

Article

Maneuver Planning for Multiple Pursuit Intelligent Surface Vehicles in a Sequence of Zero-Sum Pursuit–Evasion Games

Le Hong ^{1,2,3} , Weicheng Cui ^{2,3,*} , Hao Chen ^{2,3} , Changhui Song ^{2,3} and Weikun Li ^{2,3}

¹ Zhejiang University–Westlake University Joint Training, Zhejiang University, Hangzhou 310024, China; hongle@westlake.edu.cn

² Key Laboratory of Coastal Environment and Resources of Zhejiang Province, School of Engineering, Westlake University, 18 Shilongshan Road, Hangzhou 310024, China; chen hao@westlake.edu.cn (H.C.); songchanghui@westlake.edu.cn (C.S.); liweikun@westlake.edu.cn (W.L.)

³ Institute of Advanced Technology, Westlake Institute for Advanced Study, 18 Shilongshan Road, Hangzhou 310024, China

* Correspondence: cuiweicheng@westlake.edu.cn

Abstract: Unmanned surface pursuit is a complex and challenging engineering problem, especially when conducted by multiple intelligent surface vehicles (ISVs). To enhance the pursuit performance and facilitate strategic interaction during the target pursuit, this paper proposes a novel game theory-based maneuver planning method for pursuit ISVs. Firstly, a specific two-player zero-sum pursuit–evasion game (ZSPEG)-based target-pursuit model is formed. To ensure the vehicles reach a quick consensus, a target-guided relay-pursuit mechanism and the corresponding pursuit payoffs are designed. Meanwhile, under the fictitious play framework, the behavioral pattern and the strategies of the target could be fictitiously learned. Furthermore, mixed-strategy Nash equilibrium (MNE) is employed to determine the motions for the vehicles, the value of which is the best response in the proposed ZSPEG model. Finally, simulations verify the effectiveness of the above methods in multi-ISV surface pursuit.

Keywords: unmanned surface pursuit; maneuver planning; multiple intelligent surface vehicles; zero-sum pursuit–evasion game; fictitious play; mixed-strategy Nash equilibrium



Citation: Hong, L.; Cui, W.; Chen, H.; Song, C.; Li, W. Maneuver Planning for Multiple Pursuit Intelligent Surface Vehicles in a Sequence of Zero-Sum Pursuit–Evasion Games. *J. Mar. Sci. Eng.* **2024**, *12*, 1221. <https://doi.org/10.3390/jmse12071221>

Academic Editor: Marco Cococcioni

Received: 9 June 2024

Revised: 3 July 2024

Accepted: 19 July 2024

Published: 20 July 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

From conceptualization to realization, the utilization of multiple intelligent surface vehicles (ISVs) to conduct surface pursuits has attracted widespread attention [1–5]. However, how to effectively and efficiently pursue a moving target still poses a significant challenge for a multi-ISV system. The system necessitates numerous advanced technologies, including target recognition, target positioning, and dynamic maneuver planning [6,7]. Focusing on providing an efficient maneuver planner for the multi-ISV pursuit system, this paper aims to address the concerns regarding realizing the strategic interaction and addressing the data scarcity of the target.

There are three mainstream methods in maneuver planning for a multi-vehicle target pursuit: graph theory-based algorithms, the cooperative target hunting approach, and game theory-based methods. Based on graph theory, the prevailing approach is the Voronoi-related algorithm, which could visually represent the relationships among the motion sets for the vehicles [8,9]. However, all the graph theory-based algorithms do not directly provide information regarding strategic choices and the interaction between the pursuit vehicles and the target. As for the cooperative target hunting methods, researchers have focused on the formation control and dynamic task allocation during the planning process [10–12]. Nevertheless, these methods require detailed models for both the vehicles and the environmental information. Game theory, as a mature mathematical theory, has broad applicability and theoretical rigor in maneuver planning [13]. It offers advantages

such as equilibrium point analysis, strategy optimization, and the handling of incomplete information. To realize intelligent and cognitive interactions between the pursuit team and the target, this paper focuses on developing a novel and practical game theory-based maneuvering planning method.

Under the game theory-based target-pursuit framework, three types of game models, Stackelberg games [14], cooperative games [15], and zero-sum games (ZSGs) [16–19], demonstrate different characteristics. The Stackelberg game involves a leader–follower model, where the leader takes action first and the follower adjusts their strategy after observation [20]. Hence, the Stackelberg game-based model is suitable for simulating the dynamics in a leader–follower structure. However, the complexity of the Stackelberg game is high, requiring sophisticated real-time decision-making. Cooperative games emphasize teamwork [21]. In cooperative games, team members depend on each other. The action of one member may affect the entire team [22]. The interdependence of the members may result in a decrease in team effectiveness when one team member fails to fulfill their responsibilities. Compared with the two previously mentioned games, ZSGs are more applicable to intense competition [23]. In ZSGs, the success of one team inevitably leads to the failure of other teams. ZSGs enable participants to adopt clear confrontational strategies. Moreover, due to the relative simplicity of the ZSG models, decision-making is more explicit and easier to implement, contributing to more effective target pursuit in competitive environments.

The application of ZSGs to target pursuit manifests as zero-sum pursuit–evasion games (ZSPEG). The ZSPEG-based model comprises two players: the pursuer and the evader [16–19]. In refining the ZSPEG-based model to be more specific and tailored, researchers focus on enriching its objective function and considering various involved factors. A divide-and-conquer approach is used for a multi-player ZSPEG where the pursuers have a twofold goal [11]. To identify unknown sets of incoming attackers, a multi-model adaptive estimator is implemented in the offline design policy sets [24]. Furthermore, a bilateral adaptive parameter estimation method is adopted to deal with a multi-agent ZSPEG model with unknown general quadratic goals [25]. In recent years, various algorithms have been developed to solve ZSG-based models, such as regret matching [26], fictitious play [27], double oracle [28], etc. Among them, the most popular algorithms are regret learning-based methods. They rely on the concepts of external regret, internal regret, swap regret, and Nash equilibrium-based regret [29]. Building on this foundation, the current mainstream algorithms are the optimistic follow-the-regularized-leader [30] and the optimistic mirror descent algorithm [31]. Despite the extensive research on ZSPEG-based methods and the relevant algorithms, there is little research on applying the ZSPEG to multi-ISV target pursuit. The main reason for this is the high complexity, data scarcity, and many technical challenges in multi-ISV target pursuit. In this context, comparing other algorithms to the ZSG-based models, the fictitious play algorithm can provide more effective assistance. Firstly, fictitious play could simulate the behavior of the target, which could help the pursuit system to better understand the behavioral patterns of the target, thereby addressing the data scarcity of the target. Additionally, the fictitious play algorithm allows for the experiment to be validated and for an easier evaluation of the feasibility of different strategies and algorithms. This characteristic could prove advantageous for multi-ISV systems by mitigating the complexity during the surface pursuit. Thus, fictitious play is adopted as the solution framework for our developed ZSPEG-based, multi-ISV, target-pursuit model.

Building upon the above discussions, this paper formulates an innovative ZSPEG-based maneuver planning method for multiple-pursuit ISVs. The main contributions are summarized as follows:

- i By employing the ZSPEG framework, the surface pursuit system gains the ability to explore and analyze the intense competition between the pursuit team and the target.
- ii Through the utilization of fictitious play, the data scarcity concerning the target can be alleviated, thereby reducing the complexity associated with surface pursuit.

- iii Under the fictitious play framework, a mixed-strategy Nash equilibrium (MNE)-based decision-making process is employed. This approach could derive the best responses to maneuver the surface vehicles effectively and stably.

The subsequent sections of this paper follow a structured organization. Section 2 provides the details of the multi-ISV target-pursuit model, which is based on the ZSPEG framework. Following this, Section 3 introduces the design of a motion planner for the pursuit vehicles, utilizing the MNE methodology. Section 4 shows the execution of the simulations. Finally, Section 5 presents the conclusions drawn in the study.

2. System Model

ZSG means that, under strict competition conditions, one player gains benefits while another player suffers loss. The sum of their benefits is zero [32]. The development of a ZSPEG model for a multi-ISV pursuit system includes three foundational elements in the ZSG-based model: the two players, their strategy sets, and the payoff function.

2.1. Game Scenario and the Players

A classic ZSG model can be formulated as follows:

$$G = \{S_1, S_2, F\}, \tag{1}$$

where S_1 denotes the strategy set for player 1. It is assumed that $S_1 = \{a_1, a_2, \dots, a_m\}$. m is the positive integer. Similarly, S_2 denotes the strategy set for player 2, which is assumed to be $S_2 = \{b_1, b_2, \dots, b_n\}$. F represents the payoff matrix for player 1. As for the state in the ZSG, when $i \in \{1, 2, \dots, m\}$ and $j \in \{1, 2, \dots, n\}$, if player 1 adopts strategy a_i and player 2 adopts strategy b_j , the game state could be formulated by (a_i, b_j) . Under this game state,

the payoff for player 1 is assumed to be f_{ij} . Then, the payoff matrix $F = \begin{bmatrix} f_{11} & \dots & f_{1n} \\ \vdots & \ddots & \vdots \\ f_{m1} & \dots & f_{mn} \end{bmatrix}$

for player 1 is obtained. Different rows in F denote the different strategies of player 1, while different columns in F denote the different strategies chosen by player 2.

