Let $\boldsymbol{X}_i = (x_i, y_i, z_i, \dot{x}_i, \dot{y}_i, \dot{z}_i)^T$ ($i = P, E$) be the state of the pursuer/evader in the Hill reference frame, then the motions of the two spacecraft are given by the CW equations [10]

$$\begin{cases} \ddot{x}_i - 2\omega\dot{y}_i - 3\omega^2 x_i = a_{i,x} \\ \ddot{y}_i + 2\omega\dot{x}_i = a_{i,y} \\ \ddot{z}_i + \omega^2 z_i = a_{i,z} \end{cases} \tag{1}$$

where $\omega = \sqrt{\mu / a_r^{\,3}}$, μ is the gravitational parameter, a_r is the radius of the reference orbit, $\boldsymbol{U}_i = (a_{i,x}, a_{i,y}, a_{i,z})^T$ is the control vector (acceleration) of the pursuer/evader. Eqn (1) can be written in the state space form

$$\boldsymbol{X}_i = \boldsymbol{A}\boldsymbol{X}_i + \boldsymbol{B}\boldsymbol{U}_i \tag{2}$$

where $\boldsymbol{A}$ and $\boldsymbol{B}$ are coefficient matrices [10]. Let $\boldsymbol{X} = \boldsymbol{X}_E - \boldsymbol{X}_P = (x, y, z, \dot{x}, \dot{y}, \dot{z})^T$ be the state of the game, i.e. relative state of the purser and evader, then we have

$$\dot{\boldsymbol{X}} = \boldsymbol{A}\boldsymbol{X} + \boldsymbol{B}\boldsymbol{U}_E - \boldsymbol{B}\boldsymbol{U}_P \tag{3}$$

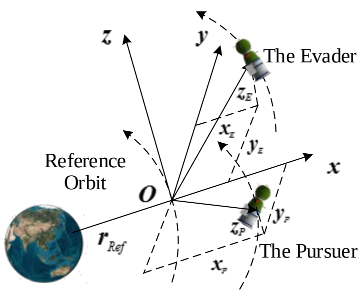


Fig. 1: Orbital pursuit-evasion in the Hill reference frame

The cost function of the pursuer to be minimized is

$$J_P = \int_0^\infty \left(\boldsymbol{X}^T(t)\boldsymbol{Q}\boldsymbol{X}(t) + \boldsymbol{U}_P^T(t)\boldsymbol{R}_P\boldsymbol{U}_P(t) - \boldsymbol{U}_E^T(t)\boldsymbol{R}_E\boldsymbol{U}_E(t) \right) dt \tag{4}$$

and the cost function of the evader is its opposite

$$J_E = -J_P \tag{5}$$

where $\boldsymbol{Q}$ is a positive semi-definite symmetric matrix, $\boldsymbol{R}_i$ ($i = P, E$) are positive definite symmetric matrices.

The optimal feedback control laws of the pursuer/evader are [11]

$$\boldsymbol{U}_P^*(t) = -\boldsymbol{K}_P\boldsymbol{X}(t) \tag{6}$$

$$\boldsymbol{U}_E^*(t) = -\boldsymbol{K}_E \boldsymbol{X}(t) \tag{7}$$

where the control gain matrices are

$$\boldsymbol{K}_P = -\boldsymbol{R}_P^{-1} \boldsymbol{B}^T \boldsymbol{P} \tag{8}$$

$$\boldsymbol{K}_E = -\boldsymbol{R}_E^{-1} \boldsymbol{B}^T \boldsymbol{P} \tag{9}$$

and $\boldsymbol{P}$ is the solution of the following algebra matrix Riccati equation

$$\boldsymbol{A}^T \boldsymbol{P} + \boldsymbol{P}\boldsymbol{A} - \boldsymbol{P}\boldsymbol{B}\boldsymbol{R}_P^{-1}\boldsymbol{B}^T \boldsymbol{P} + \boldsymbol{P}\boldsymbol{B}\boldsymbol{R}_E^{-1}\boldsymbol{B}^T \boldsymbol{P} + \boldsymbol{Q} = 0 \tag{10}$$

Pursuit-Evasion with Incomplete Information

In the complete-information game, both players have knowledge on that they are in a zero-sum game and know each other's cost function, i.e. the weighting matrices $\boldsymbol{Q}$, $\boldsymbol{R}_P$ and $\boldsymbol{R}_E$. Hence, they can compute its own and its opponent's control gains and strategies using Eqns (6)-(10). However, in realistic orbital pursuit-evasion, each player only has information on its own weighting matrices, indicating that the game has incomplete information. Let $\boldsymbol{Q}^P$, $\boldsymbol{R}_P^P$, $\boldsymbol{R}_P^P$ be the weighting matrices of the pursuer and $\boldsymbol{Q}^E$, $\boldsymbol{R}_P^E$, $\boldsymbol{R}_P^E$ are the evader's.

