

## Article

# Research on Unmanned Aerial Vehicle Cluster Collaborative Countermeasures Based on Dynamic Non-Zero-Sum Game under Asymmetric and Uncertain Information

Pengcheng Wu <sup>1,\*</sup>, Hongqiao Wang <sup>2</sup>, Gaowei Liang <sup>3</sup> and Peng Zhang <sup>4,\*</sup> <sup>1</sup> College of Automation, Northwestern Polytechnical University, Xi'an 710000, China<sup>2</sup> College of Mathematics and Statistics, Northwestern Polytechnical University, Xi'an 710000, China; 2020303240@mail.nwpu.edu.cn<sup>3</sup> College of Cyberspace Security, Northwestern Polytechnical University, Xi'an 710000, China; mion@mail.nwpu.edu.cn<sup>4</sup> Engineering Training Centre, Northwestern Polytechnical University, Xi'an 710000, China

\* Correspondence: wupengcheng29@mail.nwpu.edu.cn (P.W.); pengzhang@nwpu.edu.cn (P.Z.)

**Abstract:** Unmanned aerial vehicle (UAV) swarm coordinated confrontation is a hot topic in academic research at home and abroad, and dynamic maneuver decision-making is one of the most important research fields for UAV countermeasures. Aiming at the complexity, uncertainty and confrontation of UAV cooperative confrontation, concepts such as relative advantage degree and advantage coefficient are introduced, and game theory is used as a framework to construct a dynamic non-zero-sum game UAV cluster cooperative confrontation decision-making model, and finally convert it into an optimization problem. On this basis, using the Nash equilibrium solution method of multi-strategy fusion particle swarm algorithm, by introducing adaptive inertia weight and local mutation strategy, while enhancing the diversity of the population, it can ensure the local accurate search ability of the particle swarm. The simulation results of the example are verified. The effectiveness of the proposed model and method is confirmed.

**Keywords:** UAV colony; associated antagonists; dynamic non-zero-sum game; uncertain information; Nash equilibrium; interval probability



**Citation:** Wu, P.; Wang, H.; Liang, G.; Zhang, P. Research on Unmanned Aerial Vehicle Cluster Collaborative Countermeasures Based on Dynamic Non-Zero-Sum Game under Asymmetric and Uncertain Information. *Aerospace* **2023**, *10*, 711. <https://doi.org/10.3390/aerospace10080711>

Academic Editor: Gokhan Inalhan

Received: 27 June 2023

Revised: 31 July 2023

Accepted: 31 July 2023

Published: 15 August 2023



**Copyright:** © 2023 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

With the rapid development of UAV technology, UAVs have been widely used in agriculture, aerial photography, mapping, transportation and rescue fields [1]. On this basis, UAV cluster technology has also been developed and widely used, such as UAV cluster light show, UAV cluster confrontation and so on [2,3]. Compared with manned aircraft, UAVs are cheaper, smaller, and require less of a flight environment. UAV swarms can not only perform complex, diverse and dangerous missions in the traditional sense, such as architectural design, but also play an important role in responding to emergencies [4]. In the field of UAV cluster confrontation, currently related research is still in the early development stage, and its main difficulty lies in how to remove the traditional artificial path planning, and then realize the intelligent decision making and adaptive cooperation of the cluster itself. In addition, how to realize the optimization of the intelligent decision of cooperative maneuver confrontation in the dynamic confrontation process is also an important issue in the cooperative confrontation of UAV clusters.

From a systems science perspective [5], unmanned aerial vehicle (UAV) cluster systems are characterized by multi-platform heterogeneity, numerous task demands, input situation changes, complex tactical objectives, and coupled constraint conditions. To address these issues, it is necessary to design an autonomous decision-making and planning framework

for multi-task UAV clusters to reduce the complexity of system research. Reference [6] established a UAV cluster adversarial game model based on uncertain offensive and defensive situation information, and designed a game cost function to calculate the optimal strategy. Reference [7] proposed a multi-UAV distributed intelligent self-organization algorithm, which decomposes the optimization problem of the cluster reconnaissance-attack task into multiple local optimization problems, and realizes global optimization decision-making through information exchange between the cluster and the environment and within the cluster. Reference [8] used a deep learning method to construct a task decision-making model for typical cluster tasks such as area reconnaissance, and then optimized the decision-making model based on a genetic algorithm, providing effective support for offline learning and online decision-making of the cluster. However, existing research on UAV cluster autonomous decision-making problems is relatively scarce from a multi-task perspective.

In this paper, we apply the ideas of game theory to the UAV cluster control problem. Game theory is a modern scientific system that originated in the early 20th century and has developed into a complete and rich theoretical science after World War II. Its application to military operations has become a research hotspot for scholars at home and abroad [9–13]. Multi-UAV cooperation refers to two or more UAVs cooperating and coordinating with each other to accomplish tasks based on a certain type of motion [6]. Compared with one-on-one confrontations, the most significant difference in multi-UAV cooperative confrontations is that multiple task goals need to be addressed by allocating targets and firepower among various friendly UAVs according to our resources. However, one of the key issues in the successful completion of tasks by multiple UAVs is the problem of proper coordination between UAVs [14]. The hot research issue in the field of UAVs is how to use reasonable decision-making strategies to enable UAVs to coordinate with each other to complete complex tasks [15].

In 1998, Jun Lung Hu [16] proved that multi-intelligent collaboration converges to the Nash equilibrium point in a dynamic zero-sum game environment, which provides a theoretical basis for UAV cluster collaboration. At present, some constructive research results have been achieved in the field of UAV cluster collaborative countermeasures. Currently, there are four mainstream UAV cluster decision-making and control methods: expert system-based, population intelligence-based, neural network-based, and reinforcement learning-based [17]. Hubert H. Chin and Bechtel R J et al. [18,19] proposed an expert system-based air combat decision-making method that combines fuzzy logic and an expert knowledge base to help pilots make maneuver decisions. Bhattacharjee et al. [20] optimized multi-robot path planning by an artificial swarm method. In addition, particle swarm method [21], firefly method [22] and wolf pack method [23] are also widely used for UAV cluster coordination control. In recent years, UAV cluster coordinated countermeasures have started to be studied using the ideas of real-time analysis of posture and dynamic games. Shao et al. [4] used Bayesian inference to evaluate the air actual environment in real time by establishing a continuous decision process for multi-UAV cooperative air combat, and used the designed decision rules to make maneuver decisions. Chen Man et al. [24] built a game model for multi-drone cooperative combat missions by establishing the capability function of UAVs, and gave the finite strategy static game model and pure strategy Nash equilibrium solution method. From the above study, it can be seen that capability and real-time situational analysis for UAV clusters is the basis of the model for dynamic cooperative cluster confrontation [25].

By summarizing the above articles and existing research, the research in the field of UAV cooperative confrontation mainly has the following four characteristics:

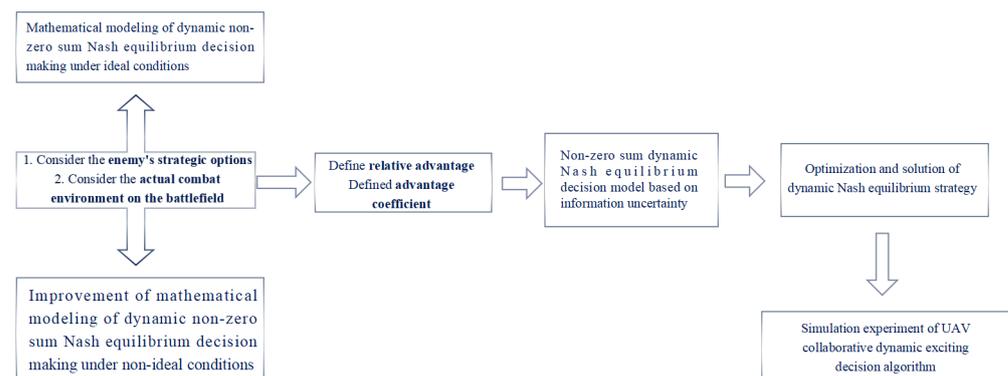
- (1). Most of the existing research in the field of UAV cooperative confrontation based on game theory assumes a zero-sum game. Characteristics of a zero-sum game.
- (2). Most of the existing research is carried out under the condition of certain information. However, due to the complexity, concealment and transient nature of the battlefield environment, the information required for most UAV operations is often uncertain. There-

fore, in the case of less information, considering the ambiguity of information in the process of UAV coordinated ground attack meets the actual combat needs.

(3). On the other hand, most of the existing research on UAV cooperative ground attack task assignment only focuses on unilateral action strategies, but due to the confrontational nature of the combat environment, it is very applicable to consider the opponent's possible defense strategies when attacking modeling and analysis through game theory.

(4). For the solution of the Nash equilibrium of the game model, the traditional intelligent optimization algorithm still has the shortcomings of weak global search ability and an easy to fall into local optimum, which urgently needs further optimization.

This article mainly discusses the above four issues, and logically it is a layer-by-layer progressive relationship, as shown in Figure 1, Section 2. The main work of this paper is to conduct multi-attribute evaluation and target strategy selection on the decision-making set through the situation analysis of both parties, and establish the dynamic non-zero and Nash equilibrium maneuver decision-making model of both parties. On the basis of the first section, the second section mainly introduces the improved dynamic non-zero and Nash equilibrium maneuver decision-making model after considering the ambiguity of information in actual combat and the possible defense strategy of the opponent. In the third section, the improved particle swarm optimization algorithm is used to solve the model proposed in the second section. The last section verifies the advantages of this model algorithm over traditional algorithms with a simulation comparison experiment.