The pursuit scenario with n ISVs pursuing a mobile target is presented in Figure 1. In accordance with this scenario, the constructed ZSSG model includes two players: the pursuit team, consisting of n vehicles, and the moving target. Then, the ZSSG-based model for the multi-ISV target-pursuit scenario could be formulated by $G = \{S_p, S_e, F\}$, where S_p and S_e denote the strategy sets for the two players. It should be noted that the red star in the figure represents the attached points of the moving target. Once the target reaches this point, the pursuit team fails the protection task. F represents the payoff matrix for the pursuit team.

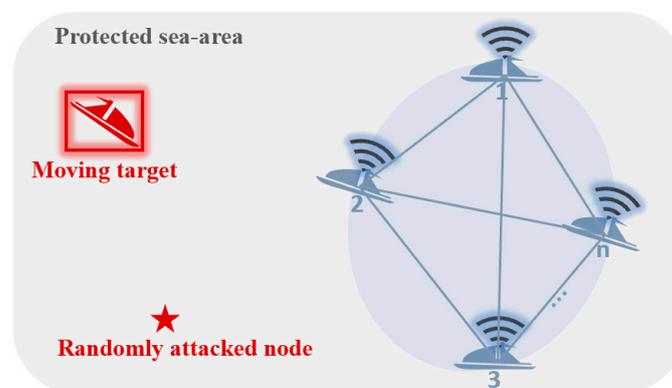


Figure 1. Schematic of the surface pursuit.

2.2. Strategy Set

Before designing the maneuver strategy sets, the kinematic model of the surface vehicle should be formulated. When the vehicle moves on the surface, three primary types of motion on the horizontal plane are typically identified: surge, sway, and yaw. The corresponding motion analysis is presented in Figure 2. The motions of the vehicle are described using two reference coordinates: the geodetic reference coordinate Σ_I and the vehicle body reference coordinate Σ_B [33]. As shown in Figure 2, when $i \in \{1, 2, \dots, n\}$, under Σ_I , the position of the i th vehicle is $x_i = [x_i, y_i]^T$. u_i is the speed at which the vehicle oscillates left and right along Y_B , which is usually set to be zero. v_i is the speed at which the boat oscillates forward and backward along X_B . ϕ_i is the yaw angle of the vehicle. v_c is the velocity of the currents at Σ_I .

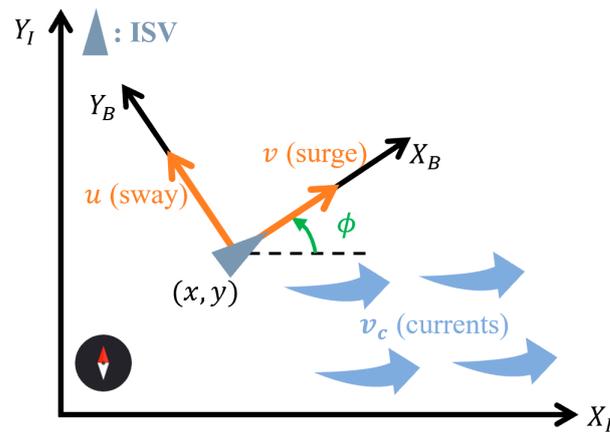


Figure 2. Top-view of the ISV motion.

Assumption 1. It is assumed that all the vehicles have a constant-speed motion, and the sway velocity for the individual vehicle is zero.

Under Assumption 1 and the discrete maneuver, the three-degrees-of-freedom kinematics model of the i th vehicle is shown as follows:

$$\begin{cases} x_i(k+1) = x_i(k) + (v_i(k)r_i(k) + v_c)\Delta t \\ r_i(k) = [\cos \phi_i(k), \sin \phi_i(k)]^T \end{cases}, \quad (2)$$

where the control variable is the unit vector of yaw angle $r_i(k)$ at the current k th step, where k is a non-negative integer. $x_i(k+1)$ represents the position of the vehicle at the next time step. $x_i(k)$ represents the position at the current k th step. Δt is the unit time period. $v_i(k)r_i(k)$ is a standard vector–scalar multiplication, where scalar v_i scales the components of vector r_i . This results in a new vector $(v_i \cos \phi_i, v_i \sin \phi_i)$, which represents the velocity of the vehicle in Cartesian coordinates.

Based on the above introduction, $r_i(k)$ is what should be determined for the pursuit vehicle in the ZSSG-based model. Then, if there are N yaw vectors for the i th vehicle to choose, there would be N^n choices in the S_p for the pursuit team. Therefore, to reduce the computational overhead, one decomposing method is adopted in this paper. The decomposing method that is employed is named the target-guided relay-pursuit (TGRP) strategy, which is elaborated in Definition 1. This strategy hinges on the principle of relay pursuit, where only one vehicle is active at any given time, while the others remain on standby, ready to take over the pursuit as the scenario evolves. There are two reasons for adopting the TGRP strategy. The core reason behind using TGRP is to dynamically assign the pursuit role to the ISV that is best positioned to continue the chase, minimizing the overall time to intercept and enhancing the strategic positioning of the pursuit team. Furthermore, if fewer ISVs are actively maneuvering at any one time, the risk of collisions is significantly decreased, enhancing safety and operational integrity. However, while

the TGRP strategy is designed to be highly effective under a wide range of conditions, it is necessary to clarify that it may not always represent the absolute optimal strategy. Its performance can be considered close to optimal in scenarios characterized by complex dynamics and the need for rapid tactical shifts.

Definition 1 (TGRP strategy [34]). *In the TGRP strategy, only one vehicle is active, while the others are stationary. This relay-pursuit mechanism could effectively avoid collisions among the vehicles. Furthermore, the distribution of the active vehicle changes over time, which depends on the outcome of the game at each time step. If the active vehicle is determined by index i^* , the heading vector r_i for the i th pursuit ISV could be obtained by Equation (3).*

$$r_i = \begin{cases} \frac{a_i}{\|a_i\|}, & \text{if } i = i^* \\ \mathbf{0} \parallel a_i, & \text{others} \end{cases} \tag{3}$$

where $a_i := x_t - x_i$ is the relative position vector from the target to the i th vehicle. $\mathbf{0} \parallel a_i$ denotes the heading angle for the inactive vehicle towards the direction of the target.

Under the TGRP strategy, the pursuit team has n total motions. When $i \in \{1, \dots, n\}$, the i th motion denotes that the i th vehicle actively pursues the target. By utilizing the fictitious play, the restricted strategy set S_t for the target is assumed to include $n + 1$ choices. The former n actions represent an evasion of the corresponding pursuers. The $n + 1$ action means that the target is moving towards its target node x_n . Then, when $j \in \{1, \dots, n + 1\}$, r_{t_j} denotes the j th maneuver strategy for the moving target. The payoff matrix F for the pursuit team is demonstrated in Table 1.

Table 1. Formulation of the payoff matrix.

Payoff	Strategy of the Invading Vehicle: S_e			
	Evading P_1	Evading P_j	Evading P_n	Towards x_a
Active P_1	f_{11}	f_{1j}	f_{1n}	$f_{1(n+1)}$
Active P_i	f_{i1}	f_{ij}	f_{in}	$f_{i(n+1)}$
Active P_n	f_{n1}	f_{nj}	f_{nn}	$f_{n(n+1)}$

2.3. Pursuit Payoff Function

In this paper, two goals are set for each pursuit vehicle. The first goal is to capture the target as quickly as possible. The second is to prevent the target from reaching its intended node. The two goals are the minimum time required for the vehicle to capture the target, and the alignment of the target with the point of attack from its current position, respectively.

2.3.1. Minimum Time-to-Capture

To calculate the minimum time-to-capture, several parameters need to be introduced. x_t , v_t , and r_t represent the location, the surge velocity, and the unit yaw vector for the target, respectively. l is the length of the vehicle.

