Journal of Systems Engineering and Electronics
Vol. 34, No. 6, December 2023, pp.1447 – 1464

Multicriteria game approach to air-to-air combat tactical decisions for multiple UAVs

JIANG Ruhao[1,2], LUO He[1,2,3], MA Yingying[1,2], and WANG Guoqiang[1,2,3,*]

1. School of Management, Hefei University of Technology, Hefei 230009, China;
2. Key Laboratory of Process Optimization & Intelligent Decision-making, Ministry of Education, Hefei 230009, China;
3. Intelligent Interconnected Systems Laboratory of Anhui Province, Hefei 230009, China

Abstract: Air-to-air combat tactical decisions for multiple unmanned aerial vehicles (ACTDMU) are a key decision-making step in beyond visual range combat. Complex influencing factors, strong antagonism and real-time requirements need to be considered in the ACTDMU problem. In this paper, we propose a multicriteria game approach to ACTDMU. This approach consists of a multicriteria game model and a Pareto Nash equilibrium algorithm. In this model, we form the strategy profiles for the integration of air-to-air combat tactics and weapon target assignment strategies by considering the correlation between them, and we design the vector payoff functions based on predominance factors. We propose a algorithm of Pareto Nash equilibrium based on preference relations using threshold constraints (PNE-PRTC), and we prove that the solutions obtained by this algorithm are refinements of Pareto Nash equilibrium solutions. The numerical experiments indicate that PNE-PRTC algorithm is considerably faster than the baseline algorithms and the performance is better. Especially on large-scale instances, the Pareto Nash equilibrium solutions can be calculated by PNE-PRTC algorithm at the second level. The simulation experiments show that the multicriteria game approach is more effective than one-side decision approaches such as multiple-attribute decision-making and randomly chosen decisions.

Keywords: tactical decision, multicriteria game, preference relation, Pareto Nash equilibrium.

DOI: 10.23919/JSEE.2023.000115

1. Introduction

With the improvement of unmanned aerial vehicle (UAV) detect ability and air-to-air missile performance, beyond visual range (BVR) combat has become an increasingly important research field in recent years [1,2]. Many institutes, such as Defense Advanced Research Projects Agency (DARPA), Aeronautics Institute of Technology (ITA), are committed to studying the field [3−5].

This paper focuses on air-to-air combat tactical decisions for multiple UAVs (ACTDMU), which is a key step in BVR combat. The ACTDMU can be described as a decision-making problem, the purpose of which is to choose a reasonable tactic to ensure that UAVs complete the attack tasks safely and efficiently. Executing a reasonable tactic means occupying a dominant position. The analysis of the overall predominance value and the threat assessment of foe UAVs in the combat situation are the premises of this problem. Moreover, the ACTDMU problem is a macrolevel and complex decision-making process. This problem has characteristics such as complex influencing factors, an antagonistic combat environment, real-time requirements, and rapidly shifting predominance values [6]. In particular, air-to-air combat tactical decisions and weapon target assignments (WTAs) are strongly correlated. Both decisions are key steps in air-to-air combat situations, and both require UAVs to cooperate with each other to maximize the overall advantage over the foe's UAVs. In addition, the time of decision-making for the two decisions overlaps, and the attack tasks for the two decisions are coupled with each other [7]. These characteristics of the ACTDMU problem pose a challenge for its solution.

In recent years, many approaches to solving the ACTDMU problem have been proposed by researchers, and they are divided into two main categories: (i) Intelligent decision-making approaches based on theories such as multi-attribute decision-making (MADM) [8], rough sets [9], and fuzzy sets [10]. In these approaches, the complexity of the influencing factors is considered, and tac-

Manuscript received August 30, 2022.
*Corresponding author.
This work was supported in part by the National Natural Science Foundation of China (71971075; 71871079; 71671059) and in part by the Anhui Provincial Natural Science Foundation (1808085MG213).