**Figure 1.** The system architecture of the article.

Based on the above analysis, this paper aims at the cooperative confrontation problem of UAV clusters. First, under ideal conditions, that is, the information of both sides of the confrontation is completely accurate and known. Through the situation analysis of both sides, the multi-attribute evaluation of the decision set and the selection of target strategies are carried out to establish a dynamic non-cooperative relationship between the two sides. A zero-sum Nash equilibrium maneuver decision model. On this basis, the ideal conditions are modified, that is, considering the actual environment of the battlefield, including the performance difference of the drones of the two sides of the game, it is impossible for us to grasp the accurate information of the opponent in time, and because the radio environment of the battlefield is relatively complex, there may be data errors caused by ground information warfare forces interfering with our information acquisition. This paper improve the dynamic non-zero and Nash equilibrium maneuver decision-making model. Afterwards, through the improved particle swarm optimization algorithm, the efficient calculation of the Nash equilibrium solution of the non-zero-sum game model is realized, and the optimal mixed strategy of both parties is obtained. Finally, the effectiveness of the proposed method is verified by numerical simulation experiments.

## 2. Mathematical Modeling of Dynamic Non-Zero-Sum Game and Nash Equilibrium Decision-Making under Ideal Conditions

Before discussing the model in this section in detail, the following definitions are given first:

**Definition 1. Non-zero-sum game [26]:** Non-zero-sum game is opposite to zero-sum game. A zero-sum game means that the sum of the interests of all players in the game is fixed, that is, if one party gains something, the other party must lose something. A non-zero-sum game means that the sum of the gains of each player under different strategy combinations is an uncertain variable, also known as a variable-sum game.

**Definition 2. Nash Equilibrium [27] (Nash Equilibrium Point):** It refers to the non-cooperative game (Non-cooperative game) involving two or more participants, assuming that each participant knows the equilibrium strategy of other participants, a conceptual solution in which no player can benefit himself by changing his strategy.

**Definition 3. The non-zero-sum attack-defense game model (NADG),** is a quaternion  $NADG = \{W, M, S, WH\}$ :

$W$  and  $M$  are collectively referred to as participants, where  $W$  represents the attacker and  $M$  represents the defender.

$S = \{S_W, S_M\}$  is the strategy set. In this paper, the strategies of all participants are the same, which are maintaining the original flight state, accelerating, decelerating, turning left, turning right, climbing and diving.

$WH = \{WH_W, WH_M\}$  is the income function matrix of both attackers and defenders.

### 2.1. Model Assumptions

(1) Under ideal conditions, it is assumed that both parties can grasp all the complete and accurate information of the other party, and the information is instantaneous, and there is no time difference in the dissemination of information.

(2) Assuming that the objective conditions (including weather, equipment, etc.) are good, the onboard computer can accurately calculate the relevant data based on the global information, and the error is negligible.

(3) Assuming that the UAV cluster flight trajectory is discretized, that is, the confrontation trajectory between the two sides is composed of several maneuvers.

### 2.2. Set of Maneuvering Strategies for Both Sides of the Game

For dynamic UAV confrontation, the strategy set of dynamic non-zero-sum game Nash equilibrium maneuvering decision needs to be established. The red and blue UAV swarms are named  $W$  and  $M$ , respectively, where the  $W$  UAV swarm consists of  $p$  individual UAVs and the  $M$  UAV swarm consists of  $q$  individual UAVs. In the process of confrontation between the two sides, by assumption (3), their flight trajectories can be considered as a combination of several consecutive maneuvers, which are maintaining the original flight state, accelerating, decelerating, turning left, turning right, climbing, and diving, in order to be recorded as  $c_1, c_2, \dots, c_7$ . In summary, the overall strategy set of the drone swarm can be expressed as

$$\begin{aligned} S_M &= \{S_{M_1}, S_{M_2}, \dots, S_{M_7}\} \\ S_W &= \{S_{W_1}, S_{W_2}, \dots, S_{W_7}\} \end{aligned} \quad (1)$$

where the maneuvering strategy of individual square drones is  $S_W = S_M = \{c_1, c_2, \dots, c_7\}$ . Since the strategy number of each individual drone is 7, the strategy numbers of drone group  $W$  and drone group  $M$  are  $n_W = 7^p$  and  $n_M = 7^q$ , respectively.

### 2.3. Maneuver State Assessment of Both Sides of the Game

In the text, based on the research of Zhang Shuo et al. [28], the situational advantage equation function is used to evaluate and qualitatively describe the state of maneuvering attributes of both players in the game. Assume that in the actual combat process, the maneuver attributes are mainly characterized qualitatively by distance advantage, speed advantage, and angle advantage, and each advantage is characterized by the existing situational advantage equation. By weighting and summing the above three advantages, the dominant state of the UAV group at a certain moment can be obtained.

Record the maneuvering attribute as  $Q$ , distance advantage  $D$ , speed advantage  $V$ , and angle advantage  $A$ .

$$Q = \{D, V, A\} \tag{2}$$

**Distance advantage**

Define  $M_D$  as the speed advantage of a single existing drone over a single enemy drone:

$$M_D = e^{-\left(\frac{D-R_0}{R_{\min}-R_{\max}}\right)^2} \tag{3}$$

where  $D$  is the Euclidean distance between the two sides of the game,  $R_{\max}$  is the maximum starting distance,  $R_{\min}$  is the minimum starting distance,  $R_0 = (R_{\min} + R_{\max})/2$  where  $R_{\max}, R_{\min}$  are, respectively, determined by the attribute parameters of the UAVs of both sides of the game.

**Speed advantage**

Define  $M_V$  as the speed advantage of your own single UAV over the target UAV:

$$M_V = \begin{cases} 0.1 & 0.1V_{W_i} \leq 0.6V_{M_j}, \\ \frac{V_{W_i}}{V_{M_j}} - 0.5 & 0.6V_{M_j} \leq V_{W_i} \leq 1.5V_{M_j}, \\ 1 & V_{W_i} \geq 1.5V_{M_j}. \end{cases} \tag{4}$$

where  $V_{w_i}$  is the individual flight speed of  $W$  UAV group and  $V_{M_j}$  is the individual flight speed of  $M$  UAV group.

From the above speed advantage function, it can be seen that when the speed  $V_{W_i}$  of the own drone is greater than 1.5 times the speed  $V_{M_j}$  of the target drone, the speed advantage  $M_V$  is the largest.

**Angle advantage**

Define  $M_A$  as the speed advantage of a single UAV of one’s own side relative to a single UAV of the enemy.

$$M_A = \frac{\frac{A_l}{180^\circ} - \frac{A_v}{180^\circ} + 1}{2} \tag{5}$$

Among them,  $A_l$  is the incident angle of the target, and  $A_v$  is the angle of view of the own side’s single UAV.  $A_l$  and  $A_v$  can be calculated from the real-time position, declination angle, pitch angle and other information of both parties. From the above angle advantage function, it can be found that when the target incident angle  $A_l$  is larger and the viewing angle  $A_v$  is smaller, the angle advantage  $M_A$  is greater.

**2.4. Establishment of the Overall Dynamics and Payoff Matrix of a Single Game per Unit of Time**

**2.4.1. Overall Posture Matrix**

From Section 2.2, the number of strategies for UAV group  $W$  and UAV group  $M$  are  $n_W = 7^P$  and  $= 7^P$ , respectively. When  $W$  adopts the  $l$  ( $l = 1, 2, \dots, n_W$ ) strategy and  $M$  adopts the  $m$ th ( $m = 1, 2, \dots, n_W$ ) strategy, the  $i$ th ( $i = 1, 2, \dots, p$ ) individual UAV in  $W$  has the distance advantage, speed advantage and angle advantage for the  $j$ th ( $j = 1, 2, \dots, q$ ) individual UAV in  $M$ . The distance advantage, speed advantage and angle advantage of the individual UAV are  $WM_{D(i,j)}, WM_{V(i,j)}, WM_{A(i,j)}$ , respectively. The overall posture can be obtained as follows.

$$WX_{(i,j)} = Wk_1M_D(i,j) + Wk_2M_V(i,j) + Wk_3M_A(i,j) \tag{6}$$

$$i = 1, 2, \dots, p, j = 1, 2, \dots, q$$

where  $Wk_1Wk_2Wk_3$  are the overall posture weighting parameters, and  $Wk_1 + Wk_2 + Wk_3 = 1$ .

Therefore, the matrix  $WX$  is the overall posture matrix of the UAV swarm  $W$  when the UAV on the  $W$  side adopts the 1st strategy and the UAV on the  $M$  side adopts the  $m$  strategy.

The matrix  $WX$  is the overall posture of  $W$  to  $M$ . Similarly, the matrix  $MX$  can be built in the same steps.  $j$ th individual of  $M(j = 1, 2, \dots, q)$ , its distance advantage, speed advantage and angle advantage to  $i$ th ( $i = 1, 2, \dots, p$ ) individual UAVs of UAV group  $W$  are  $MM_{D(i,j)}, MM_{V(i,j)}, MM_{A(i,j)}$ , and the overall posture can be calculated as

$$MX_{(j,i)} = Mk_1M_D(j, i) + Mk_2M_V(j, i) + Mk_3M_A(j, i) \tag{7}$$

$$i = 1, 2, \dots, p, j = 1, 2, \dots, q$$

where  $Mk_1Mk_2Mk_3$  are the overall posture weighting parameters, respectively, and  $Mk_1 + Mk_2 + Mk_3 = 1$ .

Therefore, matrix  $WX$  is the overall posture matrix of UAV swarm  $M$  when UAV on  $M$  side adopts the  $m$  strategy and UAV on  $W$  side adopts the  $l$ st strategy. Therefore, the matrix  $WX$  is the overall situation matrix for the UAV group  $M$  when the  $M$ -side UAV adopts the  $M$ th-th strategy and the  $W$ -side UAV adopts the  $l$ th strategy.