Then, a time metric is introduced [23]. It is associated with the minimum positive solution, $t(x_t, x_i, r_{t_j}, r_i)$, in the following equation:

$$(v_t^2 - v_i^2)t^2 + 2[v_t r_{t_j}^T a_i - l v_{r_i}]t + \|a_i\|^2 - l^2 = 0, \tag{4}$$

where $a_i := x_t - x_i$. As the time-of-capture $(T_C)_{ij}$ is bounded by the T_U , the final $(T_C)_{ij}$ can be obtained as follows:

$$(T_C)_{i,j} = \min\{t(x_t, x_i, r_{t_j}, r_i), T_U\}. \tag{5}$$

2.3.2. Avoidance of Being Attacked

There is a goal to avoid the target reaching the point of attack: x_a , $(A_N)_{:,j}$ denotes the extent to which target is heading towards x_a from its current location. Therefore, $(A_N)_{:,j}$ is assumed to be $\cos \theta$, where θ is the angle between r_{t_j} and vectors $x_a - x_t$. Therefore, $(A_N)_{:,j}$ is formed as follows:

$$(A_N)_{:,j} = \cos \left(\frac{\langle r_{t_j}, x_a - x_t \rangle}{\|r_{t_j}, x_a - x_t\|} \right), \tag{6}$$

where $(A_N)_{:,j}$ is the alignment metric for the target with respect to its attack node under the j th strategy, and $\cos \left(\frac{\langle r_{t_j}, x_a - x_t \rangle}{\|r_{t_j}, x_a - x_t\|} \right)$ represents the cosine of the angle between the target's heading vector r_{t_j} and the vector pointing from the target to the attack node $x_a - x_t$. This measures how well-aligned the target's heading direction is with the vector pointing towards the attack node. According to Equation (6), a higher value of $(A_N)_{:,j}$ (close to 1) indicates that the target is moving directly towards the attack node, suggesting a higher likelihood of an attack. A lower value (close to -1) suggests the target is moving directly away from the attack node. Values of around 0 indicate perpendicular movement, implying that the target is neither approaching nor directly retreating from the attack node. This metric helps to evaluate the level of threat posed by the target to the attack node. By continuously monitoring $(A_N)_{:,j}$, the pursuit system can prioritize targets that are better aligned with critical points, thus optimizing the pursuit strategies.

2.3.3. Goal/Payoff Function

To balance multiple critical aspects of the pursuit dynamics, the goal/payoff function (7) is formed based on the two introduced metrics: the minimum time-to-capture T_C and the avoidance of being attacked A_N . This goal function could reflect both the urgency of intercepting the target and the strategic necessity of managing risks.

$$f_{ij} = -\frac{(T_C)_{i,j}}{\max_{i,j}(T_C)} - (A_N)_{:,j}, \tag{7}$$

where the minimum time-to-capture $(T_C)_{i,j}$ is normalized with the maximum value in vector T_C . This ensures that all values of $\frac{(T_C)_{i,j}}{\max_{i,j} T_C}$ are uniformed between zero and one so that their size is similar to the values of $(A_N)_{:,j}$.

There are two main motivations for adopting this goal/payoff function. First, this function facilitates a more holistic evaluation of pursuit strategies by not only focusing on quick interception but also ensuring that the maneuvers are strategically sound and sustainable over the long term. Second, the inclusion of both T_C and A_N allows the function to adapt to different tactical situations, providing a flexible tool that can adjust to the target's behavior and environmental variables. Based on function (7), the payoff assignment is presented in the Algorithm 1, which formulates the payoff matrix under the pursuit station $(x_t, v_t, x_a, x_i, v_i)$. The calculated payoff matrix would assist in the latter solution to the proposed ZSPEG-based model.

Algorithm 1: Payoff assignment for the pursuit team

Input: x_t, v_t, x_a, x_i, v_i
Output: $F = -\frac{(T_C)}{\max_{i,j} T_C} - (A_N)$;

```

1   for  $i = 1 \rightarrow n$  do
2       for  $j = 1 \rightarrow n$  do
3            $r_t = \frac{x_t - x_i}{\|x_t - x_i\|}$ ;
4            $(T_C)_{i,j} = \min\{t(x_t, x_i, r_{t_j}), T_U\}$ ;
5            $(A_N)_{:j} = \cos\left(\frac{\langle r_{t_j}, x_a - x_t \rangle}{\|r_{t_j}, x_a - x_t\|}\right)$ ;
6       end for
7   end for
8   if  $j = n + 1$  then
9        $r_{t_j} = \frac{x_t - x_i}{\|x_t - x_i\|}$ ;
10      for  $j = 1 \rightarrow n$  do
11           $(T_C)_{i,j} = \min\{t(x_t, x_i, r_{t_j}), T_U\}$ ;
12           $(A_N)_{:j} = 1$ ;
13      end for
14  end if

```

3. Methodology

The solution to the proposed ZSPEG-based model is to calculate the accurate and stable equilibrium point. Two types of Nash equilibrium exist in the ZSG-based models: pure-strategy Nash equilibrium and MNE.

3.1. Pure-Strategy Nash Equilibrium

A pure-strategy Nash equilibrium exists when each player selects a deterministic strategy. To obtain this equilibrium, players do not employ probability distributions. Based on the definition of the Nash equilibrium in [35], the pure-strategy Nash equilibrium for the proposed ZSPEG-based model $G := \{S_p, S_t, F\}$ could be explained by Definition 2.

Definition 2 (Pure-strategy Nash equilibrium [36]). *Firstly, it is assumed that the pursuit team selects the i_* th vehicle and the target chooses motion $r_{t_{j^*}}$, which satisfies the following conditions:*

$$i_* \in \arg \max_{S_p} \min_{S_t} F_{ij}, \tag{8}$$

$$j_* \in \arg \min_{S_t} \max_{S_p} F_{ij}, \tag{9}$$

where $F_{ij} = f_i$ under the condition that $r_t = r_{t_j}$. Then, (i_*, j_*) would be a pure-strategy Nash equilibrium if, and only if, Equation (10) is satisfied.

$$\max_{S_p} \min_{S_t} F_{ij} = \min_{S_t} \max_{S_p} F_{ij}. \tag{10}$$

Corollary 1. *There exists a situation where there is no pure-strategy equilibrium value in the pursuit model.*

Proof of Corollary 1. For any fixed $i \in \{1, \dots, n\}$, Equation (11) could be established because the maximum value of F_{ij} is greater than or equal to any value of F_{ij} .

$$\min_{S_t} \max_{S_p} F_{ij} \geq \min_{S_t} F_{ij}. \tag{11}$$

For the best i' on the right-hand side of Equation (11), the corresponding best $\min_{S_t} F_{ij}$ is one of any $\min_{S_t} F_{ij}$. Therefore, Equation (12) could be obtained as follows:

$$\min_{S_t} \max_{S_p} F_{ij} \geq \max_{S_p} \min_{S_t} F_{ij}. \tag{12}$$

According to the above two equations, once the maximum value of $\min_{S_t} \max_{S_p} F_{ij}$ is greater than the value of $\min_{S_t} F_{ij}$, the following equation would be obtained, which did not satisfy the requirement of the existence of the PNE. In this situation, there is no PNE in the model.

$$\min_{S_t} \max_{S_p} F_{ij} > \max_{S_p} \min_{S_t} F_{ij}. \tag{13}$$

□

3.2. MNE

Given the potential absence of a pure-strategy Nash equilibrium, MNE would be calculated for the proposed model based on the minimax theorem. MNE denotes a scenario in a game where players employ probability distributions to select various pure strategies. The minimax theorem, proposed by Jon Von Neumann, provides a solution that incorporates strategic randomness [37].

Assume that the ISVs adopt the mixed strategy, $\mathbf{p} = [p_1, \dots, p_n]^T$, to choose their motion, where $\forall i \in \{1, \dots, n\}, 0 \leq p_i \leq 1$, and the sum of \mathbf{p} is equal to 1. Similarly, the target is assumed to adopt the mixed strategy: $\mathbf{q} = (q_1, \dots, q_{n+1})$. The constraints for \mathbf{q} are the same as that for \mathbf{p} . Therefore, under mixed strategies, the payoff for the pursuit team can be calculated by Equation (14).

$$V_p = \mathbf{p}^T \mathbf{F} \mathbf{q}. \tag{14}$$

The goal of the pursuit ISVs is to maximize its reward, while the aim of the target is to minimize the reward of the pursuit team. Hence, their payoff functions are expressed in the following two equations, respectively.

$$V_{p^*} = \max_{S_p} (\mathbf{F} \mathbf{q}), \tag{15}$$

$$V_{q^*} = \max_{S_t} (\mathbf{p}^T \mathbf{F}). \tag{16}$$

After formulating the two payoffs, the existence of MNE in the proposed model is given and explained by Theorem 1.

Theorem 1 (existence of MNE [38]). For any two-player ZSG models, the following conditions could be satisfied:

- $V_{p^*} = V_{q^*} = V_*$, where V_* is called the minimax value of the game.
- The Nash equilibrium in this ZSG model is $(\mathbf{p}^*, \mathbf{q}^*)$.