tics can be calculated rapidly by fuzzy reasoning. However, these approaches tend to make decisions from the perspective of one's own side, and the impact of the decisions of others on one's own side is ignored. It is difficult to describe the interaction between the selected strategies by using these approaches. (ii) The other decision-making approach is based on game theory [11,12]. In this approach, the strong antagonism of air combat can be considered. The strategy profiles formed by all decision makers are analyzed first. Then, the payoffs for all sides with a single objective are obtained according to the characteristics of the problem. Finally, the solutions can be calculated by constructing a Nash equilibrium [13]. Nevertheless, for this decision-making approach, it is difficult to fully represent the interaction relations between multiple influencing factors, which often leads to losing some useful information. The multicriteria game approach is an effective way to solve the ACTDMU problem. Not only the interaction relation among the complex influencing factors but also the interaction relation among all decision makers in choosing strategies is considered in this approach.

The multicriteria game was first proposed by Blackwell in 1956 [14]. It is also termed a multiobjective game or vector payoff game. Multiple objectives or lists of payoffs are considered in this game, and this game has the ability to reflect multiple influencing factors. Unlike the classical Nash equilibrium, the Pareto Nash equilibrium (PNE) is usually constructed in the multicriteria game [15]. Furthermore, various kinds of equilibria can be constructed by improving or refining the PNE according to the constraints of different multicriteria game problems. A large number of studies have focused on the definition of the equilibrium and proofs of the existence of its solutions. The Pareto-optimal security strategy equilibrium is constructed by using standard fuzzy orders to solve the multicriteria game problem with fuzzy payoffs [16]. According to the attitude toward risk of the players in the games, the G-goal security strategy equilibrium is constructed to solve the multicriteria game problem with the desired goal [17]. By considering the rational choice deviation of the decision makers, an ideal Nash equilibrium is constructed to refine the solution set of the PNE [18,19]. The properly efficient Nash equilibrium is proposed to extend and perfect the efficient Nash equilibrium, and its sufficient existence conditions are established by using maximal elements [20,21].

PNE solutions play a central role in solving multicriteria games. Nevertheless, it is a recognized difficult problem to select compromise PNE solutions, which attempt to maximize all objectives, because this involves the construction of dominance relations and tedious calculations [22]. For multicriteria games, there are few studies on easy-to-use PNE algorithms compared with the literature on the definition of equilibrium. Some studies attempt to calculate PNE solutions by linearly weighting the multicriteria game as a scalarized single-criterion game [23,24]. However, the objectives of the decision makers are conflicting and coupled in many cases. It is difficult to calculate high-quality solution sets with a wide dispersion in this way.

In addition, combined with the concept of the PNE, other approaches for solving multicriteria games can be sorted into three categories. The first algorithm is based on optimization theory and transforms the multicriteria game problem into a linear programming problem, a quadratic programming problem or another problem [25−27]. This kind of algorithm can find PNE solutions quickly, but it is usually suitable for biobjective games with small sizes. Second, artificial intelligence-based approaches are used to solve the multicriteria game. For example, the fuzzy set-based algorithm can calculate PNE solutions by constructing membership functions for fuzzy goals [28,29], and the heuristic algorithm can solve it by using a population evolution strategy [30,31]. However, these algorithms are usually used in problems with low real-time requirements, and their application range is always limited for the problem constraints. The third type of algorithm is based on theories such as MADM and fuzzy decision-making, and it finds PNE solutions by establishing dominance relations [32,33]. This approach is capable of solving non-zero-sum multicriteria game problems, and the construction of preference relations is a critical step that greatly affects the efficiency of the algorithm and the quality of the solutions.