### 2.4.2. Overall Payoff Matrix

Based on the establishment of the overall situation matrix in the previous section, the income matrices (also called payment matrices) of  $W$  and  $M$  are, respectively, established according to different strategy sets. In actual scenarios, cluster confrontation strategies can be divided into global confrontation strategies, local confrontation strategies, global penetration strategies and local penetration strategies. The profit matrix under different strategies is different. This article mainly considers the global object, so the following mainly uses  $W$  as an example to introduce the establishment process of the profit matrix under the global confrontation strategy and the profit matrix under the global penetration strategy.

- (1) Global confrontation strategy. The main objective of this strategy is to optimize the overall posture of our side, and according to this objective, the gain matrix of  $W$  can be obtained:

$$WH(l, m) = \frac{1}{pq} \sum_{i=1}^p \sum_{j=1}^q WX_{l,m}(i, j). \tag{8}$$

$$l = 1, 2, \dots, n_W, m = 1, 2, \dots, n_M.$$

- (2) Global penetration strategy. The main objective of this strategy is the worst overall posture of the opponent, and according to this objective, the gain matrix of  $W$  can be obtained:

$$WH(l, m) = -\frac{1}{pq} \sum_{i=1}^p \sum_{j=1}^q MX_{m,l}(i, j). \tag{9}$$

$$l = 1, 2, \dots, n_W, m = 1, 2, \dots, n_M.$$

Similarly, the payoff matrix of  $M$  under different strategies can be built based on the above steps. Similarly, in this paper, only the gain matrices of the global confrontation strategy and the global surprise strategy of  $M$  are illustrated.

- (3) Global confrontation strategy. The main objective of this strategy is to optimize the overall posture of our side, and according to this objective, the gain matrix of  $M$  can be obtained:

$$MH(m, l) = \frac{1}{pq} \sum_{i=1}^p \sum_{j=1}^q WX_{l,m}(j, i). \tag{10}$$

$$l = 1, 2, \dots, n_W, m = 1, 2, \dots, n_M.$$

- (4) Global penetration strategy. The main objective of this strategy is the worst overall posture of the opponent, and according to this objective, the gain matrix of  $M$  can be obtained:

$$MH(m, l) = -\frac{1}{pq} \sum_{i=1}^p \sum_{j=1}^q MX_{m,l}(j, i). \quad (11)$$

$$l = 1, 2, \dots, n_W, \quad m = 1, 2, \dots, n_M.$$

In summary, the preparation of the mathematical model for dynamic nonzero and Nash equilibrium decision making under ideal conditions is basically completed. Since the model in this section is based on the assumption of ideal conditions, the mathematical model will be improved in the next section of this paper to make it closer to the actual situation, which in turn makes the decision more accurate and informative.

### 3. Improvement of Mathematical Modeling of Dynamic Non-Zero and Nash Equilibrium Decision Making under Non-Ideal Conditions

#### 3.1. Problem Analysis

##### 3.1.1. Consider the Enemy's Strategy Choice

Now consider the following situation, when our UAV group adopts two different strategies, A and B, the maneuver advantage over the enemy is 100, and the enemy responds to our A and B with the same maneuver, respectively. In terms of strategy, the maneuvering advantages relative to our side are 60 and 80, respectively. According to the above case analysis, if we adopt different strategies based solely on our maneuvering advantage over the enemy, there is no difference between A and B strategies at this time. But in fact, when strategy A is adopted, the enemy's threat to our side is relatively small, and our optimal strategy should adopt strategy A. Therefore, based on the simple analysis of the above examples, the own side should consider the enemy's strategy when making a decision.

##### 3.1.2. Considering the Actual Battlefield Operational Environment

The model in Section 2 is based on two assumptions. However, in actual conditions, due to the performance differences between the UAVs of the two sides of the game, it is impossible to grasp the accurate information of the other side in time, and because of the more complex radio environment on the real field and because the other party may have information jamming technology, there may be errors in the acceptance of the other party information by the UAV of the already side. In summary, the assumptions of model 1 do not meet the actual combat situation and will be improved in this section.

#### 3.2. Model Revision

##### 3.2.1. Model Correction Based on Information Uncertainty

Under the assumptions in the previous section, each element of the gain matrix is a definite value, but in reality, there is a certain error in the transmission and acquisition of information, so there is a certain error in the information received by the UAV. In order to be closer to the actual situation, when calculating the posture value using the dominance function, an error factor (sign) is introduced, and (2) (3) (4) is amended to obtain the following new posture dominance function.

$$\tilde{M}_D = e^{-\left(\frac{D-R_0}{R_{\max}-R_{\min}}\right)^2} \alpha$$

$$\tilde{M}_D \in (M_{D \min}, M_{D \max}) \quad (12)$$

$$M_{D \min} = M_D \cdot \alpha_{\min}, \quad M_{D \max} = M_D \cdot \alpha_{\max}$$

where  $\alpha$  is the error factor of distance dominance,  $\tilde{M}_D$  is the probability interval information of distance dominance,  $\alpha_{\min}$  is the lower limit of error,  $\alpha_{\max}$  is the upper limit of error,  $M_{D \min}$  is the lower limit of the large probability value of distance dominance, and  $M_{D \max}$  is the upper limit of the large probability value of distance dominance. The

distance dominance accuracy values are taken randomly on this interval and obey uniform distribution.

$$\begin{aligned} \tilde{M}_V &= M_V \beta \\ \tilde{M}_V &\in (M_{V \min}, M_{V \max}) \\ M_{V \min} &= M_V \cdot \beta_{\min}, M_{V \max} = M_V \cdot \beta_{\max} \end{aligned} \tag{13}$$

where  $\beta$  is the error factor of speed advantage,  $\tilde{M}_V$  is the probability interval information of distance advantage,  $\beta_{\min}$  is the lower error limit,  $\beta_{\max}$  is the upper error limit,  $M_{V \min}$  is the lower limit of the large probability value of speed advantage, and  $M_{V \max}$  is the lower limit of the large probability value of speed advantage. The accurate value of speed advantage is taken randomly on this interval and obeys uniform distribution.

$$\begin{aligned} \tilde{M}_A &= [(A_1/180^\circ) - (A_v/180^\circ) + 1]/2\gamma \\ \tilde{M}_A &\in (M_{A \min}, M_{A \max}) \\ M_{A \min} &= M_A \cdot \gamma_{\min}, M_{A \max} = M_A \cdot \gamma_{\max} \end{aligned} \tag{14}$$

After the above corrections, a new overall posture dominance function is obtained.

$$\tilde{T} = k_1 \tilde{M}_D + k_2 \tilde{M}_V + k_3 \tilde{M}_A \tag{15}$$

Since  $\tilde{M}_D, \tilde{M}_V, \tilde{M}_A$  are all interval numbers,  $\tilde{T}$  is also an interval number.

### 3.2.2. Non-Zero and Dynamic Nash Equilibrium Decision Model Based on Information Uncertainty

The analysis in this section is based on the assumptions of Section 3.2.1. Due to information uncertainty, each element in the resulting payoff matrix is an interval number, and the matrix  $\tilde{W}L$  is the matrix of the overall posture function when  $W$  takes the  $m$  strategy and  $M$  takes the 1st strategy as follows.

$$\tilde{W}L = \begin{pmatrix} \tilde{W}L_{a11} & \tilde{W}L_{a12} & \cdots & \tilde{W}L_{a1n} \\ \tilde{W}L_{a21} & \tilde{W}L_{a22} & \cdots & \tilde{W}L_{a2n} \\ \vdots & \vdots & \cdots & \vdots \\ \tilde{W}L_{an1} & \tilde{W}L_{an2} & \cdots & \tilde{W}L_{apq} \end{pmatrix} \tag{16}$$

$$= \begin{pmatrix} (WL_{\min}^{a11}, WL_{\max}^{a11}) & (WL_{\min}^{a12}, WL_{\max}^{a12}) & \cdots & (WL_{\min}^{a1n}, WL_{\max}^{a1n}) \\ (WL_{\min}^{a21}, WL_{\max}^{a21}) & (WL_{\min}^{a22}, WL_{\max}^{a22}) & \cdots & (WL_{\min}^{a2n}, WL_{\max}^{a2n}) \\ \vdots & \vdots & \vdots & \vdots \\ (WL_{\min}^{an1}, WL_{\max}^{an1}) & (WL_{\min}^{an2}, WL_{\max}^{an2}) & \cdots & (WL_{\min}^{ann}, WL_{\max}^{ann}) \end{pmatrix} \tag{17}$$

The payoff interval matrix is built on the basis of the overall posture matrix, and the specific process is similar to Section 2.4.2 and will not be repeated here, the payoff matrix is as follows.

$$\tilde{W}H = \begin{pmatrix} \tilde{W}H_{a11} & \tilde{W}H_{a12} & \cdots & \tilde{W}H_{a1n} \\ \tilde{W}H_{a21} & \tilde{W}H_{a22} & \cdots & \tilde{W}H_{a2n} \\ \vdots & \vdots & \cdots & \vdots \\ \tilde{W}H_{an1} & \tilde{W}H_{an2} & \cdots & \tilde{W}H_{apq} \end{pmatrix} \tag{18}$$

After obtaining the gain interval matrix, based on the consideration of Section 3.2.1, in order to better measure the true posture of the already side's UAV, not only the gain interval matrix of the already side is needed, but also the gain interval matrix of the opposite side  $\tilde{M}H$  is needed to calculate the gain interval matrix of the enemy, and  $WH_{ij}$  and  $MH_{ji}$  are analyzed and compared, and the specific analysis and comparison process is as follows: according to the matrix (true value matrix),  $WH_{ij}$  can be compared with  $MH_{ji}$  in two cases,

which are if  $WH_{ij}$  is greater than  $MH_{ji}$  and  $WH_{ij}$  is less than  $MH_{ji}$ . If  $WH_{ij}$  is greater than  $MH_{ji}$ , there are three more cases as follows [3].