Proof of Theorem 1. Suppose that $(\tilde{\mathbf{p}}, \tilde{\mathbf{q}})$ is an MNE in this model. Then, the payoff for the ISV is obtained by Equation (17).

$$V_* = \tilde{\mathbf{p}}^T \mathbf{F} \tilde{\mathbf{q}}. \tag{17}$$

As both the sum of \tilde{p} and \tilde{q} are equal to one, according to the above equations, Equation (18) could be obtained.

$$V_{\tilde{q}} = \min_{s_t} \left(\tilde{p}^T F \right) \leq \tilde{p}^T F \tilde{q} \leq \max_{s_p} \left(F \tilde{q} \right) = V_{\tilde{p}}. \tag{18}$$

Based on Corollary 1, it could be inferred that $\min_{s_t} \left(\tilde{p}^T F \right) \geq \max_{s_p} \left(F \tilde{q} \right)$. Therefore, the value of $V_{\tilde{q}}$ is equal to $V_{\tilde{p}}$. Equation (19) is subsequently formulated.

$$V_{\tilde{q}} = \min_{s_t} \left(\tilde{p}^T F \right) = \tilde{p}^T F \tilde{q} = \max_{s_p} \left(F \tilde{q} \right) = V_{\tilde{p}}. \tag{19}$$

It could be concluded that when (\tilde{p}, \tilde{q}) is an MNE, \tilde{p} and \tilde{q} must be the optimal strategies for the pursuit team and the target, respectively. \square

Definition 3 (best response strategy [39]). Suppose that p_{-*} is the mixed strategy set other than p_* ; then, the best response strategy for p_* is formulated by Equation (20).

$$BR_{p_*} := \arg \max_{p \in \Delta_i} u(p_*, p_{-*}), \tag{20}$$

$$s.t. \forall p' \in \Delta, u(p_*, p_{-*}) \geq u(p', p_{-*})$$

Notations:

- The equation must be constructed from the expected reward u .
- BR_{p_*} takes the mixed strategy in the p_{-*} as the independent variable. Correspondence: $\Delta_{-*} \rightarrow \Delta$.
- All the mixed strategies in the BR_{p_*} have no differences.

Corollary 2. p_* is the best response to q_* , and vice versa.

Proof of Corollary 2. Based on Equation (19), the following equations could be formulated:

$$V_{p_*} = \min_{s_t} \left(p^T F \right), \tag{21}$$

$$V_{q_*} = \max_{s_p} \left(F q \right). \tag{22}$$

The expected reward in Definition 3 for p_* and the q_* can be calculated by Equations (23) and (24), respectively.

$$u(p_*, p_{-*}) = p_*^T F q, \tag{23}$$

$$u(q_*, q_{-*}) = p^T F q_*. \tag{24}$$

The best response for p_* and q_* is denoted by Equations (25) and (26).

$$BR_{p_*} = \arg \max u(p_*, p_{-*}), \tag{25}$$

$$BR_{q_*} = \arg \max u(q_*, q_{-*}). \tag{26}$$

Based on Equation (19), Equations (27) and (28) can be established.

$$\max u(p_*, p_{-*}) = \min p_*^T F q, \tag{27}$$

$$\max u(q_*, q_{-*}) = \min p^T F q_*, \tag{28}$$

where the solution to the $\min_{S_p} p_*^T F q$ is $q = q_*$, and the solution to the $\max_{S_t} p^T F q_*$ is $p = p_*$. Finally, Equations (29) and (30) can be formulated.

$$BR_{p_*} = \arg \max u(p_*, p_{-*}) = q_*, \tag{29}$$

$$BR_{q_*} = \arg \max u(q_*, q_{-*}) = p_*. \tag{30}$$

Based on the above equations and proofs, q_* is the best response to p_* , and vice versa. \square

3.3. Computing Pursuit Motions

As MNE proved to be the best response in the proposed model, the solution to the $\max_{S_p}(Fq)$ or $\min_{S_t}(p^T F)$ can provide the optimal and stable motions for the vehicles. The calculation process is as follows. First of all, $p^T F q$ is set as V . Then, as the values in q are summed to be 1, the minimum value of $p^T F q$ is equal to the minimum value in the vector of $(p_*^T F)$. Therefore, solving the $\max_{S_p} F q$ problem is equivalent to solving the following linear programming problem (31):

$$\begin{aligned} & \max V, \\ & \text{s.t. } p^T F \geq V1, \\ & p_1 + \dots + p_n = 1, \\ & p_i \geq 0, i \in \{1, \dots, n\}. \end{aligned} \tag{31}$$

With the assistance of Algorithm 1 in generating the payoff matrix F , the MNE-based maneuver planning algorithm is shown in Algorithm 2. Based on the MNE framework derived from game-theoretic principles, Algorithm 2 aims to dynamically adjust the strategies of multiple pursuit agents based on real-time feedback from the operational environment and the target's maneuvers. The algorithm calculates the Nash equilibriums for different strategic setups, allowing each pursuit vehicle to select the best possible action that maximizes the collective chances of success.

Algorithm 2: MNE-based maneuver planner for pursuit vehicles.

Input: $x_t, v_t, x_n, x_i, v_i, i \in \{1, \dots, n\}$;
Output: r_i ;

```

1  if  $t \leq T_S$  and  $d_i \leq d_c$  and  $d_t \leq d_a$  then
2      formulate  $F$  according to Algorithm 1;
3      solve the problem (31);
4      obtain MNE  $p_*$ ;
5       $i^* = \arg \max (p_*)$ ;
6  end if
7  if  $i = i^*$  then
8       $r_i = \frac{x_t - x_i}{\|x_t - x_i\|}$ ;
9  else
10      $r_i = 0 \rightarrow 0 \parallel \frac{x_t - x_i}{\|x_t - x_i\|}$ ;
11 end if

```

3.4. Step-by-Step Process

Based on the above introduction, the flowchart illustrating the different phases of the algorithm in Section 3 is provided in Figure 3.



Figure 3. The overall step-by-step process of the proposed algorithm.

4. Simulations

In this section, we present the results obtained from the simulations and discuss their implications. The objective is to evaluate the performance of the proposed ZSPEG-based model and its effectiveness in achieving stable and accurate equilibrium points. We will analyze the outcomes of different scenarios and compare them with baseline methods to demonstrate the advantages of our method. The results will be examined in terms of time-to-capture, maneuver efficiency, etc.

To set the parameters before the simulations, a series of parameters were chosen based on their relevance to the performance and functionality of the ISV during pursuit tasks. The chosen parameters are presented in Table 2. The selection criteria for these parameters were as follows: (i) Sway velocity of the vehicle (v)—the sway velocity range of 5 m/s to 15 m/s was determined by referring to the currently used ISVs, ensuring that the vehicle could adapt to the various operational speeds required for different scenarios. (ii) Sway velocity of the target (v_t)—a fixed value of 10 m/s was chosen based on the average target speeds observed in similar pursuit and evasion scenarios. This value balances the difficulty for the ISV to capture the target while allowing for a meaningful analysis of the algorithm's effectiveness. (iii) Length of the vehicle (l)—a standard length of 4 m was selected to represent the typical dimensions of ISVs used in practical applications, ensuring the results are applicable to real-world scenarios. (iv) Intercepted distance to target (d_c)—the value of 12 m was derived based on the length of the vehicle, indicating the optimal interception distance for successful captures while minimizing the risk of overshooting or collision. (v) Attack distance of target (d_a)—an attack distance of 6 m was chosen based on the ISV's common response capabilities, ensuring effective engagement. (vi) Safe distance between the ISVs (d_s)—a safe distance of 9 m was determined to prevent collisions between multiple ISVs operating within the same area, allowing for coordinated maneuvers. (vii) Interval time (Δt)—an interval time of 0.01 s was selected to ensure precise and responsive updates to the ISV's path-planning algorithm, enabling real-time adjustments. (viii) Survival time of target (T_S)—the survival time of 2.0 s was established to assess the ISV's rapid response capabilities, providing a challenging yet achievable timeframe for the pursuit. (IX) Protected area (A)—the protected area of 40 m \times 80 m was set to represent a controlled environment where the ISVs and targets operate, allowing for consistent and repeatable testing conditions.

Table 2. Parameters chosen for the monomer pursuit.