In the ACTDMU problem, with the increase in the number of tactics and objectives, the degree of difficulty of solving it increases sharply [22]. High-quality PNE solutions obtained with existing approaches always take a few seconds or even longer. It is difficult to meet the real-time requirements of this problem. Otherwise, combined with knowledge of air combat, it is necessary to build a model considering the characteristics of the ACTDMU problem. Thus, there is a gap in the usability of the approaches for solving the ACTDMU problem. In this paper, we propose a multicriteria game approach to the ACTDMU problem. This approach consists of a multicriteria game model and a PNE algorithm. The main highlights of this work are as follows:

(i) Strategy profiles for the integration of air-to-air

combat tactics and WTA strategies are constructed by considering the correlation between them.

(ii) Vector payoff functions of multicriteria games are designed based on predominance factors considering the complexity of the influencing factors and strong antagonism in air-to-air combat.

(iii) A algorithm of PNE based on preference relations using threshold constraints (PNE-PRTC) is proposed.

(iv) The equilibrium solution obtained by the PNE-PRTC algorithm is proven to be a refinement of the PNE solution.

The remainder of this paper is arranged as follows: Section 2 briefly describes the ACTDMU problem. In Section 3, a multicriteria game model for ACTDMU is established. The PNE-PRTC algorithm for solving the multicriteria game model is proposed in Section 4. Numerical experiments and simulation experiments are given in Section 5. Section 6 concludes the paper.

2. Problem formulation

ACTDMU is a macroscopic decision-making problem in beyond-visual-range combat. First, it is necessary to fully analyze the predominance value in a combat situation and evaluate the threat levels of foe UAVs. Then, tactics affected by multiple-decision information are compared and selected according to superiority. Finally, a reasonable tactic is executed to maximize the overall advantage of our UAV swarm against the foe's targets. The characteristics of the ACTDMU problem require that decisions be made as quickly as possible. Fig. 1 shows that a selected tactic is executed in the ACTDMU problem.