For each player, even it does not know the other player's weighting matrices, it can make an assumption that the other player is employing the same weighting matrices as itself. In other words, it assumes that the other player is playing a zero-sum game with it. Under this assumption, the player can compute a control gain based on its own weighting matrices and perform it during the game. We call such a control strategy as the "zero-sum strategy". We will show that such a zero-sum strategy could lead to performance loss of the players in incomplete-information pursuit-evasion.

Control Policy Design

Incomplete Information Online Estimation

From the perspective of the pursuer, its goal is to estimate the weighting matrices of the evader, so as to obtain the evader's control gain. Let $\hat{\boldsymbol{Q}}^E$, $\hat{\boldsymbol{R}}_P^E$, $\hat{\boldsymbol{R}}_E^E$ be the pursuer's estimation of the evader's weighting matrices, which are assumed in the following form

$$\hat{\boldsymbol{Q}}^E = \begin{bmatrix} q_r \boldsymbol{I}_{3\times3} & 0 \\ 0 & q_v \boldsymbol{I}_{3\times3} \end{bmatrix}, \quad \hat{\boldsymbol{R}}_P^E = \boldsymbol{I}_{3\times3}, \quad \hat{\boldsymbol{R}}_E^E = r\boldsymbol{I}_{3\times3} \tag{11}$$

Assuming the evader is performing a "zero-sum strategy", the pursuer can obtain a guess of the evader's control gain and control law as follows

$$\hat{\boldsymbol{K}}_E = -\hat{\boldsymbol{R}}_E^{E-1} \boldsymbol{B}^T \hat{\boldsymbol{P}} \tag{12}$$

$$\hat{\boldsymbol{U}}_E(t) = -\hat{\boldsymbol{K}}_E \boldsymbol{X}(t) \tag{13}$$

by solving the following algebra matrix Riccati equation

$$\boldsymbol{A}^T \hat{\boldsymbol{P}} + \hat{\boldsymbol{P}}\boldsymbol{A} - \hat{\boldsymbol{P}}\boldsymbol{B}\hat{\boldsymbol{R}}_P^{E-1}\boldsymbol{B}^T \hat{\boldsymbol{P}} + \hat{\boldsymbol{P}}\boldsymbol{B}\hat{\boldsymbol{R}}_E^{E-1}\boldsymbol{B}^T \hat{\boldsymbol{P}} + \hat{\boldsymbol{Q}} = 0 \tag{14}$$

where a hat $\wedge$ means the pursuer's estimation.

Substituting Eqn (13) to Eqn (3), the pursuer can formulate a control system described by

$$\dot{\boldsymbol{X}} = \boldsymbol{A}\boldsymbol{X} - \boldsymbol{B}\hat{\boldsymbol{K}}_E \boldsymbol{X} - \boldsymbol{B}\boldsymbol{U}_P \tag{15}$$

Let $\boldsymbol{Y} = (x, y, z, \dot{x}, \dot{y}, \dot{z}, q_r, q_v, r)^T$ be the augmented state vector of the game and substitute Eqns (12)-(14) to Eqn (15), then Eqn (15) can be written as

$$\dot{\boldsymbol{Y}}(t) = \boldsymbol{f}(\boldsymbol{Y}(t), \boldsymbol{U}_P(t), t) \tag{16}$$

Hence, for the pursuer, estimating the weighting matrices information, i.e. q_r, q_v and r, becomes to estimate the augmented game state $\boldsymbol{Y}$, which is governed by Eqn (16). Furthermore, the observation vector is the relative state of the two spacecraft, i.e. the observation model is

$$\boldsymbol{Z} = \boldsymbol{h}(\boldsymbol{Y}) = (x, y, z, \dot{x}, \dot{y}, \dot{z})^T \tag{17}$$

In this paper, we utilize UKF to perform the estimation.