Parameters	Definitions	Values
v	Sway velocity of the vehicle	5–15 m/s
v_t	Sway velocity of the target	10 m/s
l	Length of the vehicle	4 m
d_c	Intercepted distance to target	12 m
d_a	Attack distance of target	6 m
d_s	Safe distance between the ISVs	9 m
Δt	Interval time	0.01 s
T_S	Survival time of target	2.0 s
A	Protected area	40 m × 80 m

First of all, to compare the proposed ZSPEG-based planning scheme in both the single-ISV scenario and the multi-ISV scenario, the strategy set and the reward function for the single-ISV pursuit scenario are constructed as follows. In the single-ISV pursuit scenario, the heading vector of the ISV is limited to within their field of vision constraints, which are denoted by FoV_p . Therefore, r_p , the unit yaw vector for the pursuit vehicle, is limited to the range of $\left[-\frac{FoV_p}{2}, \frac{FoV_p}{2}\right]$. Furthermore, the maneuver planning for the ISV is closely related to its minimum steering angle (MSA). Then, under the constraints of the FoV and MSA , there could be $\lfloor FoV_p/MSA_p \rfloor$ options for ISV to move. Hence, the i th strategy for the pursuit vehicle is denoted by $r_{p_i} = i - \frac{FoV_p}{2}, i \in \{1, 2, \dots, \lfloor FoV_p/MSA_p \rfloor\}$. Similarly, the maneuver strategy set for the moving target could be assumed by $r_{t_j} = j - \frac{FoV_T}{2}, j \in \{1, 2, \dots, \lfloor FoV_T/MSA_T \rfloor\}$. The payoff function for the single-ISV target pursuit is designed to consider two factors: the distance between the pursuit ISV and the target, and the distance between the target and the point of attack. These two factors could be formulated by the following equations:

$$(D_P)_{i,j} = -\frac{\tan\left(d_p^1(r_{p_i}, r_{t_j}) - d_p^0\right)}{K_A}, \tag{32}$$

$$(D_T)_{i,j} = -\frac{\tan\left(d_t^1(r_{t_j}) - d_t^0\right)}{K_B}, \tag{33}$$

where $(D_P)_{i,j}$ denotes the pursuit feedback of the distance between the vehicle and the target. $(D_T)_{i,j}$ denotes the attacking feedback of the distance between the target and the intended node. $d_p^1(r_{p_i}, r_{t_j})$ and $d_t^1(r_{t_j})$ denote the corresponding distance in the future situation after adopting the related motions. d_p^0 and d_t^0 represent the distances in the current situation. Then, the payoff for the pursuit ISV can be obtained by Equation (34).

$$(F)_{i,j} = \lg\left((D_P)_{i,j} + (D_T)_{i,j} + K_C\right), \tag{34}$$

where $K_A, K_B,$ and K_C are all constants.

To generate and analyze the trajectories under the single-ISV pursuit model, the initial position for the moving target is set to along the boundary of the protected area. The position of the unknown attacked node should be within the protected area. Thus, it is assumed that $x_t(0) = (80,5), x_a = (20,10), x_p(0) = (40,22), FoV_T = 120^\circ,$ and $MSA_T = 1^\circ$. After being solved by Algorithm 2, under different FoV_p and MSA_p configurations, nine trajectories in the single-ISV pursuit are shown in Figure 4. Although the rational motions were sequentially generated for the ISV in time, the vehicle fails to capture the target in all the situations. If the ISV does not fulfil the necessary conditions, the target could gain the initiative to attack its intended node. Therefore, it is an inevitable trend that multi-ISV pursuit replaces single-ISV pursuit.

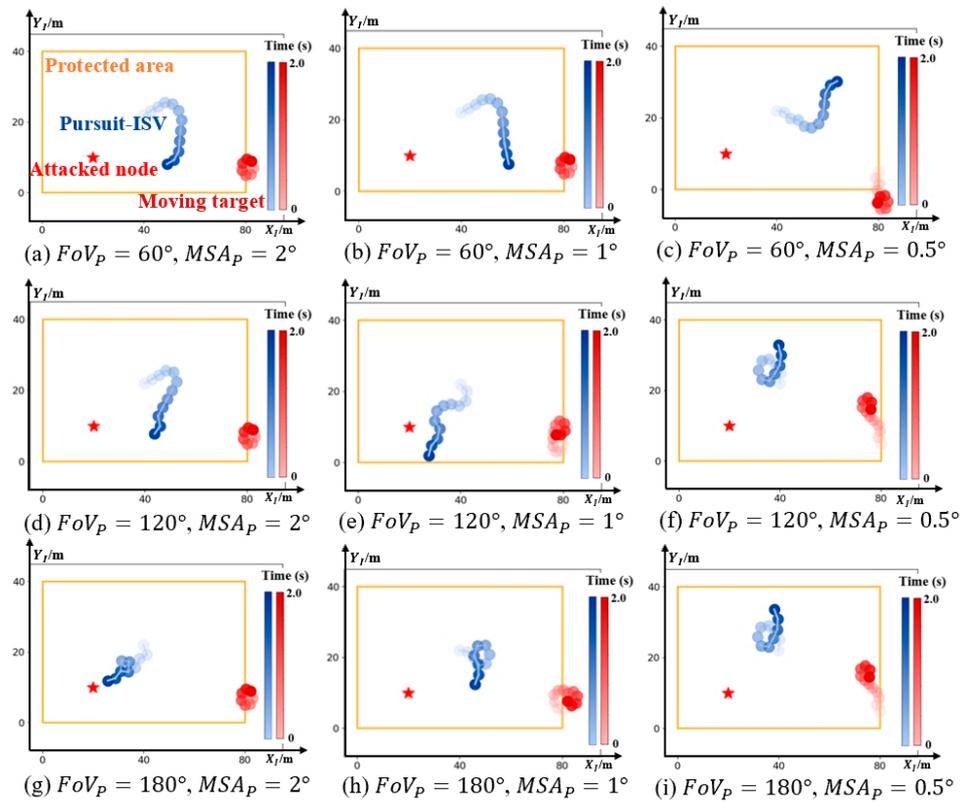


Figure 4. Parameter assumption for the monomer pursuit.

4.1. Model Evaluation

In this subsection, we evaluate our proposed ZSPEG-based model by comparing it with two other target-pursuit methods through a series of simulations. The compared methods include the following:

Graph theory-based: There are many versions of Voronoi algorithms being applied in a multi-robot pursuit scenario [40]. Among these Voronoi-based methods, the GAM–Voronoi algorithm outperforms the others in a bounded convex environment, which utilizes a global area-minimization (GAM) strategy [41]. This was chosen as the representative of the Voronoi-related algorithms.

Cooperative target hunting: To compare our method with strategy-based cooperative pursuit methods, a popular method, called Cooperative Hunting based on the Dynamic Hunting Points Allocation (CH-DHPA) [42–44], is selected. This method transforms the target pursuit to the dynamic hunting points allocation problem.

Our Proposed Method (TGRP-MNE): This method aims to achieve stable and accurate equilibrium points by implementing the ZSPEG-based model.

The reasons why the GAM–Voronoi algorithm and CH-DHPA were chosen for the comparison with our proposed methods include two main criteria. First, both the GAM–Voronoi and CP-DHPA methods address the core issue of maneuver planning in pursuit scenarios, similar to our study’s focus. This relevance ensures that the comparisons are directly applicable and meaningful to the field of pursuit–evasion games involving multiple ISVs. Second, by comparing these methods, this study could have significant theoretical implications regarding the applicability of game theory versus graph theory and cooperative strategies in real-world scenarios.

The required parameters, the pros, and the cons of the three methods are shown in Table 3. To ensure the fairness of the experiment, while providing our method with additional parameters and information, we established ideal environmental conditions for the other two methods. The relevant environmental elements and geographic information were set to ideal situations devoid of external disturbances. Furthermore, we also idealized

the communication situation for these two methods. As they involve distributed communication, effective communication among multiple ISVs is crucial during task allocation and team collaboration. Hence, to enhance fairness during the comparison simulation, signal transmission was also configured to be in an ideal state for the methods of GAM-Voronoi and CP-DHPA. Based on the above settings, simulations were conducted under the fixed initial positions for the ISVs. The relevant initial positions were set as follows: $x_1(0) = (40,30)$, $x_2(0) = (40,10)$, $x_3(0) = (50,20)$, and $x_4(0) = (30,20)$. Furthermore, $x_t(0)$ and $x_n(0)$ were randomly chosen. The settings for the other parameters are shown in Table 2. For every group, the simulations were carried out 1000 times. The assessed metrics include the capture success rate— $r_s(\%)$; the rate of nodes being attacked— $r_a(\%)$; the number of collisions— C ; and the approximate average energy consumption per successful capture— AEC (m·s).

Table 3. Preliminaries of the three methods.

Methods	Required Parameters	Pros	Cons
GAM-Voronoi	(i) Spatial positioning parameters; (ii) Target attributes; (iii) Initial conditions.	Effectively partitions space to address unknown target movements	Requires more geographic information and robust communications
CP-DHPA	(i) Pursuit team attributes; (ii) Target attributes; (iii) Environmental factors; (iv) Allocation rules.	(i) Emphasizes teamwork and flexibility; (ii) Utilizes team resources efficiently;	Requires more internal team coordination and robust communications
TGRP-MNE (ours)	(i) Behavioral patterns of the target; (ii) Pursuit team attributes; (iii) Rules of the game.	(i) Reflects the competitive interaction; (ii) Fictitiously plays the strategies and the motions of the target; (iii) Maintains high flexibility;	Requires more parameters and pursuit information

The selected three parameters were computed as follows: (i) The capture success rate $r_s(\%)$ —this parameter represents the percentage of simulations where the pursuit strategy successfully led to the target being captured within the defined constraints and conditions. It is computed by dividing the number of successful captures by the total number of simulation runs and then multiplying by 1000 to express this as a percentage. (ii) The rate of nodes being attacked $r_a(\%)$ —similar to the success rate, the avoidance rate measures the percentage of simulations in which the target successfully evaded capture for the duration of the simulation or until it reached a safe zone. This is calculated by counting the instances of successful evasion, dividing this by the total simulations, and converting the result into a percentage. (iii) The average energy consumption per successful capture AEC (m·s)—to compute the AEC , the total energy consumed in each run is summed up and then divided by the number of runs to obtain the average. This provides a measure of how much energy, on average, is required for a vehicle to participate in a pursuit scenario over the course of the simulations. Mathematically, it can be represented as $AEC = \frac{\sum_{i=1}^n E_i}{n}$, where E_i is the energy consumed in the i th simulation and n is the total number of simulations. This calculation helps to understand the energy efficiency of the different tactical approaches under various operational conditions, which is crucial for optimizing the pursuit strategies in terms of both effectiveness and sustainability.

