SCIENCE CHINA
Information Sciences

• MOOP •

September 2024, Vol. 67, Iss. 9, 194201:1–194201:3
https://doi.org/10.1007/s11432-024-4134-1

A mixed Nash equilibrium solution for visibility-based pursuit-evasion game with multiple obstacles

Shaoming BU, Li LIANG* & Youqing WANG

College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029*, China*

Received 24 January 2024/Revised 3 April 2024/Accepted 3 June 2024/Published online 20 August 2024

This study primarily investigates a visibility-based pursuit-evasion game with multiple obstacles. The pursuer aims to maintain surveillance over the evader as long as possible, while the evader's objective is to escape the pursuer's surveillance as quickly as possible. We use the explicit strategy method to analyze the winning conditions and optimal strategy of the player in the case of corner obstacles and propose a 'Present-Future-Past' (PFP) loop structure to extend the strategy to the case of multiple obstacles. The action value functions are designed to evaluate various strategies of pursuer and evader, and the optimal strategy is determined through mixed Nash equilibrium theory. This method has high scalability and can be used to solve decision problems in multi-agent and multi-task scenarios.

The surveillance problem has attracted the attention of the robotics community due to its potential applications [1]. The visibility-based pursuit-evasion game is a significant tool for solving surveillance problems, in which the pursuer robot aims at maintaining surveillance on the evader robot, while the evader robot attempts to escape the surveillance by using obstacles. In relevant research, Murrieta et al. [2] introduced the concepts of strong mutual visibility and accessibility, and the results showed that the problem of whether the pursuer and evader can maintain strong mutual visibility is a completely Nondeterministic Polynomial problem. Bhattacharya et al. [3] provided optimal strategies in a single corner and proposed the concept of star-shaped regions. Zou et al. [4] proposed a general framework for this problem and made improvements. Researchers have incorporated the concept of iterated games into multiplayer games and demonstrated the convergence of multiplayer systems, offering new insights for research in this field [5]. However, there is little detailed research on solving surveillance problems in arbitrary polygonal multi-obstacle environments.

This study addresses a visibility-based pursuit-evasion game with multiple obstacles, and the primary contributions are categorized into three key aspects.

(1) In the case of a corner, we provide an analytical expression for the barrier and divide the workspace into winning regions for the evader and pursuer respectively.

(2) We introduce a PFP structure that accounts for the influence of future and past events on current decision-making. We transform the visibility-based pursuit-evasion game into an optimization problem for selecting the optimal strategy, thereby reducing the complexity of the problem.

(3) We incorporate mixed Nash equilibrium into the visibility-based pursuit-evasion game and design a action value function to solve the Nash equilibrium strategy in the context of multi-obstacle scenarios.

Problem formulation. As shown in Figure 1, the pursuer and evader move in a two-dimensional environment with obstacles, indicated by the brown-colored areas. The purpose of the pursuer is to monitor the evader, while the evader attempts to escape the pursuer's surveillance by using the obstacles. Establishing a coordinate system with A as the origin, the dynamics of the system are described as

$$\begin{cases} \dot{r_i} = -V_i \cos\theta_i, \\ \dot{\phi_i} = \dfrac{V_i \sin\theta_i}{r_i}, \end{cases} (i \triangleq P, E), \tag{1}$$

where V is the speed, θ is the angle between the velocity and the polar radius, with the positive direction being counterclockwise. We use P to represent the pursuer, whose polar coordinates are represented by (r_P,ϕ_P), and use E to represent the evader, whose polar coordinates are represented by (r_E,ϕ_E). The ray representing the negative direction of the x-axis is e_1, and the ray used to represent the visible area boundary of the pursuer is e_2. The angle of the corner A is β_A and the angle of the corner B is β_B. When the evader enters the blind area of visual field, presented by the gray area, it will escape the pursuer's surveillance, which is formed by the relative position of the corner and the pursuer. We define the Star Region of corner A as η_A, shown by the light blue-colored area, consisting of the reverse extension line of corner A. When the pursuer enters η_A, it can see both sides of A, and the evader cannot use this corner to evade the surveillance of the pursuer. In an open environment, the shortest distance between two points is $d(X, Y) = \|X-Y\|_2$, and the shortest distance from a point to a segment, ray, line or region is $d(X, e) = \min_{Y\in e} \|X - Y\|_2$, where e represents the segment, ray, line or region. In instances where an obstacle hinders the direct path between points X and e, the expression $d(X, e)$ denotes the minimal rectilinear distance between X and e. This distance is defined by the shortest polyline that avoids traversing through obstacles.