**Definition 4. Relative dominance:** Assuming that the upper limit of the interval of  $\widetilde{WH}_{ij}$  is higher than the upper limit of the interval of  $\widetilde{MH}_{ij}$ , there are only three cases:  $\widetilde{WH}_{ij}$  and  $\widetilde{MH}_{ij}$  have no intersection (Figure 2a),  $\widetilde{WH}_{ij}$  and  $\widetilde{MH}_{ij}$  have intersection and B is not completely contained in  $\widetilde{WH}_{ij}$  (Figure 2b),  $\widetilde{MH}_{ij}$  is completely contained in  $\widetilde{WH}_{ij}$  (Figure 2c). According to these three situations, this paper defines the result calculated by Equation (19) as the relative dominance of  $\widetilde{WH}_{ij}$  to  $\widetilde{MH}_{ij}$ .

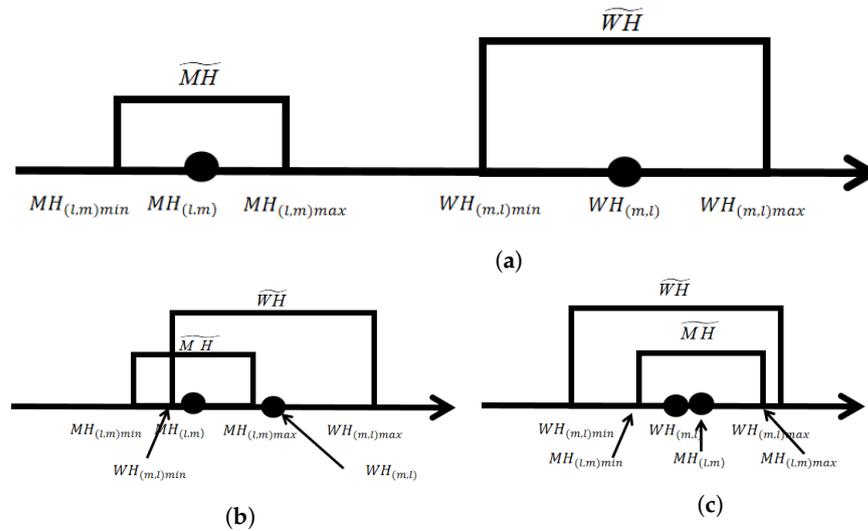


Figure 2. Three situations of interval  $MH$  relative to interval  $WH$ . (a–c) see Definition 4.

$$P_{\widetilde{WH}_{ij} > \widetilde{MH}_{ji}} = \begin{cases} 1, & MH_{ji} \max \leq WH_{ij} \min; \\ \frac{\delta_{\max}^1 - \delta_{\max}^2}{\delta_{\max}^1 - \delta_{\min}^1} + \frac{\delta_{\max}^2 - \delta_{\min}^1}{\delta_{\max}^1 - \delta_{\min}^1} \cdot \frac{\delta_{\min}^1 - \delta_{\min}^2}{\delta_{\max}^2 - \delta_{\min}^2} + 0.5 \frac{\delta_{\max}^2 - \delta_{\min}^1}{\delta_{\max}^1 - \delta_{\min}^1} \cdot \frac{\delta_{\max}^2 - \delta_{\min}^1}{\delta_{\max}^2 - \delta_{\min}^2}, & MH_{ji} \min < WH_{ij} \min < MH_{ji} \max \\ \frac{\delta_{\max}^1 - \delta_{\max}^2}{\delta_{\max}^1 - \delta_{\min}^1} + 0.5 \frac{\delta_{\max}^2 - \delta_{\min}^1}{\delta_{\max}^1 - \delta_{\min}^1}, & WH_{ij} \min \leq MH_{ji} \min < MH_{ji} \max \end{cases} \quad (19)$$

**Definition 5. Dominance coefficient,** a value measuring the degree of dominance of an existing UAV over an opposing UAV, judged by the degree of relative dominance, calculated as follows.

$$\psi_{ij} = \begin{cases} 1, & 0.9 < H_{ij} < 1; \\ 0.9, & 0.6 < H_{ij} < 0.9; \\ 0.8, & 0 < H_{ij} < 0.6 \end{cases} \quad (20)$$

According to the definition of the dominance coefficient, if the dominance coefficient is 1, when both sides take the 1st strategy the  $m$ th strategy, respectively, under the premise of our dominance, the overall posture of the already side is absolutely dominant relative to the overall posture of the enemy, and the opposite side is less threatening to us when we complete the 1st strategy; if the dominance coefficient is 0.9, it means that the overall posture of the already side is relatively dominant relative to the overall posture of the opposite side, and if the dominance coefficient is 0.8, it means that the overall posture of the opponent is generally superior to the overall posture of the opposite side, and the opponent needs to pay a larger price while completing their tactical moves.

According to Equations (10) and (16), we can obtain the gain interval matrix  $\widetilde{WH}$  and  $\widetilde{MH}$  for  $W$  and  $M$ , respectively, after which we can derive the dominance coefficient matrix of the already party’s UAV group  $W$  according to Equations (17) and (18). Using the

dominance coefficients, the matrix  $WH$  can be modified, which in turn yields the matrix  $\widehat{WH}$  as follows.

$$\widehat{WH} = \begin{pmatrix} \widehat{WH}_{11} & \widehat{WH}_{12} & \cdots & \widehat{WH}_{1n_M} \\ \widehat{WH}_{21} & \widehat{WH}_{22} & \cdots & \widehat{WH}_{2n_M} \\ \vdots & \vdots & \cdots & \vdots \\ \widehat{WH}_{n_W1} & \widehat{WH}_{n_W2} & \cdots & \widehat{WH}_{n_Wn_M} \end{pmatrix} \tag{21}$$

$$= \begin{pmatrix} WH_{11} \cdot \psi_{11} & WH_{12} \cdot \psi_{12} & \cdots & WH_{1n_M} \cdot \psi_{n_M} \\ WH_{21} \cdot \psi_{21} & WH_{22} \cdot \psi_{22} & \cdots & WH_{2n_M} \cdot \psi_{2n_M} \\ \vdots & \vdots & \cdots & \vdots \\ WH_{n_W1} \cdot \psi_{n_W1} & WH_{n_W2} \cdot \psi_{n_W2} & \cdots & WH_{n_Wn_M} \cdot \psi_{n_Wn_M} \end{pmatrix} \tag{22}$$

After building the improved payoff matrices  $\widehat{WH}$  and  $\widehat{WH}$  for  $W$  and  $M$ , respectively, the confrontation between the two sides of the game is a typical non-zero-sum game due to the different nature and purpose of the drone swarms on both sides of the confrontation

**Definition 6 ([2]). Mixed strategy Nash equilibrium:** Let a non-cooperative game in which there are  $n$  insiders involved, where the pure strategy of insider  $i$  is denoted as  $s^i = \{s_1^i, s_2^i, \dots, s_{m_i}^i\}$  and the mixed strategy of  $i$  is defined as  $x_i = \{(x_{i1}, x_{i2}, \dots, x_{im_i}) \mid x_{ij} \geq 0, \sum_{j=1}^{m_{ij}} X_{ij} = 1\}$ , i.e., the inning chooses the  $j$ th strategy with the probability of  $x_{ij}$ . If a mixed strategy combination  $X^* = \{x_1^*, x_2^*, \dots, x_n^*\}$  satisfies  $\mu_i(X^*) \geq \mu_i(x_i, X_{-i}^*) (i = 1, 2, \dots, n)$ , where  $\mu_i$  denotes the payoff function of  $i$  and  $X_{-i}^* = \{x_1^*, x_2^*, \dots, x_{i-1}^*, x_{i+1}^*, \dots, x_n^*\}$ , i.e., a single inning changes its strategy and its payoff value does not increase, then  $X^*$  is said to be the Nash equilibrium of this non-zero-sum game.

From Definition 6, let the pure strategy sets of  $W$  and  $M$  be  $S_l^W \in S_W$  and  $S_m^M \in S_M$ , respectively, and the probabilities of individual UAVs of the already party to choose the corresponding strategies are  $x_l (l = 1, 2, \dots, n_W)$  and  $y_m (m = 1, 2, \dots, n_G)$ . In summary, the mixed strategy of  $W$  and  $G$  can be expressed in the following form.

$$X = \left\{ x \in R^{n_W} \mid \sum_{l=1}^{n_W} x_l = 1, x_l > 0, l = 1, 2, \dots, n_W \right\} \tag{23}$$

$$Y = \left\{ x \in R^{n_M} \mid \sum_{m=1}^{n_M} x_m = 1, x_m > 0, m = 1, 2, \dots, n_M \right\},$$

Based on the characteristics of non-zero-sum games, the following theorems exist.