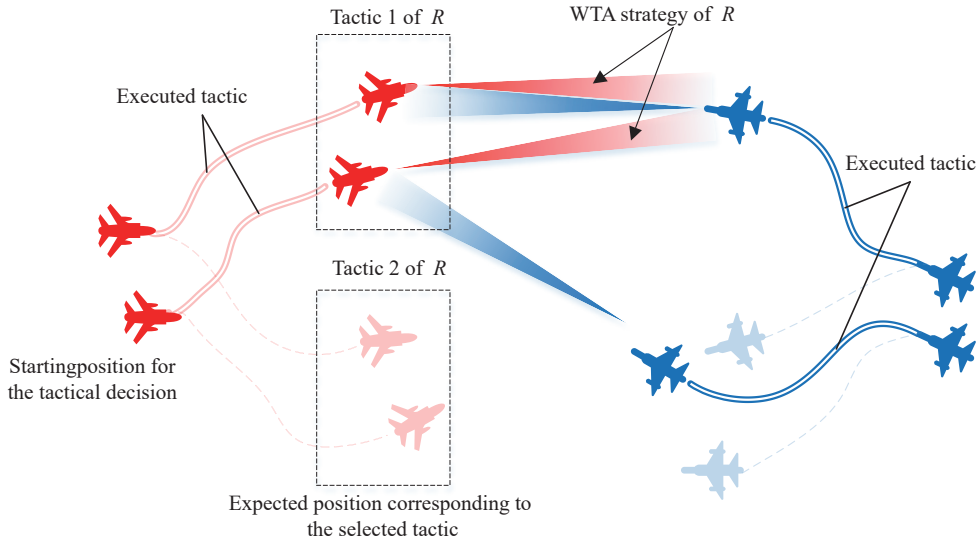


Fig. 1 Executing the tactics in the ACTDMU problem

A set of tactics for our UAV formation R and a set of tactics for the foe's UAV formation B are recorded as $R_{\text{tac}} = \{r_m\}_{m\in M}$ and $B_{\text{tac}} = \{b_n\}_{n\in N}$, respectively. It is assumed that R and B make decisions simultaneously at the beginning of tactical decision-making and fly from the starting position to the expected position corresponding to the requirement of the tactics. The position and state sets corresponding to the tactics selected by R and B are denoted as $R_{\text{pos}} = \{p^p_{r_m}\}_{m\in M, p\in P}$ and $B_{\text{pos}} = \left\{p^q_{b_n}\right\}_{n\in N, q\in Q}$, respectively. The position and state of a single UAV in R and B are recorded as $p^p_{r_m} = (x, y, z, v, \theta, \phi)_{p^p_{r_m}}$ and $p^q_{b_n} = (x, y, z, v, \theta, \phi)_{p^q_{b_n}}$. x, y, z are the position coordinates, v is the speed, and θ and ϕ express the pitch angle and yaw angle, respectively. The detailed parameter definitions of this problem description are shown in Table 1.

Table 1 Parameter definitions of the ACTDMU problem

Variable	Description
R	Our UAV formation
B	Foe UAV formation
r_m	The mth tactic of R, $m \in M$
b_n	The nth tactic of B, $n \in N$
R_{tac}	A set of tactics for R
B_{tac}	A set of tactics for B
p	R's UAV number p, $p \in P$
q	B's UAV number q, $q \in Q$
$p^p_{r_m}$	The position and state of p when r_m is selected
$p^q_{b_n}$	The position and state of q when b_n is selected
R_{pos}	The position and state sets of R's UAVs for the sets of tactics
B_{pos}	The position and state sets of B's UAVs for the sets of tactics

3. Multicriteria game model

Considering the correlation between air-to-air combat tactical decisions and WTAs, this paper integrates WTA strategies into air-to-air combat tactics, and a multicriteria game model is developed for the ACTDMU problem. We describe a finite multicriteria two-person game model by the tuple $G=\langle P,O,F\rangle$. $P=\{R,B\}$ is the set of players, where R and B are two players. $O=\bar{R}_{\text{tac}}\times\bar{B}_{\text{tac}}$ denotes the finite set of strategy profiles for the integration of air-to-air combat tactics and WTA strategies. $\bar{R}_{\text{tac}}$ and $\bar{B}_{\text{tac}}$ are denoted as the sets of air-to-air combat tactics incorporating the WTA strategies of R and B, respectively. $F=\{\boldsymbol{F}_R,\boldsymbol{F}_B\}$ indicates the vector payoff function, where $\boldsymbol{F}_R$ is for R and $\boldsymbol{F}_B$ is for B.

3.1 Strategy profiles

First, the sets of air-to-air combat tactics incorporating the WTA strategies of R and B are constructed, and then the set of strategy profiles for integrating air-to-air combat tactics and WTA strategies are formed.

(i) Sets of air-to-air combat tactics incorporating WTA strategies

In this paper, air-to-air combat tactics and WTA strategies are combined to construct air-to-air combat tactics incorporating WTA strategies, which are denoted by $r_{r_mw_c}$ for R. $r_{r_mw_c}$ indicates that tactic r_m is selected and WTA strategy w_c is implemented, where w_c is the cth WTA strategy of R. Similarly, $b_{b_nv_d}$ denotes that tactic b_n is selected and WTA strategy v_d is implemented, where v_d is the dth WTA strategy of B. A pair of strategy profiles for air-to-air combat tactics incorporating WTA strategies is recorded as $o_{\tau_{mc}\tau_{nd}}=(r_{r_mw_c},b_{b_nv_d})$.

(ii) Sets of strategy profiles for integrating air-to-air combat tactics and WTA strategies

Sets of air-to-air combat tactics incorporating WTA strategies $\bar{R}_{\text{tac}}=\{r_{r_mw_c}\}_{m\in M,c\in C}$ and $\bar{B}_{\text{tac}}=\{b_{b_nv_d}\}_{n\in N,d\in D}$ are constructed. The number of air-to-air combat tactics incorporating WTA strategies is expressed as $\left|\bar{R}_{\text{tac}}\right|=|R_{\text{tac}}\times W_R|$ and $\left|\bar{B}_{\text{tac}}\right|=|B_{\text{tac}}\times V_B|$, where W_R and V_B are the one-to-one attack relations of all UAVs for R to B and B to R [34], where $W_R=\{w_c\}_{c\in C}$ and $V_B=\{v_d\}_{d\in D}$. C and D are the numbers of WTA strategies for R and B, respectively.