* Corresponding author (email: lianglibuct@buct.edu.cn)

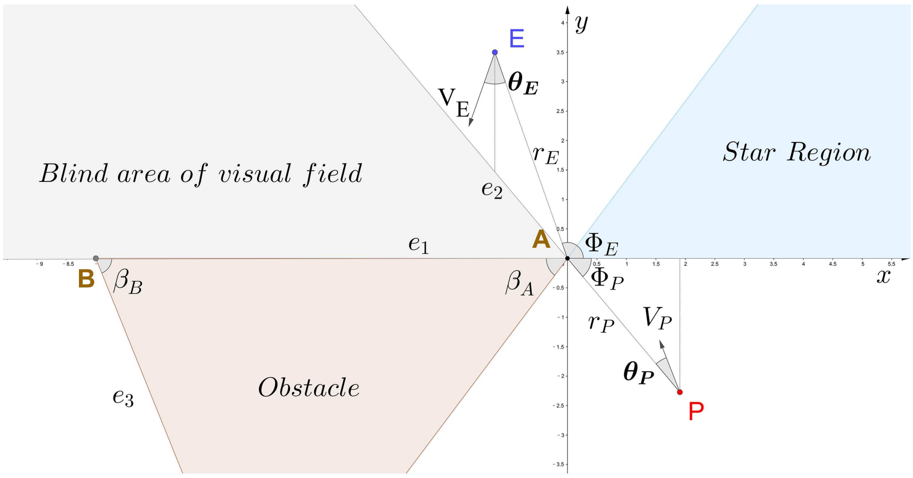

Figure 1 (Color online) A schematic diagram of the visibility-based pursuit-evasion game.

The PFP loop structure. We first use geometric analysis to develop the winning conditions and strategies for both pursuers and evaders in the case of corner obstacles, with $d_1 = |\frac{V_E}{V_P} r_P \sin\phi_P|$ and $d_2 = |\frac{V_E}{V_P} r_P|$. The evader winning region for the corner obstacles is described as $d(E, e_1) < d_1, \phi_P \in [-\frac{\pi}{2}, 0]$; $d(E, e_2) < d_2, \phi_P \in [\beta_A - \pi, -\frac{\pi}{2}]$.

When $\phi_P \in [-\frac{\pi}{2}, 0]$, the strategy of the evader θ_E can be described as $\phi_E - \frac{\pi}{2}$ if $d(E, e_1) < d_1$, $\phi_E \in [\frac{\pi}{2}, \pi + \phi_P]$; 0 if $d(E, e_1) < d_1$, $\phi_E \in [\frac{\pi}{2} + \phi_P, \frac{\pi}{2}]$ or $\phi_E \in [\beta_A - \pi, \frac{\pi}{2} + \phi_P]$; or $\phi_E - \phi_P - \frac{\pi}{2}$ if $d(E, e_1) \geqslant d_1$, $\phi_E \in [\frac{\pi}{2} + \phi_P, \pi + \phi_P]$. The strategy of the pursuer θ_P can be described as $\phi_P + \frac{\pi}{2}$ if $d(E, e_2) \geqslant d_1$ or $d(E, A) < d_2$; or $\arcsin(\frac{V_E r_P \sin\theta_E}{V_P r_E})$ if $d(E, e_2) < d_1$, $d(E, A) \geqslant d_2$.

When $\phi_P \in [\beta_A - \pi, -\frac{\pi}{2}]$, the strategy of the evader θ_E can be described as 0 if $d(E, A) < d_2$ or $\phi_E \in [\beta_A - \pi, \frac{\pi}{2} + \phi_P]$; or $\phi_E - \phi_P - \frac{\pi}{2}$ if $d(E, A) \geqslant d_2$, $\phi_E \in [\frac{\pi}{2} + \phi_P, \pi + \phi_P]$. The strategy of the pursuer θ_P can be described as 0 if $d(E, A) < d_2$ or $d(E, e_2) \geqslant d_2$; or $\arcsin(\frac{V_E r_P \sin\theta_E}{V_P r_E})$ if $d(E, A) \geqslant d_2$, $d(E, e_2) < d_2$.