**Theorem 1 ([29]).** For any mixed strategy  $(X, Y)$ , there exists a Nash equilibrium solution  $(X^*, Y^*)$  satisfying the following conditions.

$$\begin{aligned} X^T F H Y^* &\leq X^{*T} F H Y^*, \\ X^{*T} F H^T Y &\leq X^{*T} F H^T Y^* \end{aligned} \tag{24}$$

The Nash equilibrium solution obtained at this point is the optimal strategy of the nonzero-sum game for  $W$  and  $M$ . In order to quickly solve the Nash equilibrium solution satisfying Equation (15), Equation (15) is transformed into the following optimization problem.

$$\begin{aligned} &\min E^*(X^*, Y^*) \\ \{ X^* = \{ x^* \in R^{n_F} \mid \sum_{l=1}^{n_F} x_l^* = 1, x_l^* \geq 0, l = 1, 2, \dots, n_F \}, \\ Y^* = \{ y^* \in R^{n_G} \mid \sum_{m=1}^{n_G} y_m^* = 1, y_m^* \geq 0, m = 1, 2, \dots, n_G \}, \end{aligned} \tag{25}$$

$$E^*(X^*, Y^*) = \max\left\{\max_{l=1,2,\dots,n_F}\left(FH_l Y^* - X^{*T} F H Y^*\right), 0\right\} + \max\left\{\max_{m=1,2,\dots,n_G}\left(X^{*T} G H_m^T - X^{*T} G H^T Y^*\right), 0\right\} \quad (26)$$

where the mixed strategy in which the function  $E(X^*, Y^*)$  or  $E^*(X^*, Y^*)$  takes the minimum value of 0 is the Nash equilibrium point of the original non-zero-sum game problem.

### 3.3. Dynamic Nash Equilibrium Decision Model

The non-zero-sum game process of W and M within the unit process is given in the first two sections, and in the actual situation, the two drone swarms are a dynamic game process, which is combined via several unit processes, so a dynamic Nash equilibrium decision model is needed. The specific steps of the model are as follows.

(1) Set the simulation parameters and initial conditions according to the actual situation of the UAVs on both sides of the game, and specify the unit step size and the maximum number of iterations.

(2) According to the real-time state parameters of both sides, under the combination of different strategies, the overall state advantage function of individual UAVs is calculated according to Equations (6)–(9), so as to obtain the gain matrix of both sides and finally obtain the single-step non-zero-sum game model.

(3) Calculate the Nash equilibrium solution of the non-zero-sum game according to Equations (15) and (16), and obtain the optimal hybrid strategy of both sides of the single-step game.

(4) Calculating the position and other state parameters of each UAV where it will be located next from the step length specified in step (1).

(5) According to the state parameters calculated in step (4), determine whether the game is over, if the conditions for the game to proceed are still met, return to step (2), if the conditions for the game to end have been met, proceed to step (6).

(6) The game ends, and the results of both sides of the game are derived.

## 4. Optimization and Solution of Dynamic Nash Equilibrium Strategy

For the intelligent algorithm solution of non-cooperative game model Nash equilibrium, predecessors have carried out a lot of related work. The research results show that the PSO algorithm is excellent in finding Nash equilibrium [30]. Therefore, this paper uses the PSO algorithm to solve the Nash equilibrium. In view of the disadvantages of the classic PSO algorithm, such as premature convergence and low precision, this paper makes certain improvements to the PSO algorithm.

### 4.1. Classical Particle Swarm Algorithm

Section 3.3 gives the Nash equilibrium maneuver decision model and its solution process for the dynamic non-zero-sum game of cooperative UAV swarm confrontation under asymmetric uncertain information, the core of which is in the solution of the optimization problem Equation (15) with step (3). Therefore, how to find the optimal solution of Equation (15) quickly and efficiently becomes the key to the optimization of dynamic non-zero-sum Nash equilibrium strategies.

In this paper, the search capability of PSO is improved by optimizing the adjustment of inertia weights in it based on the use of the classical particle swarm optimization (PSO) algorithm, which is an optimization algorithm that simulates the foraging process of a flock of birds, where the potential solution of each optimization problem is a bird in the search space, called a particle.

Assuming that the search space is a multidimensional space of  $D$  dimensions, the velocity vector and position vector of a particle can be defined as follows, respectively:

$$\begin{aligned} X_i &= (x_{i1}, x_{i2}, \dots, x_{iD}), i = 1, 2, \dots, N \\ V_i &= (v_{i1}, v_{i2}, \dots, v_{iD}), i = 1, 2, \dots, N \end{aligned} \quad (27)$$

The equation for its velocity and position update is

$$\begin{aligned} v_{id} &= \omega v_{id} + c_1 r_1 (p_{id} - x_{id}) + c_2 r_2 (p_{gd} - x_{gd}) \\ x_{id} &= x_{id} + v_{id} \end{aligned} \quad (28)$$

where  $c_1, c_2$  are called learning factors, also known as acceleration constants,  $r_1, r_2$  are random numbers in the range  $[0, 1]$ ,  $P_{id}$  is the optimal position searched so far by the  $i$ th particle, i.e., the individual extremum, and  $P_{gd}$  is the optimal position searched so far by the whole particle population, i.e., the global extremum, and  $\omega$  is the inertia weight.

#### 4.2. Control of Inertia Weight of Particle Swarm Algorithm

As can be seen from the above equation, the inertia weight controls the influence of the previous variable on the current variable. If  $\omega$  is larger, it can search the area that the particle failed to reach before, which makes the global search ability of the whole algorithm enhanced, on the contrary, if  $\omega$  is smaller, the particle mainly searches within the area of the current solution, and the local search ability is enhanced. Therefore, in the field environment, the need to solve the Nash equilibrium solution quickly and accurately is crucial to the field situation, and in the classical PSO,  $\omega$  is a constant, which cannot adapt to the needs of the solution in different situations, so it needs to be improved.

There are three common methods to improve  $\omega$  in PSO, adaptive weight method, random weight method and linear decreasing weight method.

##### 4.2.1. Linear Decreasing Method

This method addresses the phenomenon that the PSO algorithm is prone to premature maturity and to oscillation near the global optimal solution at a later stage, even though the inertia weights are decreasing one at a time in accordance with linearity from large to small, and the variation formula is

$$\omega = \omega_{\max} - \frac{t \times (\omega_{\max} - \omega_{\min})}{t_{\max}} \quad (29)$$

where  $\omega_{\max}$  is the maximum value of inertia weight,  $\omega_{\min}$  is the minimum value of inertia weight,  $t_{\max}$  is the total number of iteration steps, and  $t$  is the number of current iteration.

##### 4.2.2. Adaptive Modification Weighting Method

###### 1. Adjustment according to the distance of global optimum

According to previous studies, the size of the inertia weight is considered to be related to its distance from the global optimum in some papers, and related papers suggest that the inertia weights of different particles proposed in the previous section not only decrease linearly with the increase in the number of iterations, but also decrease with the increase in the distance from the global optimum solution, because the probability of searching for a better solution is greater when the particle is closer to the global optimum solution, so it is necessary to decrease the inertia weight to achieve the purpose of enhancing the local search ability of the particle. In summary, the inertia weights change dynamically depending on the particle positions. And the linear equations generally cannot meet the requirements, so most of them currently use the coefficient formula of nonlinear dynamic inertia weights, which is as follows.

$$\omega = \begin{cases} \omega_{\min} - \frac{(\omega_{\max} - \omega_{\min}) \times (f - f_{\min})}{f_{\text{avg}} - f_{\min}}, & f \leq f_{\text{avg}} \\ \omega_{\max}, & f > f_{\text{avg}} \end{cases} \quad (30)$$

where  $f$  is the objective function value of the current step,  $f_{\text{avg}}$  and  $f_{\min}$  are the minimum and average values of the objective function values of all particles in the current step, respectively.

As can be seen from the above equation,  $\omega$  is constantly adjusted as the dispersion of particles changes, and the inertia weight will be reduced when the target value of particles is more dispersed; when the target value of particles is more concentrated, the inertia weight increases.

2. Adjust the weights according to the degree of early convergence and adaptation value.

The improvement method adjusts its own weights according to the degree of early proficiency of the particle population and the adaptation value of the objective function of each particle, as follows.

$$\omega = \begin{cases} \omega - (\omega - \omega_{\min}) \left| \frac{f_i - f'_{\text{avg}}}{f_m - f'_{\text{avg}}} \right| & f_m > f_i > f'_{\text{avg}} \\ \omega, & f_m > f_i > f'_{\text{avg}} \\ 1.5 - \frac{1}{1+k_1 \cdot \exp(-k_2 \cdot \Delta)}, & f_i > f'_{\text{avg}} \end{cases} \quad (31)$$

where  $f_i$  is the adaptation value of particle  $p_i$ , and  $f_m$  is the optimal particle fitness. Then, the average fitness of the particle swarm is  $f_{\text{avg}} = \frac{1}{n} \sum_{i=1}^n f_i$ . The particle fitness values that are better than the average fitness values are averaged, denoted as  $f'_{\text{avg}}$ , and define  $\Delta = |f_m - f'_{\text{avg}}|$ .  $k_1, k_2$  are called regulation parameters,  $k_1$  is used to regulate the upper limit of  $\omega$  and  $k_2$  is used to control  $\omega = 1.5 - \frac{1}{1+k_1 \cdot \exp(-k_2 \cdot \Delta)}$  of the regulation capacity.

From the definition of  $\Delta$ , it is known that if the particles in the swarm are too dispersed,  $\Delta$  becomes larger, thus  $\omega$  decreases, which enhances the local search ability of the swarm; if the particles in the swarm are too concentrated,  $\Delta$  becomes smaller, thus  $\omega$  increases, which enhances the global search ability of the swarm to help jump out the local optimal solution effectively.

In this paper, the adaptive modified weight method is used to improve the PSO algorithm.