Optimal Control Policy Design

According to Eqn (15), the pursuer formulates a control system based on its estimation of the evader's control gain. To this end, designing the control law of the purser is to solve a one-sided optimization problem for it, where its cost function is

$$J_P{}' = \min \int_0^\infty \left(\boldsymbol{X}^T(t)\boldsymbol{Q}^P\boldsymbol{X}(t) + \boldsymbol{U}_P^T(t)\boldsymbol{R}_P^P\boldsymbol{U}_P(t) \right) dt \tag{18}$$

Hence, according to the optimal control theory [12], the pursuer's optimal control law is

$$\boldsymbol{U}_P^*(t) = -\boldsymbol{K}_P\boldsymbol{X}(t) \tag{19}$$

where $\boldsymbol{K}_P$ is the solution of the following algebra matrix Riccati equation

$$(\boldsymbol{A} - \boldsymbol{B}\hat{\boldsymbol{K}}_E)^T\boldsymbol{P} + \boldsymbol{P}(\boldsymbol{A} - \boldsymbol{B}\hat{\boldsymbol{K}}_P) + \boldsymbol{P}\boldsymbol{B}\boldsymbol{R}_P^{P-1}\boldsymbol{B}^T\boldsymbol{P} + \boldsymbol{Q}^P = 0 \tag{20}$$

Simulation Results

Simulation Setup

The reference orbit is set as a near-Earth circular orbit with its altitude 400 km. The initial conditions of the pursuer and evader are set as $\boldsymbol{X}_P^0$ = (-6.0 km, -16.0 km, 4.0 km, -9.0 m/s, 13.6 m/s, 0 m/s) and $\boldsymbol{X}_E^0$ = (0 km, 0 km, 0 km, 0 m/s, 0 m/s, 0 m/s) respectively. Table 1 contains the weighting matrices of the pursuer and evader. This setup indicates that the evader concerns more on its relative state with the pursuer ($\|\boldsymbol{Q}^E\| > \|\boldsymbol{Q}^P\|$) while the concerns more on the fuel consumption ($\|\boldsymbol{R}_E^P\| > \|\boldsymbol{R}_E^E\|$). The evader always adopts a zero-sum strategy in the game, while for the pursuer we consider and compare three control strategies:

1) Complete information. The pursuer knows and adopts the same weighting matrices as the evader. This scenario is regarded as a baseline for comparison.
2) Incomplete information with zero-sum strategy. The pursuer assumes the evader has the same weighting matrices as it and adopts a zero-sum strategy.
3) Incomplete information with online information estimation, i.e. our proposed method.

Table 1: Weighting matrices of the pursuer and evader

Weighting matrices	$\boldsymbol{Q}$	$\boldsymbol{R}_P$	$\boldsymbol{R}_E$
Pursuer	$0.8\boldsymbol{I}_{6\times6}$	$\boldsymbol{I}_{3\times3}$	$\sqrt{2.4}\boldsymbol{I}_{3\times3}$
Evader	$\boldsymbol{I}_{6\times6}$	$\boldsymbol{I}_{3\times3}$	$\sqrt{2}\boldsymbol{I}_{3\times3}$

Results

Fig. 2 shows the time history of the pursuer's cost function when using the three different control strategies. As can be seen in Fig. 1, the final cost of the pursuer when employing a zero-sum strategy ($J^{f}_{P,zero} = 1.73\times10^{-5}$) is larger than that of the complete information case ($J^{f}_{P,complete} = 1.56\times10^{-5}$). This indicates that lack of game information, i.e. weighting matrices of the pursuer, leads the pursuer to a higher cost for interception. Nevertheless, thanks to online information estimation, the final cost ($J^{f}_{P,estimation} = 1.62\times10^{-5}$) is diminished compared with the zero-sum strategy case, albeit still larger than the complete information case. Hence, the proposed online game information estimation method is shown to be able to overcome the preformation loss of the pursuer due to incomplete information.