On this basis, sets of strategy profiles for the integration of air-to-air combat tactics and WTA strategies $\bar{R}_{\text{tac}}\times\bar{B}_{\text{tac}}=O$ are established. Notably, $o_{\tau_{mc}\tau_{nd}}\in O$.

3.2 Vector payoff functions

In the ACTDMU problem, the vector payoff functions consist of two sets of objective functions according to the validity of the fighting effect and the threat assessment of foe UAVs. The two sets of objective functions are constructed by using four predominance factors: speed predominance, angle predominance, altitude predominance, and distance predominance. Each objective function in a set is constructed by using one predominance factor. For example, for speed predominance, R's objective function of the validity of the fighting effect is denoted as f_{R1}. For angle predominance, altitude predominance and distance predominance, R's objective functions of the validity of the fighting effect are denoted as f_{R2}, f_{R3} and f_{R4}, respectively. Similarly, R's objective functions of the threat assessment of the foe UAVs in terms of speed predominance, angle predominance, altitude predominance and distance predominance are denoted as f_{R5}, f_{R6}, f_{R7} and f_{R8}, respectively. R's vector payoff function $\boldsymbol{F}_R=(f_{R1},f_{R2},\cdots,f_{R8})^t$ is composed of eight objective functions. Similar to that of R, B's vector payoff function $\boldsymbol{F}_B=(f_{B1},f_{B2},\cdots,f_{B8})^t$ is constructed.

The jth objective function of i is expressed as

$$f_{ij}=\begin{cases}\dfrac{1}{1+\mathrm{e}^{-(V_i^j/V_{-i}^j)}}, & j=1,2,3,4\\ \dfrac{1}{1+\mathrm{e}^{-((1-V_{-i}^{8-j})/(1-V_i^{8-j}))}}, & j=5,6,7,8\end{cases}\tag{1}$$

where $-i$ is the opponent of i. j denotes an objective, $j=1,2,\cdots,8$. V_i^j represents the validity of the fighting effect of i for the jth objective, and $1-V_{-i}^{8-j}$ represents the safety level, which is calculated by a threat assessment of the foe UAVs. Moreover, V_i^j/V_{-i}^j represents the relative validity of the fighting effect of i for the jth objective, and $\left(1-V_{-i}^{8-j}\right)/\left(1-V_i^{8-j}\right)$ represents the safety level of i for the jth objective. The sigmoid function is used to map the ratio results from 0 to 1.

$V_i^j(r_{r_mw_c},b_{b_nv_d})$ denotes the validity of the fighting effect of i for the jth objective when a pair of strategy profiles $(r_{r_mw_c},b_{b_nv_d})$ is selected. $V_i^j(r_{r_mw_c},b_{b_nv_d})$ is calculated by the multilevel information fusion method [35], and the calculation process is shown as follows:

$$V_{ih}^j(r_{r_mw_c},b_{b_nv_d})=v_{k_1h}^j\oplus v_{k_2h}^j\oplus\cdots\oplus v_{k_lh}^j,\tag{2}$$

$$V_i^j(r_{r_mw_c},b_{b_nv_d})=V_{ih}^j(r_{r_mw_c},b_{b_nv_d})\oplus\cdots\oplus V_{iH}^j(r_{r_mw_c},b_{b_nv_d}),\tag{3}$$

where $\oplus$ is the fusion operator of evidence theory [36]. $V_{ih}^j(r_{r_mw_c},b_{b_nv_d})$ represents the first-level fusion of the predominance value for jth objective of i in (2). $V_i^j(r_{r_mw_c},b_{b_nv_d})$ represents the second-level fusion of the first-level fusion value for the jth objective of i in (3), that is, the validity of the fighting effect of i.