Polygon obstacle can be represented as a combination of multiple corner obstacles. In a polygon obstacle scenario, if the evader cannot avoid surveillance using the current corner (the 'Present' corner), it can start preparing for the next corner, which is named the 'Future' corner. When the pursuer enters the Star Region of the 'Present' corner, it will be transformed into a 'Past' corner, while the previous 'Future' corner becomes the new 'Present' corner. This creates a cyclic structure of PFP. The evader can use the 'Present' and 'Future' corner obstacles to avoid surveillance of the pursuer. 'Present' corner can directly use the strategies mentioned above, we only discuss the strategies of the evader and pursuer for the 'Future' corner. Let 'Present' corner be denoted as A and 'Future' corner as B, with a ray designated as e_3 originating from B along another edge. Due to B's role in preventive measures, the objectives for both evader and pursuer are to enter or prevent entry into the evader winning region formed by B, which can be expressed as $d(E, e_3) < \left|\frac{V_E}{V_P}\right| d(P, \eta_B)$, before B is transformed into the 'Present' corner. The specific strategy for B ('Future' corner) can be articulated through by the following equations, with

$$r_0 = -\frac{d(A, B)\cos(\beta_B)}{\cos(\beta_B + \phi_E)} \tag{2}$$

and

$$\theta_0 = \arccos\left(\frac{r_E^2 + d(E, B)^2 - d(A, B)^2}{2d(E, B) r_E}\right). \tag{3}$$

When $\beta_B \in [\frac{\pi}{2}, \pi]$ and $\phi_P \in [\frac{\pi}{2} - \beta_B, 0]$, the pursuer's strategy can be represented as $\theta_P = \beta_B + \phi_P - \frac{\pi}{2}$, and the evader's strategy θ_E can be represented as follows: 0 if $\phi_E \in [\beta_A - \pi, 0]$; θ_0 if $\phi_E \in [0, \frac{3\pi}{2} - \beta_B]$ or $r_E < r_0$, $\phi_E \in [\frac{3\pi}{2} - \beta_B, \pi + \phi_P]$; or $\beta_B + \phi_E + \frac{\pi}{2}$ if $r_E \geqslant r_0$, $\phi_E \in [\frac{3\pi}{2} - \beta_B, \pi + \phi_P]$.

When $\beta_B \notin [\frac{\pi}{2}, \pi]$ or $\phi_P \notin [\frac{\pi}{2} - \beta_B, 0]$, the pursuer's strategy can be represented as $\theta_P = 0$, and the evader's strategy θ_E can be represented as follows: 0 if $\phi_E \in [\beta_A - \pi, 0]$; or θ_0 if $\phi_E \in [0, \pi + \phi_P]$.

Strategy balance with multiple obstacles. In a multi-obstacle environment, both pursuer and evader can form N strategies, where N is less than or equal to the number of corners. 'PFP' structure can also be used to construct each corner as 'Present', 'Future', or 'Past'. It needs to be emphasized that when considering the 'Future' corner, the influence of the 'Present' corner on the critical point of visibility needs to be taken into account.

Thus, we have transformed the visibility-based pursuit-evasion game into an optimization problem. Given the rationality of both the pursuer and the evader, their strategies adapt in response to each other's actions, making the identification of Nash equilibrium crucial in this study. Leveraging the 'PFP' framework, we convert this game into a finite game, where each participant is associated with a finite set of pure strategies. In non-cooperative games, the existence of Nash equilibrium is not guaranteed for finite games with pure strategies. However, for mixed games, it is established that every finite game has an equilibrium point [6]. The mixed Nash equilibrium yields mixed strategies, where players choose different pure strategies with specific probability distributions when making decisions, thereby preventing any player from unilaterally improving their situation by changing their strategy combination. The introduction of mixed strategies allows game theory to better describe situations in the real world, as in practical scenarios, players may make random choices based on various factors rather than always sticking to a fixed, determinate strategy.

We use mixed Nash equilibrium theory to present the balance strategy for both pursuer and evader. We construct the interaction between both parties as a zero-sum game and formulate the action value function as

$$\begin{cases} f_E(\alpha, F) = \alpha \left(\dfrac{d(F,P)}{d(F,E)}\right)^{n\cdot\mathrm{sgn}(\alpha)}, \\ f_P(\alpha, F) = -f_E(\alpha, F), \end{cases} \tag{4}$$

where F represents the corner which the evader or the pursuer relies, and n is a hyperparameter with $n > 0$. The values of α is obtained through geometric analysis for regional division and it can be delineated in a unified form. For the 'Present' corner,

$$\alpha = \frac{V_E \sin(\phi_E - \theta_E)}{r_E \sin\phi_E} - \frac{V_P \sin(\phi_P - \theta_P)}{r_P \sin\phi_P}. \tag{5}$$

For the 'Future' corner,

$$\alpha = \frac{V_E}{d(E, e_3)} - \frac{V_P}{d(P, \eta_B)}. \tag{6}$$

The key point to note here is that the '$d(\cdot)$' is not the shortest straight-line distance, but rather the distance defined by the shortest polyline that avoids traversing through obstacles.