## 5. Simulation Experiments of UAV Cooperative Dynamic Maneuver Decision Algorithm

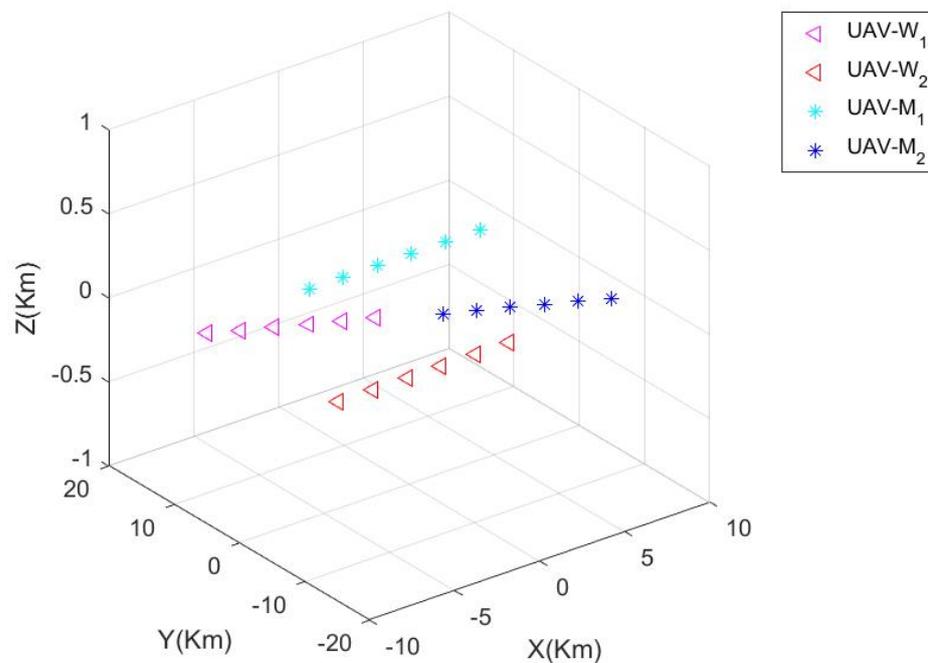
Based on the UAV cluster adversarial algorithm model based on asymmetric uncertain information environment proposed in the previous two sections, numerical simulation experimental results are presented in this section to verify the validity of the model. The simulation experiments apply the game confrontation steps in Section 3.3, in which the adaptive algorithm in Section 4.2.2 is used to optimize the speed of the algorithm in “calculating the Nash equilibrium solution of the non-zero-sum game and obtaining the optimal hybrid strategy for both sides of the single-step game”.

In order to compare the superiority of the proposed algorithm,  $W$  uses the “global adversarial strategy” of the UAV cluster adversarial algorithm based on asymmetric uncertain information environment proposed in this paper, and  $M$  uses the classical global adversarial strategy based on the maximum–minimum pure strategy [31]. The following Table 1 shows the initial parameters of the numerical simulation experiment.

Substituting the initial conditions into Equations (3)–(5), we can see that  $M$  has a clear advantage in angular posture in the initial stage of the confrontation. In the total 40 steps of the confrontation, they can be roughly divided into “posture equilibrium phase”, “cooperative confrontation phase” and “absolute advantage phase”, and part of the confrontation process is shown in Figures 3–8. The red “ $\Delta$ ” is the path trajectory position of  $W$  at each step, and the blue “\*” is the path trajectory position of  $M$  at each step. The confrontation ends when one side has an absolute advantage over the other side or reaches the maximum number of confrontation steps.

**Table 1.** Numerical experiments against double-fire initial basic parameters.

Parameters		W1	W2	M1	M2
Drone Number					
Initial location		(−400 m, 0 m, 1000 m)	(−400 m, 200 m, 1000 m)	(400 m, 0 m, 1000 m)	(400 m, 200 m, 1000 m)
Initial speed		200 m/s	200 m/s	200 m/s	200 m/s
Initial deflection angle		60°	60°	−118°	−118°
Initial pitch angle		3°	3°	5°	5°
Acceleration			60 m/s		60 m/s
Deceleration			60 m/s		60 m/s
Time per unit	Maximum deflection angle variable		30°		30°
	Maximum pitch angle variable		10°		10°
	Overall situational weighting parameters	Wk1 = 0.3, Wk2 = 0.2, Wk3 = 0.5		Mk1 = 0.3, Mk2 = 0.2, Mk3 = 0.5	
	Confrontation step limit	40 step			
	Unit decision time	1 s			



**Figure 3.** Simulation of UAV cooperative dynamic maneuvering decision algorithm (step 5).

The first stage is the situational balance stage (see Figures 3 and 4). The initial positions of the two sides are far apart. In the fifth step, the two sides enter the opponent’s combat radius, and a confrontation situation occurs. W<sub>1</sub> and M<sub>1</sub> are close to each other and confront each other, and W<sub>2</sub> and M<sub>2</sub> approach each other against each other, as shown in Figure 3.

When the first stage reaches the 10th step, W<sub>1</sub> adjusts the pitch angle by a large margin in order to increase the advantage of the situation, and quickly raises the height of the UAV. Since M<sub>1</sub> is relatively close to W<sub>1</sub>, in order to maintain the overall advantage of M and W, M<sub>1</sub> also quickly adjusts the pitch angle and follows W<sub>1</sub> to reach a higher altitude; while M<sub>2</sub> is responsible for continuing to approach W<sub>2</sub> to expand its own advantages. At this time, due to the hysteresis of the M<sub>1</sub> response, W<sub>1</sub> has a greater advantage over M<sub>1</sub>; while W<sub>2</sub> is in a tracked state, therefore, M<sub>2</sub> has a greater advantage over W<sub>2</sub>. In the first stage, the two sides have not yet formed a cooperative confrontation, and are still in the stage of balanced confrontation. The confrontation process of step 10 is shown in Figure 3.

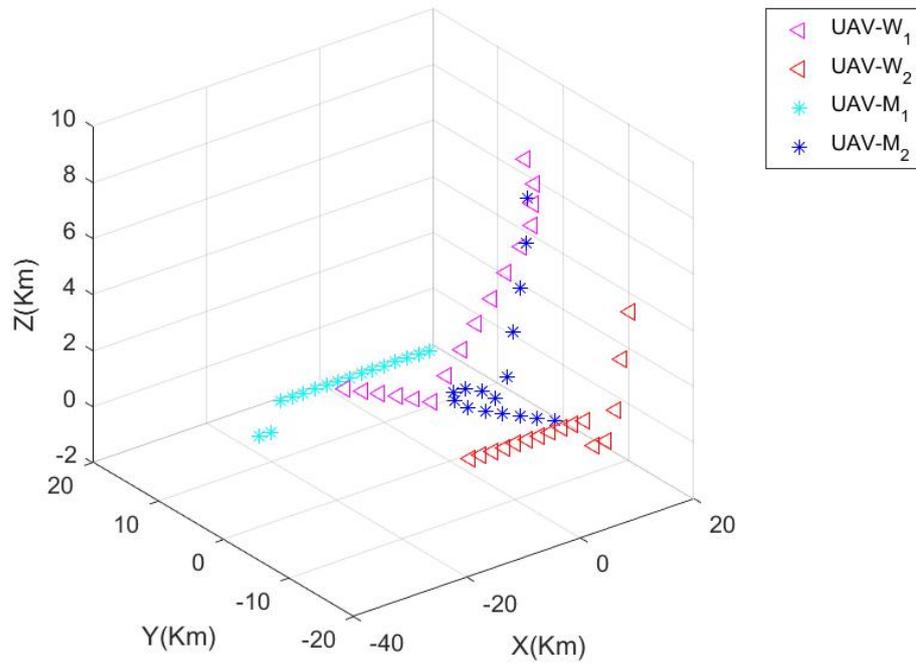


Figure 4. Simulation of UAV cooperative dynamic maneuvering decision algorithm (step 10).

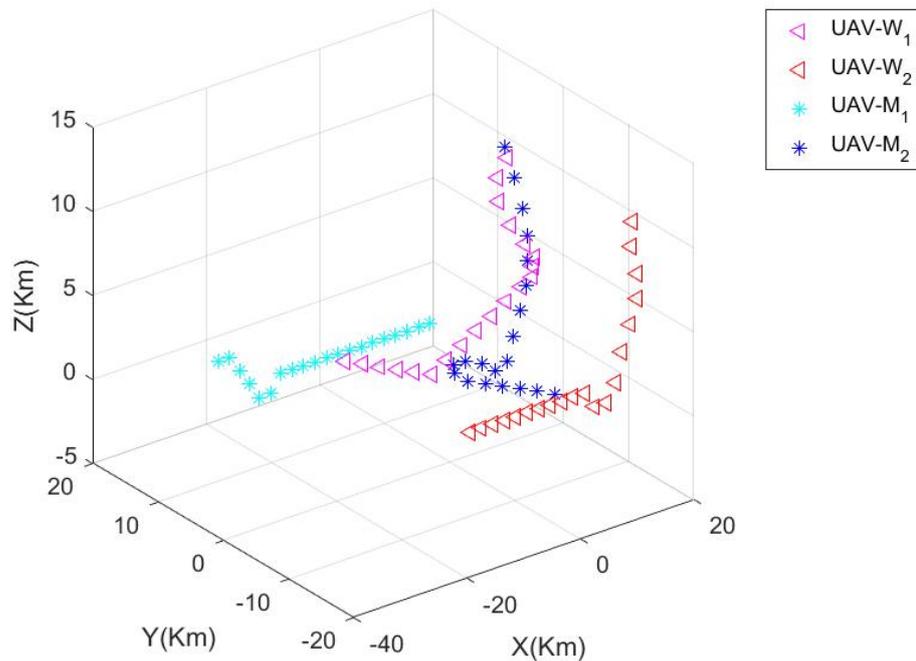


Figure 5. Simulation of UAV cooperative dynamic maneuvering decision algorithm (step 14).