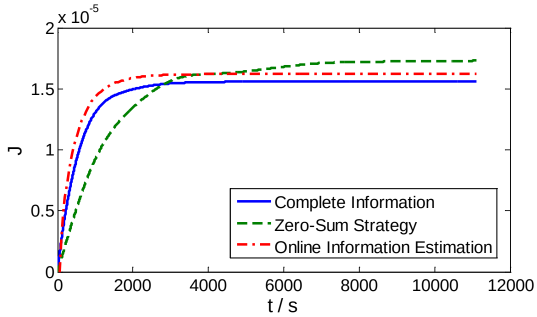


Fig. 2: Time history of the pursuer's cost function

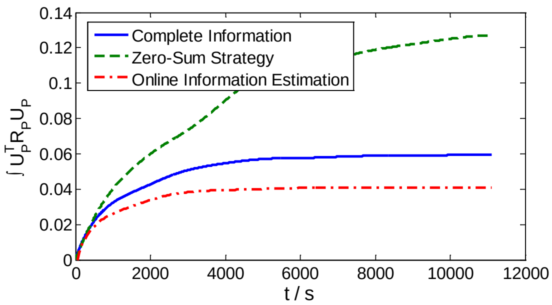
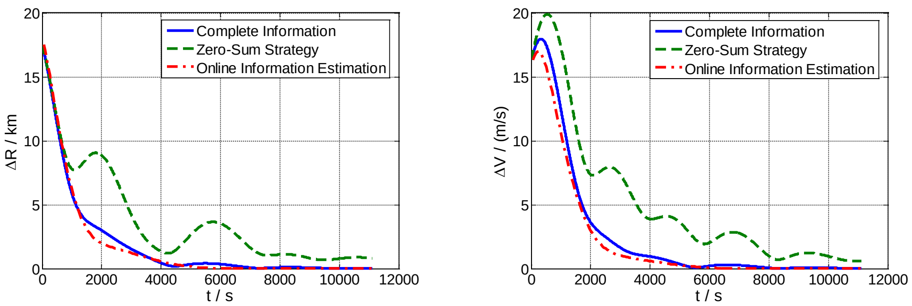


Fig. 3: Accumulated fuel consumption of the pursuer

The accumulated fuel consumption of the pursuer and the relative distance and velocity of the two spacecraft are shown in Fig. 3 and Fig. 4 respectively. According to the two figures, it can be analysed that with online information estimation, the pursuer's decreasing its relative distance and velocity to the evader is much faster than that of the zero-sum strategy, resulting a faster end of the game, i.e. interception with the evader. Hence, the accumulated fuel consumption is smaller than both the complete information case and zero-sum strategy case.



(a) Relative distance (b) Relative velocity

Fig. 4: Relative state of the pursuer and evader

Conclusion

In this paper, an online game information estimation method is developed for orbital pursuit-evasion with incomplete information in the Hill reference frame. Optimal control law is designed from the perspective of the pursuer based on its estimation of the evader's cost function and control gains. The method is shown in simulation to outperform the zero-sum strategy and can overcome the preformation loss of the pursuer in incomplete-information orbital pursuit-evasion. Future work shall develop a method to design optimal control laws for more general incomplete-information pursuit-evasion games, instead of only the weighting matrices information are not complete.

References

1. Basar, T. and Olsder G.J., *Dynamic Noncooperative Game Theory*, Second Edition, SIAM, Philadelphia, 1999.

2. Jagat, A. and Sinclair, A.J., “Nonlinear Control for Spacecraft Pursuit-Evasion Game Using the State-Dependent Riccati Equation Method”, *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 53, No. 6, 2017, pp. 3032-3042.

3. Shen, H. and Casalino, L., “Revisit of the Three-Dimensional Orbital Pursuit-Evasion Game”, *Journal of Guidance, Control, and Dynamics*, Vol. 41, No. 8, 2018, pp. 1823-1831.

4. Ehtamo, H. and Raivio, T., “On Applied Nonlinear and Bi-Level Programming for Pursuit-Evasion Games”, *Journal of Optimization Theory and Applications*, Vol. 108, No. 1, 2001, pp. 65-96.

5. Lachner, R., Michael, H.B. and Hans J.P, “Three-Dimensional Air Combat: Numerical Solution of Complex Differential Games”, *New Trends in Dynamic Games and Applications*, Birkhäuser Boston, 1995, pp. 165-190.

6. Pontani, M. and Conway, B.A., “Numerical Solution of the Three-Dimensional Orbital Pursuit–Evasion Game”, *Journal of Guidance Control and Dynamics*, Vol. 32, No. 2, 2009, pp. 474-487.

7. Li, Z.Y., Zhu, H., Yang, Z. and Luo, Y.Z., “A Dimension-Reduction Solution of Free-Time Differential Games for Spacecraft Pursuit-Evasion”, *Acta Astronautica*, 2019, accepted.

8. Cavalieri, K.A., Sataky, N. and Hurtado, J.E., “Incomplete Information Pursuit-Evasion Games with Uncertain Relative Dynamics”, *AIAA Guidance, Navigation, and Control Conference*, National Harbor, Maryland, 2014.

9. Satak, N. and Hurtado, J.E., “A Framework for Behavior Learning in Differential Games”, 52*nd Aerospace Sciences Meeting*, National Harbor, Maryland, 2014.

10. Vallado, D.A., *Fundamentals of Astrodynamics and Applications*, Fourth Edition, Microcosm Press, Portland, 2013.

11. Engwerda, J., *LQ Dynamic Optimization and Differential Games*, John Wiley & Sons, England, 2005.

12. Macki, J. and Aaron S., *Introduction to Optimal Control Theory*, Springer Science & Business Media, 2012.