Taking the example of calculating $V_R^1(r_{r_mw_c},b_{b_nv_d})$ for the first objective, an attack of multiple R UAVs $k_1,k_2,\cdots,k_l$ on one of B's UAVs h is regarded as an

attack group. In terms of speed predominance for the first objective, $v^1_{k_1h}, v^1_{k_2h}, \cdots, v^1_{k_lh}$ are the validities of the fighting effect of $k_1, k_2, \cdots, k_l$ on h in this attack group [37]. $V^j_{ih}(r_{r_mw_c}, b_{b_nv_d})$ is obtained by a first-level fusion of $v^j_{k_1h}, v^j_{k_2h}, \cdots, v^j_{k_lh}$. Then, the second-level is calculated for the $V^j_i(r_{r_mw_c}, b_{b_nv_d})$ value in each group to obtain R's validity of the fighting effect for the first objective. Analogously, the validities of the fighting effect for the other objectives can also be calculated.

For the set of strategy profiles for integrating air-to-air combat tactics and WTA strategies, the vector payoff functions of R and B for every objective are constructed. Then, the vector payoff matrices shown in (4) and (5) can be calculated.

$$\boldsymbol{F}_R = \begin{bmatrix} \begin{pmatrix} f_{R1}(r_{r_1w_1}, b_{b_1v_1}) \\ f_{R2}(r_{r_1w_1}, b_{b_1v_1}) \\ \vdots \\ f_{R8}(r_{r_1w_1}, b_{b_1v_1}) \end{pmatrix} & \begin{pmatrix} f_{R1}(r_{r_1w_1}, b_{b_2v_2}) \\ f_{R2}(r_{r_1w_1}, b_{b_2v_2}) \\ \vdots \\ f_{R8}(r_{r_1w_1}, b_{b_2v_2}) \end{pmatrix} & \cdots & \begin{pmatrix} f_{R1}(r_{r_1w_1}, b_{b_Nv_D}) \\ f_{R2}(r_{r_1w_1}, b_{b_Nv_D}) \\ \vdots \\ f_{R8}(r_{r_1w_1}, b_{b_Nv_D}) \end{pmatrix} \\ \begin{pmatrix} f_{R1}(r_{r_2w_2}, b_{b_1v_1}) \\ f_{R2}(r_{r_2w_2}, b_{b_1v_1}) \\ \vdots \\ f_{R8}(r_{r_2w_2}, b_{b_1v_1}) \end{pmatrix} & \begin{pmatrix} f_{R1}(r_{r_2w_2}, b_{b_2v_2}) \\ f_{R2}(r_{r_2w_2}, b_{b_2v_2}) \\ \vdots \\ f_{R8}(r_{r_2w_2}, b_{b_2v_2}) \end{pmatrix} & \cdots & \begin{pmatrix} f_{R1}(r_{r_2w_2}, b_{b_Nv_D}) \\ f_{R2}(r_{r_2w_2}, b_{b_Nv_D}) \\ \vdots \\ f_{R8}(r_{r_2w_2}, b_{b_Nv_D}) \end{pmatrix} \\ \vdots & \vdots & \ddots & \vdots \\ \begin{pmatrix} f_{R1}(r_{r_Mw_C}, b_{b_1v_1}) \\ f_{R2}(r_{r_Mw_C}, b_{b_1v_1}) \\ \vdots \\ f_{R8}(r_{r_Mw_C}, b_{b_1v_1}) \end{pmatrix} & \begin{pmatrix} f_{R1}(r_{r_Mw_C}, b_{b_2v_2}) \\ f_{R2}(r_{r_Mw_C}, b_{b_2v_2}) \\ \vdots \\ f_{R8}(r_{r_Mw_C}, b_{b_2v_2}) \end{pmatrix} & \cdots & \begin{pmatrix} f_{R1}(r_{r_Mw_C}, b_{b_Nv_D}) \\ f_{R2}(r_{r_Mw_C}, b_{b_Nv_D}) \\ \vdots \\ f_{R8}(r_{r_Mw_C}, b_{b_Nv_D}) \end{pmatrix} \end{bmatrix} \tag{4}$$