The second phase is the postural equilibrium phase (see Figures 5 and 6). At the 14th step,  $M_2$  is still at a disadvantage relative to  $W_2$ , but because  $M_1$  could not form an absolute advantage over  $W_1$  alone due to the large distance and the limitation of the single-step deflection angle change,  $M_2$  gave up the strategy of fighting  $M_2$  alone and makes a strategic adjustment of approaching  $W_1$  and cooperating with  $M_1$  to surround and fight  $W_1$ . And  $W_2$  could not change the direction in time to assist  $W_1$  due to the long distance; at this time,  $M_1$  and  $M_2$  as a whole maintain the advantage over  $W_1$ , as shown in Figure 5.

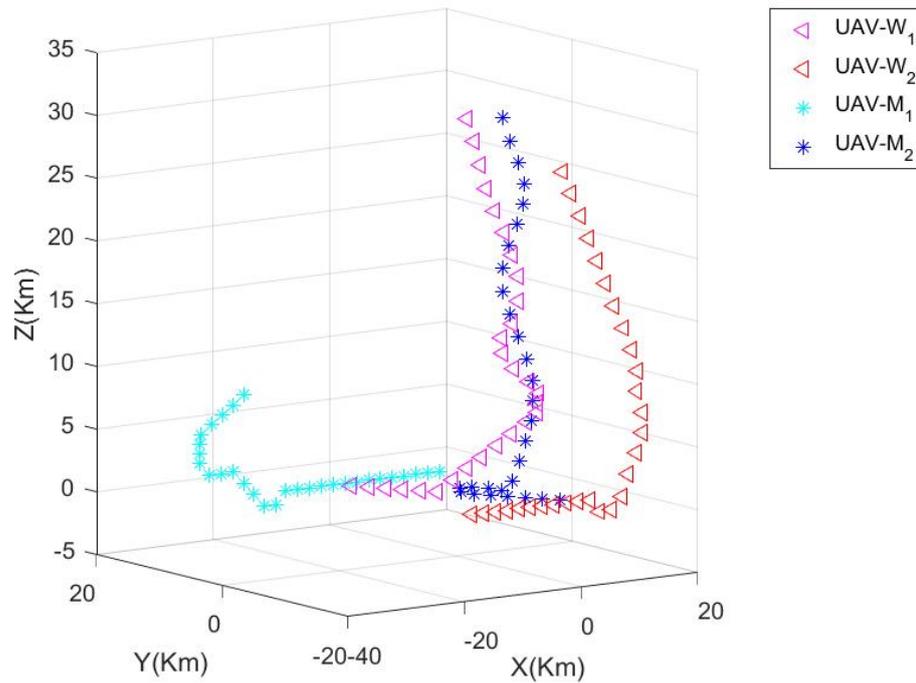


Figure 6. Simulation of UAV cooperative dynamic maneuvering decision algorithm (step 20).

Figure 6 shows the 20th step,  $W_2$  collaborates with  $W_1$  to approach  $M_2$ , at this time  $W_1$  completely removes  $M_1$ , and the  $M_2$  form is intertwined, the advantage of both sides is still not obvious, because the distance  $W_2$  in a shorter number of steps fails to approach  $M_2$  in time, and  $W_2$  together form a pincer attack on  $M_2$ . As shown in Figure 6, at this time, the local gradually forms two against one state, and  $W$ 's overall posture advantage gradually forms.

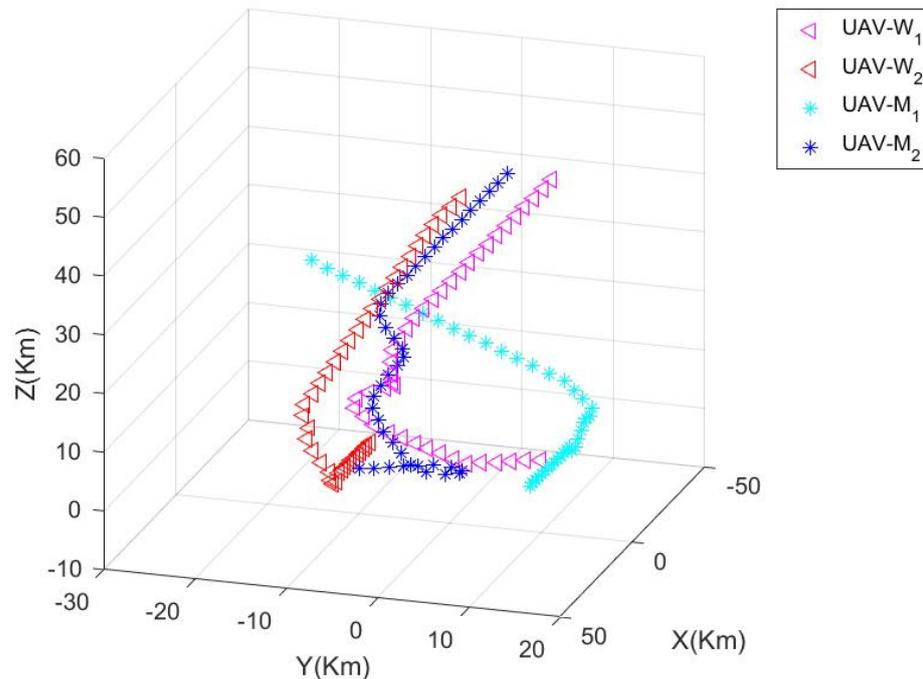
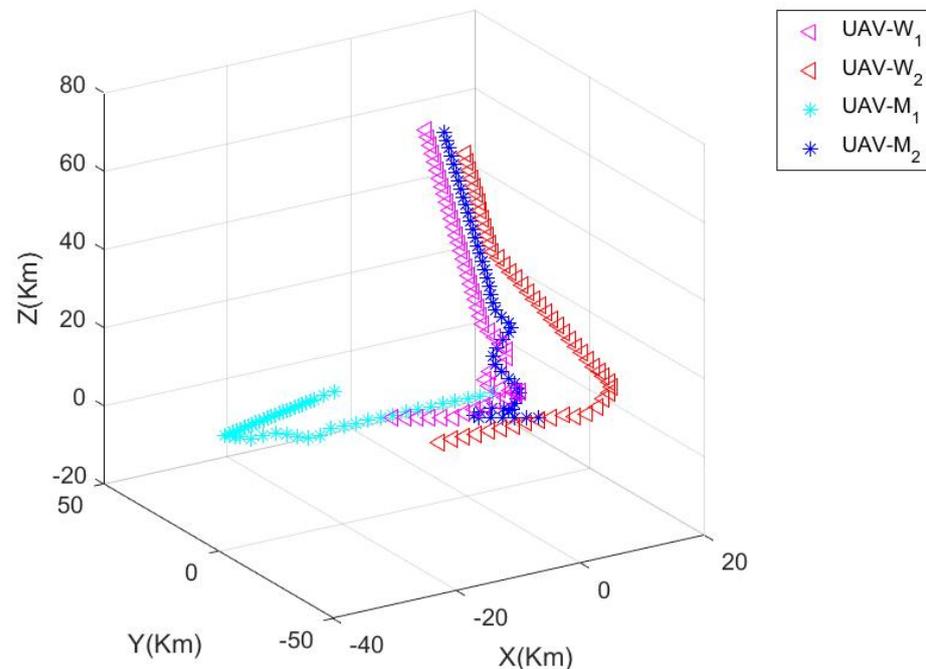


Figure 7. Simulation of UAV cooperative dynamic maneuvering decision algorithm (step 31).

The third phase is the absolute advantage phase (see Figures 7 and 8), in which  $W$  maintains the absolute advantage of the overall posture. In step 31 (see Figure 7),  $W$  maintains the posture of pinning  $M_2$ , with two chasing one, and the threat to  $W$  basically

disappears as  $M_1$  maintains a large distance between  $M_1$  and  $W$ .  $W$  maintains the absolute advantage of the overall posture.



**Figure 8.** Simulation of UAV cooperative dynamic maneuvering decision algorithm (step 36).

Figure 8 shows the course of the confrontation between the two sides at step 36. In terms of the overall confrontation posture,  $W$  maintains the absolute advantage of the overall posture into the third phase. At this point, the confrontation ends.

## 6. Conclusions

In this paper, a more practical UAV cluster cooperative adversarial decision algorithm based on a dynamic non-zero-sum game under uncertain asymmetric information is proposed based on the actual situation. This article mainly completed the following work:

1. Firstly, based on the actual situation of the actual field, the posture advantage of the adversarial parties under ideal conditions is calculated, and thereafter the gain matrix of the adversarial parties under ideal conditions is further calculated.
2. Secondly, considering the uncertainty of both adversaries in acquiring information and the complexity and variability of the real-time field situation, the information acquired by both adversaries is not the exact value, and the gain matrices of both sides of the game are modified. Then, the particle swarm algorithm is improved to solve the dynamic non-zero and Nash equilibrium maneuver decision model efficiently and quickly to obtain the optimal hybrid strategy based on the actual situation.
3. Finally, a 2-to-2 unmanned cluster cooperative countermeasures simulation experiment is given to verify the superiority and realism of the algorithm proposed in this paper to solve the UAV cooperative countermeasures problem.

Based on the idea of non-zero-sum game, this model regards the solution of Nash equilibrium as the maneuvering action of the UAV cluster. Compared with the traditional model, it has the following advantages:

1. Optimal Strategy: A Nash Equilibrium is an optimal strategy in which each player cannot improve his individual payoff by changing his own strategy given the strategies of the other players. This means that the Nash equilibrium strategy adopted by drone clusters is not easy to be defeated, and it is a relatively stable strategy. It is more suitable for objects such as drones that require high stability.