The simulation results under the varying survival times, the varying pursuit velocities, and the different areas are shown in Tables 4–6, respectively. It should be noted that only the CP-DHPA method has the collision parameter C . This is because the CP-DHPA method requires additional collision avoidance measures, for which we need to analyze

its effectiveness in avoiding collisions. However, in the GAM–Voronoi method, collision avoidance is inherently built into the method, preventing any collisions from occurring. In our TGRP–MNE method, a relay–pursuit strategy is adopted. Therefore, collisions between pursuit vehicles are not a concern, either. As presented by these three tables, our method outperforms the other two methods in achieving a relatively high success rate of capture r_s and the lowest average energy consumption AEC . To be specific, when comparing Tables 4 and 5, we observed that the rate of nodes being attacked, r_a , in the CP–DHPA method is lower than that in our proposed TGRP–MNE method. This phenomenon indicates that that our TGRP–MNE method might implement a more effective or aggressive pursuit strategy. Such a strategy could better block the routes of the target or more accurately predict the target’s movements, but may not provide sufficient protection for the target’s attack point. Furthermore, according to Table 6, our proposed method demonstrates a superior performance compared to GAM–Voronoi and CP–DHPA in terms of its success rate and energy consumption across all area scales. Based on these findings, if node protection is the primary goal in surface pursuit, CP–DHPA would be the preferred choice among the three methods. However, in general situations, achieving a high and stable success rate of capture is paramount for the surface pursuit. With adaptability to factors T_S , v_p , and the area scale, and the ability to achieve a high and stable success rate of capture and low energy consumption, our method can be considered the optimal choice for surface pursuit.

Table 4. Comparison under a varying T_S with $v_p = 15$ m/s.

T_S (s)	TGRP–MNE (Ours)			GAM–Voronoi			CP–DHPA			
	r_s (%)	r_a (%)	AEC (m·s)	r_s (%)	r_a (%)	AEC (m·s)	r_s (%)	r_a (%)	C	AEC (m·s)
1.0	48.7	8.0	134.21	51.0	11.4	520.92	78.3	7.3	113	497.61
1.5	84.6	10.1	184.07	78.8	9.8	707.13	89.4	10.0	101	484.65
2.0	87.8	12.8	201.75	84.6	14.4	792.63	82.7	6.6	109	478.98
2.5	88.6	11.2	206.06	86.2	13.8	801.63	80.5	7.0	125	509.58
3.0	87.8	12.2	200.48	85.9	14.1	809.64	83.7	6.3	100	466.83

Table 5. Comparison under a varying v_p with $T_S = 2.0$ s.

T_S (s)	TGRP–MNE (Ours)			GAM–Voronoi			CP–DHPA			
	r_s (%)	r_a (%)	AEC (m·s)	r_s (%)	r_a (%)	AEC (m·s)	r_s (%)	r_a (%)	C	AEC (m·s)
5.0	63.0	18.3	256.59	60.4	21.4	1084.4	83.3	13.1	7	820.89
7.5	73.4	18.8	245.07	66.3	17.8	992.16	84.1	11.7	31	742.42
10.0	80.6	16.9	237.92	71.4	18.1	873.45	84.9	9.5	55	613.80
12.5	84.0	14.6	217.19	83.4	11.8	834.48	84.4	8.7	69	519.21
15.0	87.8	12.8	201.75	84.6	14.4	792.63	82.7	6.6	109	478.98

Table 6. Comparison under different areas.

$A(m^2)$	TGRP-MNE (Ours)			GAM-Voronoi			CP-DHPA			
	$r_s(\%)$	$r_a(\%)$	AEC (m·s)	$r_s(\%)$	$r_a(\%)$	AEC (m·s)	$r_s(\%)$	$r_a(\%)$	C	AEC (m·s)
40 × 40	92.0	8.0	82.7	85.7	14.3	490.05	82.0	5.0	130	81.12
40 × 60	89.9	10.0	95.72	79.6	16.3	710.73	72.0	7.0	210	129.24
40 × 80	86.8	7.7	122.15	60.6	18.1	963.54	84.0	2	120	253.80
40 × 80	87.8	12.8	201.75	84.6	14.4	792.63	82.7	6.6	109	478.98

4.2. Sensitive Analysis

In this subsection, the sensitivity of our method under different conditions is tested and analyzed. The conducted experiments include the following: (i) analysis of n , T_S , and v_p , where we tested our methods under multiple scenarios with different initial conditions, including the number of vehicles n , the survival time T_S , and the pursuit velocity v_p ; (ii) system stability, where we assessed the stability of the pursuit system by observing the consistency of the pursuit strategies over extended simulation runs with different target strategy sets. These experiments could allow for a comprehensive evaluation of the proposed method, demonstrating its advantages and identifying potential areas for improvement.

Analysis of n , T_S , and v_p : To explore the impact of variables on the performance, simulations are conducted using different numbers of vehicles n , various survival times T_S , and changing pursuit velocities v_p . The results are shown in Figure 5. In Figure 5a, once T_S exceeds 1.6 s, all the success rates would be stable, with a small fluctuation of around 85%. As shown in Figure 5b, there were no significant differences or changes among the data. The rates of attack were all below 15%. The above observations indicate that when the survival time exceeds a certain value, further increasing T_S no longer contributes to improvements in the performance. According to the results shown in Figure 5c,d, with the pursuit velocity increases, the capture success rate in all the teams increases, while the rates of the node being attacked all decrease. To obtain a higher success rate and a lower rate of being attacked in a pursuit team, improving the pursuit velocity could be helpful. Furthermore, the team with $n = 5$ always achieves the best pursuit performance due to its having the highest number of vehicles. In sum, the variation in the team-scale and the pursuit velocity has a more significant impact on the model’s performance compared to the changes in time. Therefore, during the optimization of the system, particular attention should be paid to the n and v_p adjustments.

System stability: To determine how the strategy set of the target affects the pursuit, the strategy sets for the target are redesigned. For a pursuit team with $n = 3$, the strategy sets for the target are reformulated by the two newly designed sets: S_{ig} and S_{tc} . The first strategy set is based on the ZSG model, which was introduced in Section 2.2. S_{tc} is more complex than the ZSG-based strategy set and the original one. The detailed strategies in S_{tc} are shown as follows, and are also displayed in Figure 6:

- r_{t_1} : the target evades from the nearest ISV.
- r_{t_2} : the target adopts the collective evasion strategy, which is explained in Definition 4.
- r_{t_3} : the target heads directly toward its attacked node, denoted by the red star in Figure 6.
- r_{t_4} : the direction of the target is the angle bisector formed by r_{t_1} and r_{t_2} .