As mentioned earlier, the strategy sets for both pursuer and evader consist of N strategies each, resulting in $N \times N$ possible action scenarios. By evaluating each action scenario with the action value function, we can construct an $N \times N$ matrix, represented as:

$$A_{f_E} = \begin{matrix} \\ E_1 \\ E_2 \\ \vdots \\ E_N \end{matrix} \begin{matrix} \begin{matrix} P_1 & P_2 & \cdots & P_N \end{matrix} \\ \begin{bmatrix} f_E(\alpha_{11}) & f_E(\alpha_{12}) & \cdots & f_E(\alpha_{1N}) \\ f_E(\alpha_{21}) & f_E(\alpha_{22}) & \cdots & f_E(\alpha_{2N}) \\ \vdots & \vdots & \ddots & \vdots \\ f_E(\alpha_{N1}) & f_E(\alpha_{N2}) & \cdots & f_E(\alpha_{NN}) \end{bmatrix} \end{matrix}. \tag{7}$$

The variables $E_1, E_2, \cdots, E_N$ and $P_1, P_2, \cdots, P_N$ represent the strategies of the evader and pursuer formed by N corners. The expression $f_E(\alpha_{XY})$ signifies the value for the evader when the evader adopts the strategy of corner X and the pursuer adopts the strategy of corner Y, depicting the outcome of this particular action scenario. It is important to note that when both pursuer and evader formulate strategies based on different corners, two action value functions, $f_E(\alpha, X)$ and $f_E(\alpha, Y)$, are formed. We choose the one most favorable to the evader, denoted as $f_E(\alpha_{XY}) = max(f_E(\alpha, X), f_E(\alpha, Y))$.

Utilizing the theory of mixed Nash equilibrium, the mixed strategy for the evader is determined as $p_E = (p_{E1}, p_{E2}, \cdots, p_{EN})$, where p_{Ej} denotes the probability that evader executing the strategy E_j. Consequently, the expected payoff for the evader, denoted as $U_E(p) = p_E A_{f_E} p_P^T$, is characterized by the inequality $U_E(p_E, p_P) \geqslant U_E(p'_E, p_P)$, where p'_E represents any probability distribution over the N strategies for the evader. For the pursuer, $U_P(p) = -U_E(p)$, where $U_P(p)$ satisfies $U_P(p_E, p_P) \geqslant U_P(p_E, p'_P)$. Therefore, we obtain the optimal mixed strategy p_E and p_P.

Subsequently, the evader and pursuer each choose among their N strategies with probabilities p_E and p_P, respectively. It is worth noting that after both the pursuer and evader execute a strategy, a re-analysis and game take place. As time progresses, the states of the pursuer and evader change. Since both parties continuously adjust their strategies based on the actions of the other, this process occurs continuously. Through multiple instances of static games, the aim is to address dynamic game problems.

Conclusion. We provided a solution for visibility-based pursuit-evasion games with multiple obstacles. By employing explicit strategy analysis, we extended the corner model to a multi-obstacle case. The evaluation of several strategies for both pursuer and evader was conducted using the action value function. We identified the optimal strategies through the application of mixed Nash equilibrium theory. The proposed method has high scalability and is suitable for decision-making problems in multi-task scenarios involving agents. In the future, we intend to employ Q-learning as an alternative approach to the current action value function, addressing disturbances introduced by additional variables.

Acknowledgements This work was supported in part by National Key Research and Development Program of China (Grant No. 2022ZD0119701), in part by National Natural Science Foundation of China (Grant No. 62103045), and in part by Fundamental Research Funds for the Central Universities (Grant No. buctrc202228).

Supporting information Videos and other supplemental documents. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Xi L, Peng Z, Jiao L, et al. Smooth quadrotor trajectory generation for tracking a moving target in cluttered environments. Sci China Inf Sci, 2021, 64: 172209

2 Murrieta-Cid R, Monroy R, Hutchinson S, et al. A complexity result for the pursuit-evasion game of maintaining visibility of a moving evader. In: Proceedings of 2008 IEEE International Conference on Robotics and Automation, Pasadena, 2008. 2657–2664

3 Bhattacharya S, Hutchinson S. A cell decomposition approach to visibility-based pursuit evasion among obstacles. Int J Robot Res, 2011, 30: 1709–1727

4 Zou R, Bhattacharya S. On optimal pursuit trajectories for visibility-based target-tracking game. IEEE Trans Robot, 2019, 35: 449–465

5 Zhu Y, Xia C, Chen Z. Nash equilibrium in iterated multiplayer games under asynchronous best-response dynamics. IEEE Trans Automat Control, 2023, 68: 5798–5805

6 Nash J. Non-cooperative games. Annal Math, 1951, 54: 286–295