2. **Predictability:** Nash equilibrium can be used to predict and understand various game scenarios, especially in complex environments where multiple parties interact. By analyzing and calculating the Nash equilibrium, the drone swarm can speculate on the possible behavior and outcomes of the participants, so as to better predict and plan the strategy of the swarm.

3. **Stability:** Nash equilibrium is theoretically stable. Even in the face of some disturbances or external pressures, participants will tend to stick to their equilibrium strategies. This stability can help maintain a balanced state of the game and reduce possible conflicts and confrontations. In a more complex actual environment, external disturbances or changes are very frequent. If the UAV cluster changes its strategy too frequently, it will cause damage to itself.

Although this paper has performed some research on the UAV swarm confrontation decision-making problem based on incomplete information, there are still some unresolved problems:

1. In this paper, incomplete information refers to unknowable information such as enemy strategy, enemy revenue, and the partially observable environment when making simultaneous decisions, and does not discuss in depth the unavailable information and untrustworthy information of incomplete information.

2. This paper simplifies the flight process in the UAV swarm air-to-air confrontation environment, and does not take into account the flight characteristics of the UAV itself. In future work, we will focus on a swarm confrontation that is closer to the real environment, including specific attack processes such as missile attacks. In addition, the complex environmental characteristics of the real battlefield have not been fully considered, such as electromagnetic space, weather effects, etc.

3. In the process of UAV swarm confrontation, problems such as large-scale (number greater than 1000) will lead to joint state-action dimensional explosion and dynamic changes in the number of UAVs (scalability) are not considered, which does not meet the requirements of swarm operations actual needs.

To sum up, although this paper has made further research on the basis of existing research, there is still a big gap with the actual environment. The above points will be the key issues of our next research.

**Author Contributions:** Conceptualization, P.W. and H.W.; Methodology, P.W.; Software, P.W.; Validation, P.W.; Formal analysis, P.W.; Resources, P.W. and P.Z.; Writing—Original draft, P.W.; Writing—review & editing, P.W., H.W., G.L. and P.Z.; Visualization, H.W.; Supervision, G.L.; Funding acquisition, P.W. and P.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** Innovation and Entrepreneurship Fund of the Student Work Department of the Party Committee of Northwestern Polytechnical University (NO.2023-CXCX-021), Higher Education Research Fund, Northwestern Polytechnical University (NO.GJJJM202405) and Northwestern Polytechnical University Education and Teaching Reform Research Fund (NO.2023JGZ25).

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Fan, B.; Li, Y.; Zhang, R.; Fu, Q. Review on the technological development and application of UAV systems. *Chin. J. Electron.* **2020**, *29*, 199–207. [[CrossRef](#)]
2. Wang, Y.; Zhang, W.; Li, Y. An efficient clonal selection algorithm to solve dynamic weapon-target assignment game model in UAV cooperative aerial combat. In Proceedings of the 2016 35th Chinese Control Conference (CCC), Chengdu, China, 27–29 July 2016; pp. 9578–9581.
3. Wang, Y.; Zhao, F.; Wang, H.; Hu, X.; Zhang, W.; Zhang, T.; Feng, Y. Research Status and Development of Graded-index Multimode Fiber Mode-locking Technique. *Acta Photonica Sin.* **2020**, *49*, 34–44. [[CrossRef](#)]
4. Shao, J.; Xu, Y.; Luo, D. Cooperative combat decision-making research for multi UAVs. *Inf. Control* **2018**, *47*, 347–354.
5. Niu, Y.; Xiao, X.; Ke, G. Operation concept and key techniques of unmanned aerial vehicle swarms. *Natl. Def. Sci. Technol.* **2013**, *34*, 37–43.

6. Chen, X.; Liu, M.; Hu, Y.-X. Study on UAV Offensive/Defensive Game Strategy Based on Uncertain Information. *Acta Armamentarii* **2012**, *33*, 1510.
7. Zhen, Z.; Zhu, P.; Xue, Y.; Ji, Y. Distributed intelligent self-organized mission planning of multi-UAV for dynamic targets cooperative search-attack. *Chin. J. Aeronaut.* **2019**, *32*, 2706–2716. [[CrossRef](#)]
8. Xu, J.; Guo, Q.; Xiao, L.; Li, Z.; Zhang, G. Autonomous decision-making method for combat mission of UAV based on deep reinforcement learning. In Proceedings of the 2019 IEEE 4th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), Chengdu, China, 20–22 December 2019; Volume 1, pp. 538–544.
9. Han, Y.; Yan, J.; Chen, R.; Li, J.; Sun, S.; Lin, Y. Improved Game Theory Based Targets Assigning for Ship-based UAV Formation Coordinated Air-to-Sea Attack. *Fire Control Command Control* **2016**, *41*, 65–70.
10. Yao, Z.; Li, M.; Chen, Z.; Zhou, R. Mission decision-making method of multi-aircraft cooperatively attacking multi-target based on game theoretic framework. *Chin. J. Aeronaut.* **2016**, *29*, 1685–1694. [[CrossRef](#)]
11. Zhang, L.; Zhang, A. Study on air formation to ground attack-defends decision-making in uncertainty. *Syst. Eng. Electron.* **2009**, *31*, 411–415.
12. Cruz, J.; Simaan, M.A.; Gacic, A.; Liu, Y. Moving horizon Nash strategies for a military air operation. *IEEE Trans. Aerosp. Electron. Syst.* **2002**, *38*, 989–999. [[CrossRef](#)]
13. Cruz, J.B.; Simaan, M.A.; Gacic, A.; Jiang, H.; Letellier, B.; Li, M.; Liu, Y. Game-theoretic modeling and control of a military air operation. *IEEE Trans. Aerosp. Electron. Syst.* **2001**, *37*, 1393–1405. [[CrossRef](#)]
14. Johnson, L.; Ponda, S.; Choi, H.L.; How, J. Asynchronous decentralized task allocation for dynamic environments. In *Infotech@Aerospace 2011*; American Institute of Aeronautics and Astronautics (AIAA): St. Louis, MO, USA, 2011; p. 1441.
15. Long, T.; Shen, L.; Zhu, H.; Niu, Y. Distributed task allocation & coordination technique of multiple UCAVs for cooperative tasks. *Acta Autom. Sin.* **2007**, *33*, 731.
16. Hu, J.; Wellman, M.P. Multiagent reinforcement learning: Theoretical framework and an algorithm. In Proceedings of the ICML, Madison, WI, USA, 24–27 July 1998; Volume 98, pp. 242–250.
17. Xuan, S.; Zhou, H.; Ke, L. Review of UAV Swarm Confrontation Game. *Command Inf. Syst. Technol.* **2021**, *12*, 27–31.
18. Chin, H.H. Knowledge-based system of supermaneuver selection for pilot aiding. *J. Aircr.* **1989**, *26*, 1111–1117. [[CrossRef](#)]
19. Bechtel, R.J. *Air Combat Maneuvering Expert System Trainer*; Technical Report; Merit Technology Inc.: Plano, TX, USA, 1992.
20. Bhattacharjee, P.; Rakshit, P.; Goswami, I.; Konar, A.; Nagar, A.K. Multi-robot path-planning using artificial bee colony optimization algorithm. In Proceedings of the 2011 Third World Congress on Nature and Biologically Inspired Computing, Salamanca, Spain, 19–21 October 2011; pp. 219–224.
21. Butenko, S.; Murphey, R.; Pardalos, P.M. *Cooperative Control: Models, Applications and Algorithms*; Springer Science & Business Media: Amsterdam, The Netherlands, 2013; Volume 1.
22. Lukasik, S.; Zak, S. Firefly algorithm for continuous constrained optimization tasks. In Proceedings of the ICCCI, Wroclaw, Poland, 5–7 October 2009; pp. 97–106.
23. Wu, H.S.; Zhang, F.; Wu, L. New swarm intelligence algorithm-wolf pack algorithm. *Syst. Eng. Electron.* **2013**, *35*, 2430–2438.
24. Chen, X.; Li, G.; Zhao, L. Research on UCAV game strategy of cooperative air combat task. *Fire Control Command Control* **2018**, *43*, 17–23.
25. Li, B.; Wang, H.; Mu, S. Situation Reasoning Based on Ontology Modeling. *Aeronaut. Sci. Technol.* **2021**, *32*, 80–90.
26. Duffy, J. *Game Theory and Nash Equilibrium*; Lakehead University: Thunder Bay, ON, Canada, 2015; pp. 19–20.
27. Osborne, M.J.; Rubinstein, A. *A Course in Game Theory*; MIT Press: Cambridge, MA, USA, 1994; p. 14.
28. Liu, K.; Cao, X.; Li, Y.; Fang, D. Research on Cooperative Countermeasures of UAV cluster based on dynamic non-zero-sum game. *Aviat. Sci. Technol.* **2022**, *33*, 20–36. [[CrossRef](#)]
29. Wang, Y.; Zhang, W.; Fu, L.; Huang, D.G.; Li, Y. Nash equilibrium strategies approach for aerial combat based on elite re-election particle swarm optimization. *Control Theory Appl.* **2015**, *32*, 857–865.
30. Pavlidis, N.G.; Parsopoulos, K.E.; Vrahatis, M.N. Computing nash equilibria through computational intelligence methods. *J. Comput. Appl. Math.* **2005**, *1*, 113–136. [[CrossRef](#)]
31. Muhammad, R.; Saeed, M.; Ali, B.; Ahmad, N.; Ali, L.; Abdal, S. Application of Interval Valued Fuzzy Soft Max-Min Decision Making Method. *Int. J. Math. Res.* **2020**, *9*, 11–19.

**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.