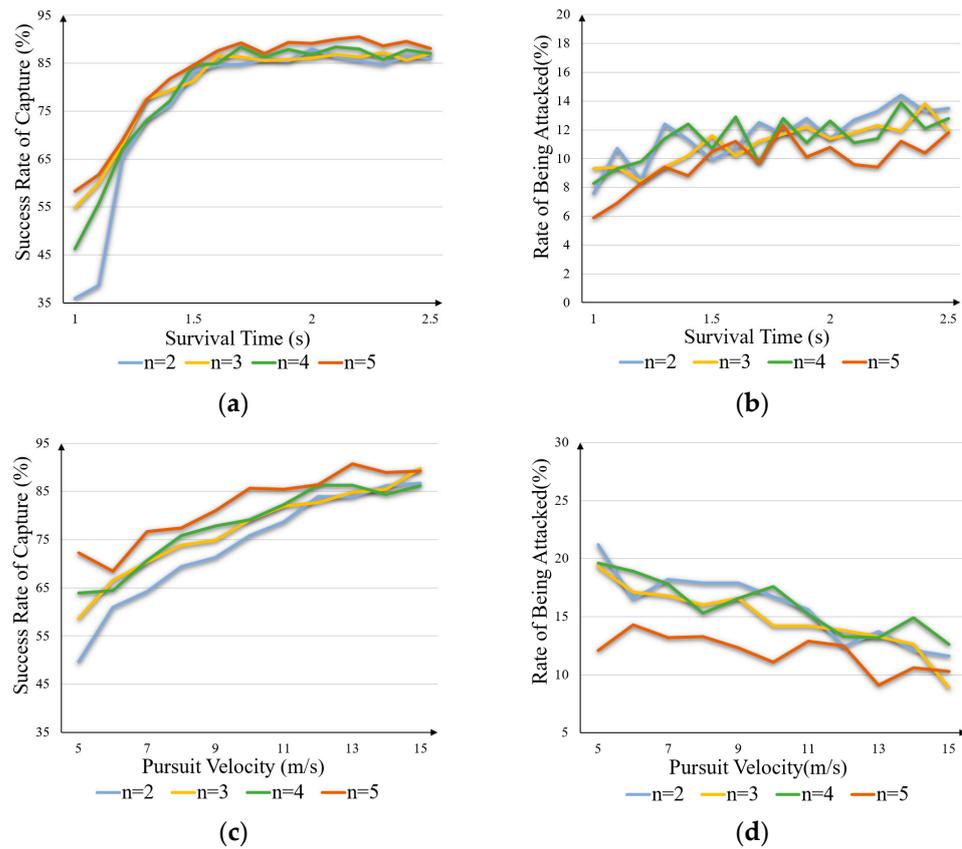


Figure 5. Performance of the pursuit team with different N under (a,b) varying survival times and (c,d) varying pursuit velocities.

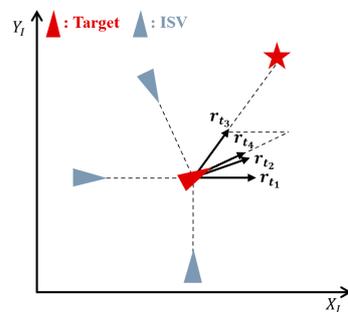


Figure 6. The enriched motions for the target.

Definition 4 (collective evasion strategy). From the perspective of the target, all angles formed between the two adjacent ISVs are taken into account in the collective evasion strategy. Additionally, the moving direction of the target is the parallelogram of the maximum angle formed by the two adjacent ISVs. The purpose of this strategy is to enable the target to move immediately away from the entire pursuit team, rather than just from one ISV. The calculation for this evasion motion is as follows: Firstly, $\alpha_i := \angle(x_t, x_i)$ is set to represent the angle of the vector $(x_t - x_i)$. $\beta_i := \alpha_{i+1} - \alpha_i$ denotes the angle between two adjacent ISVs. If $i = n$, α_{i+1} is equal to α_1 . Therefore, r_{t_2} could be obtained by Equation (35).

$$r_{t_2} = [\cos(\beta_{i_{mh}} + \beta_{i_m}), \sin(\beta_{i_{mh}} + \beta_{i_m})]^T \tag{35}$$

where i_m is the index when β_i is taken to its maximum value. $\beta_{i_{mh}}$ is half of the β_{i_m} .

With changing survival times, simulations were carried out under $S_t = S_{tg}$, $S_t = S_{tc}$ and $S_t = S_{to}$, respectively. The results are shown in Figure 7. In terms of the success rate of

capture and the rate of being attacked, it is apparent that there is no significant difference among the three strategy sets. This discovery underscores the robustness of our proposed method, showcasing its ability to maintain a consistent performance despite variations in the target's strategy set.

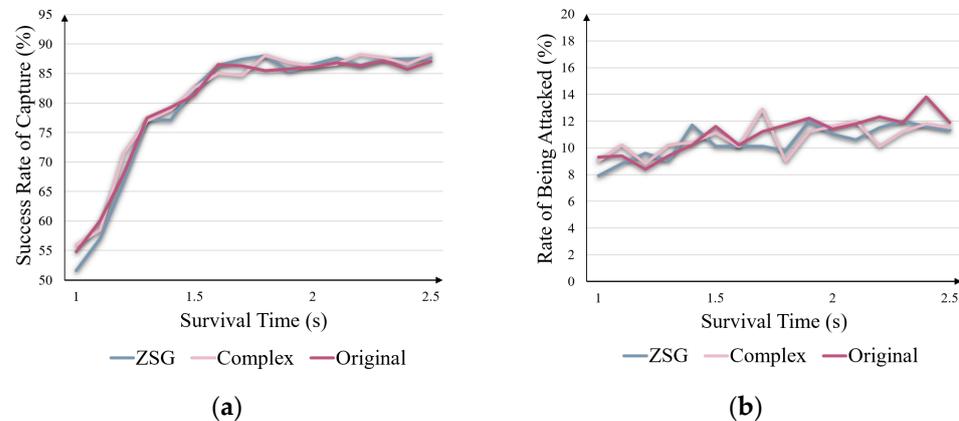


Figure 7. Performance for the three strategy sets for the target under (a,b) varying survival times.

5. Conclusions

This paper introduced a novel ZSPEG-based maneuvering method for multiple ISVs to pursue a moving target. The proposed method addresses the challenge of realizing the interaction between the pursuit team and the target and mitigating the data scarcity of the target. At first, the original $n + 1$ non-cooperative multi-ISV game model was reformulated into a two-player ZSPEG model. Based on the fictitious play solution, an MNE-based decision-making process was designed to provide the vehicles with rational and stable motions. Finally, the simulation results show that the proposed method achieves a comprehensively outstanding performance, obtaining high capture success rates, low energy costs, generalization, and adaptability. Nevertheless, the proposed model depends on centralized computing to achieve a consensus among multiple ISVs, which limits its adaptability to more intricate surface environments. To enhance the navigation of the surface vehicles under constrained surface communication conditions, future research will integrate decentralized computing into the proposed method.

Author Contributions: Conceptualization, L.H., W.C., H.C., C.S. and W.L.; methodology, L.H., W.C. and H.C.; software, L.H.; validation, L.H.; formal analysis, L.H., W.C. and H.C.; investigation, W.C., C.S. and W.L.; resources, W.C. and C.S.; data curation, L.H.; writing—original draft preparation, L.H. and W.C.; writing—review and editing, L.H., W.C., H.C. and C.S.; visualization, L.H.; supervision, W.C. and C.S.; project administration, W.C.; funding acquisition, W.C. and C.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Zhejiang Key R&D Program No. 2021C03157, start-up funding from Westlake University under grant number 041030150118 and Scientific Research Funding Project of Westlake University under Grant No. 2021WUFP017.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Acknowledgments: The authors would like to thank the anonymous reviewers for their valuable insights and feedback.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Bovcon, B.; Muhovič, J.; Vranac, D.; Mozetič, D.; Perš, J.; Kristan, M. MODS—A USV-oriented object detection and obstacle segmentation benchmark. *IEEE Trans. Intell. Transp. Syst.* **2021**, *23*, 13403–13418. [[CrossRef](#)]

2. Peng, Z.; Wang, J.; Wang, D.; Han, Q.-L. An overview of recent advances in coordinated control of multiple autonomous surface vehicles. *IEEE Trans. Ind. Inform.* **2020**, *17*, 732–745. [[CrossRef](#)]
3. Zhou, X.; Wu, P.; Zhang, H.; Guo, W.; Liu, Y. Learn to navigate: Cooperative path planning for unmanned surface vehicles using deep reinforcement learning. *IEEE Access* **2019**, *7*, 165262–165278. [[CrossRef](#)]
4. Li, Y.; Li, X.; Wei, X.; Wang, H. Sim-real joint experimental verification for an unmanned surface vehicle formation strategy based on multi-agent deterministic policy gradient and line of sight guidance. *Ocean Eng.* **2023**, *270*, 113661. [[CrossRef](#)]
5. Chen, J.; Wang, J.; Hou, X.; Fang, Z.; Du, J.; Ren, Y. Advance into ocean: From bionic monomer to swarm intelligence. *Acta Electron. Sin.* **2021**, *49*, 2458–2467.
6. Wolf, M.T.; Assad, C.; Kuwata, Y.; Howard, A.; Aghazarian, H.; Zhu, D.; Lu, T.; Trebi-Ollennu, A.; Huntsberger, T. 360-degree visual detection and target tracking on an autonomous surface vehicle. *J. Field Robot.* **2010**, *27*, 819–833. [[CrossRef](#)]
7. Dai, S.-L.; He, S.; Ma, Y.; Yuan, C. Cooperative learning-based formation control of autonomous marine surface vessels with prescribed performance. *IEEE Trans. Syst. Man Cybern. Syst.* **2021**, *52*, 2565–2577. [[CrossRef](#)]
8. Pehlivanoglu, Y.V. A new vibrational genetic algorithm enhanced with a Voronoi diagram for path planning of autonomous UAV. *Aerosp. Sci. Technol.* **2012**, *16*, 47–55. [[CrossRef](#)]
9. Pierson, A.; Wang, Z.; Schwager, M. Intercepting rogue robots: An algorithm for capturing multiple evaders with multiple pursuers. *IEEE Robot. Autom. Lett.* **2016**, *2*, 530–537. [[CrossRef](#)]
10. Farinelli, A.; Iocchi, L.; Nardi, D. Distributed on-line dynamic task assignment for multi-robot patrolling. *Auton. Robot.* **2017**, *41*, 1321–1345. [[CrossRef](#)]
11. Makkapati, V.R.; Tsiotras, P. Optimal evading strategies and task allocation in multi-player pursuit–evasion problems. *Dyn. Games Appl.* **2019**, *9*, 1168–1187. [[CrossRef](#)]
12. Qu, X.; Gan, W.; Song, D.; Zhou, L. Pursuit-evasion game strategy of USV based on deep reinforcement learning in complex multi-obstacle environment. *Ocean Eng.* **2023**, *273*, 114016. [[CrossRef](#)]
13. Ge, J.; Tang, L.; Reimann, J.; Vachtsevanos, G. Suboptimal approaches to multiplayer pursuit-evasion differential games. In Proceedings of the AIAA Guidance, Navigation, and Control Conference and Exhibit, Keystone, CO, USA, 21–24 August 2006.
14. Zhang, Y.; Zhang, P.; Wang, X.; Song, F.; Li, C.; Hao, J. An open loop Stackelberg solution to optimal strategy for UAV pursuit-evasion game. *Aerosp. Sci. Technol.* **2022**, *129*, 107840. [[CrossRef](#)]
15. Tan, M.; Shen, H. Three-Dimensional Cooperative Game Guidance Law for a Leader–Follower System with Impact Angles Constraint. *IEEE Trans. Aerosp. Electron. Syst.* **2024**, *60*, 405–420. [[CrossRef](#)]
16. Raivio, T.; Ehtamo, H. Visual aircraft identification as a pursuit-evasion game. *J. Guid. Control Dyn.* **2000**, *23*, 701–708. [[CrossRef](#)]
17. Carr, R.W.; Cobb, R.G.; Pachter, M.; Pierce, S. Solution of a pursuit–evasion game using a near-optimal strategy. *J. Guid. Control Dyn.* **2018**, *41*, 841–850. [[CrossRef](#)]
18. Nikooeinejad, Z.; Delavarkhalafi, A.; Heydari, M. A numerical solution of open-loop Nash equilibrium in nonlinear differential games based on Chebyshev pseudospectral method. *J. Comput. Appl. Math.* **2016**, *300*, 369–384. [[CrossRef](#)]
19. Sun, S.; Zhang, Q.; Loxton, R.; Li, B. Numerical solution of a pursuit-evasion differential game involving two spacecraft in low earth orbit. *J. Ind. Manag. Optim.* **2015**, *11*, 1127–1147. [[CrossRef](#)]
20. Huang, K.; Ma, C.; Li, C.; Chen, Y.-H. High-order robust control and stackelberg game-based optimization for uncertain fuzzy pmsm system with inequality constraints. *ISA Trans.* **2023**, *134*, 451–459. [[CrossRef](#)]
21. Ma, C.; Huang, K.; Wu, Q.; Sun, H. Cooperative game-based optimization of flexible robust constraint following control for spacecraft rendezvous system with uncertainties. *IEEE Trans. Syst. Man Cybern. Syst.* **2023**, *53*, 6849–6860. [[CrossRef](#)]
22. Ruan, W.; Sun, Y.; Deng, Y.; Duan, H. Hawk-pigeon game tactics for unmanned aerial vehicle swarm target defense. *IEEE Trans. Ind. Inform.* **2023**, *19*, 11619–11629. [[CrossRef](#)]
23. Selvakumar, J.; Bakolas, E. Feedback strategies for a reach-avoid game with a single evader and multiple pursuers. *IEEE Trans. Cybern.* **2019**, *51*, 696–707. [[CrossRef](#)] [[PubMed](#)]
24. Shaferman, V.; Shima, T. Cooperative multiple-model adaptive guidance for an aircraft defending missile. *J. Guid. Control Dyn.* **2010**, *33*, 1801–1813. [[CrossRef](#)]
25. Cheng, L.; Yuan, Y. Adaptive multi-player pursuit–evasion games with unknown general quadratic objectives. *ISA Trans.* **2022**, *131*, 73–82. [[CrossRef](#)] [[PubMed](#)]
26. Brown, N.; Sandholm, T. Regret transfer and parameter optimization. In Proceedings of the AAAI Conference on Artificial Intelligence, Québec City, QC, Canada, 27–31 July 2014.
27. Bailey, J.; Piliouras, G. Fast and furious learning in zero-sum games: Vanishing regret with non-vanishing step sizes. *Adv. Neural Inf. Process. Syst.* **2019**, *32*.
28. Huang, Y.; Zhuang, L.; Zhao, C.; Liu, H. Efficient Double Oracle for Extensive-Form Two-Player Zero-Sum Games. In Proceedings of the International Conference on Neural Information Processing, Virtual Event, 22–26 November 2022.
29. Zhang, M.; Zhao, P.; Luo, H.; Zhou, Z.-H. No-regret learning in time-varying zero-sum games. In Proceedings of the International Conference on Machine Learning, Baltimore, MD, USA, 17–23 July 2022.
30. Vlatakis-Gkaragkounis, E.-V.; Flokas, L.; Lianias, T.; Mertikopoulos, P.; Piliouras, G. No-regret learning and mixed nash equilibria: They do not mix. *Adv. Neural Inf. Process. Syst.* **2020**, *33*, 1380–1391.
31. Anagnostides, I.; Farina, G.; Panageas, I.; Sandholm, T. Optimistic mirror descent either converges to nash or to strong coarse correlated equilibria in bimatrix games. *Adv. Neural Inf. Process. Syst.* **2022**, *35*, 16439–16454.

32. Abu-Khalaf, M.; Lewis, F.L.; Huang, J. Neurodynamic programming and zero-sum games for constrained control systems. *IEEE Trans. Neural Netw.* **2008**, *19*, 1243–1252. [[CrossRef](#)]
33. Gu, N.; Wang, D.; Peng, Z.; Wang, J.; Han, Q.-L. Advances in line-of-sight guidance for path following of autonomous marine vehicles: An overview. *IEEE Trans. Syst. Man Cybern. Syst.* **2022**, *53*, 12–28. [[CrossRef](#)]
34. Pan, T.; Yuan, Y. A region-based relay pursuit scheme for a pursuit–evasion game with a single evader and multiple pursuers. *IEEE Trans. Syst. Man Cybern. Syst.* **2022**, *53*, 1958–1969. [[CrossRef](#)]
35. Frihauf, P.; Krstic, M.; Basar, T. Nash equilibrium seeking in noncooperative games. *IEEE Trans. Autom. Control* **2011**, *57*, 1192–1207. [[CrossRef](#)]
36. Rosenthal, R.W. A class of games possessing pure-strategy Nash equilibria. *Int. J. Game Theory* **1973**, *2*, 65–67. [[CrossRef](#)]
37. Von Neumann, J.; Morgenstern, O. *Theory of Games and Economic Behavior*; Princeton University Press: Princeton, NJ, USA, 1944.
38. Harsanyi, J.C. Games with randomly disturbed payoffs: A new rationale for mixed-strategy equilibrium points. *Int. J. Game Theory* **1973**, *2*, 1–23. [[CrossRef](#)]
39. Wu, Z.; Li, Q.; Wu, W.; Zhao, M. Crowdsourcing model for energy efficiency retrofit and mixed-integer equilibrium analysis. *IEEE Trans. Ind. Inform.* **2019**, *16*, 4512–4524. [[CrossRef](#)]
40. Zhou, Z.; Zhang, W.; Ding, J.; Huang, H.; Stipanović, D.M.; Tomlin, C.J. Cooperative pursuit with Voronoi partitions. *Automatica* **2016**, *72*, 64–72. [[CrossRef](#)]
41. Huang, H.; Zhang, W.; Ding, J.; Stipanović, D.M.; Tomlin, C.J. Guaranteed decentralized pursuit-evasion in the plane with multiple pursuers. In Proceedings of the IEEE Conference on Decision and Control and European Control, Orlando, FL, USA, 12–15 December 2011.
42. Hung, N.T.; Rego, F.F.; Pascoal, A.M. Cooperative distributed estimation and control of multiple autonomous vehicles for range-based underwater target localization and pursuit. *IEEE Trans. Control Syst. Technol.* **2021**, *30*, 1433–1447. [[CrossRef](#)]
43. Sun, Z.; Sun, H.; Li, P.; Zou, J. Cooperative strategy for pursuit-evasion problem with collision avoidance. *Ocean Eng.* **2023**, *279*, 114476. [[CrossRef](#)]
44. Li, R.-Z.; Yang, H.-Z.; Xiao, C.-S. Cooperative hunting strategy for multi-mobile robot systems based on dynamic hunting points. *Control Eng. China* **2019**, *26*, 510.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.