$$\boldsymbol{F}_B = \begin{bmatrix} \begin{pmatrix} f_{B1}(r_{r_1w_1}, b_{b_1v_1}) \\ f_{B2}(r_{r_1w_1}, b_{b_1v_1}) \\ \vdots \\ f_{B8}(r_{r_1w_1}, b_{b_1v_1}) \end{pmatrix} & \begin{pmatrix} f_{B1}(r_{r_1w_1}, b_{b_2v_2}) \\ f_{B2}(r_{r_1w_1}, b_{b_2v_2}) \\ \vdots \\ f_{B8}(r_{r_1w_1}, b_{b_2v_2}) \end{pmatrix} & \cdots & \begin{pmatrix} f_{B1}(r_{r_1w_1}, b_{b_Nv_D}) \\ f_{B2}(r_{r_1w_1}, b_{b_Nv_D}) \\ \vdots \\ f_{B8}(r_{r_1w_1}, b_{b_Nv_D}) \end{pmatrix} \\ \begin{pmatrix} f_{B1}(r_{r_2w_2}, b_{b_1v_1}) \\ f_{B2}(r_{r_2w_2}, b_{b_1v_1}) \\ \vdots \\ f_{B8}(r_{r_2w_2}, b_{b_1v_1}) \end{pmatrix} & \begin{pmatrix} f_{B1}(r_{r_2w_2}, b_{b_2v_2}) \\ f_{B2}(r_{r_2w_2}, b_{b_2v_2}) \\ \vdots \\ f_{B8}(r_{r_2w_2}, b_{b_2v_2}) \end{pmatrix} & \cdots & \begin{pmatrix} f_{B1}(r_{r_2w_2}, b_{b_Nv_D}) \\ f_{B2}(r_{r_2w_2}, b_{b_Nv_D}) \\ \vdots \\ f_{B8}(r_{r_2w_2}, b_{b_Nv_D}) \end{pmatrix} \\ \vdots & \vdots & \ddots & \vdots \\ \begin{pmatrix} f_{B1}(r_{r_Mw_C}, b_{b_1v_1}) \\ f_{B2}(r_{r_Mw_C}, b_{b_1v_1}) \\ \vdots \\ f_{B8}(r_{r_Mw_C}, b_{b_1v_1}) \end{pmatrix} & \begin{pmatrix} f_{B1}(r_{r_Mw_C}, b_{b_2v_2}) \\ f_{B2}(r_{r_Mw_C}, b_{b_2v_2}) \\ \vdots \\ f_{B8}(r_{r_Mw_C}, b_{b_2v_2}) \end{pmatrix} & \cdots & \begin{pmatrix} f_{B1}(r_{r_Mw_C}, b_{b_Nv_D}) \\ f_{B2}(r_{r_Mw_C}, b_{b_Nv_D}) \\ \vdots \\ f_{B8}(r_{r_Mw_C}, b_{b_Nv_D}) \end{pmatrix} \end{bmatrix} \tag{5}$$

4. Algorithm design

It is necessary to study an algorithm to obtain the high-quality PNE solutions at a fast speed because of the ACT-DMU problem's real-time requirements. The general framework of the PNE-PRTC algorithm for the multicriteria game model is introduced in Subsection 4.1. Subsection 4.2 proves that the solution obtained by the PNE-PRTC algorithm is a refinement of the PNE solution.

4.1 PNE-PRTC algorithm

There are conflicts, coupling and incommensurability among the multiple objectives of the multicriteria game model. It is difficult to accurately obtain the opponent's objective weight. To solve the multicriteria game model, the PNE-PRTC is proposed in this paper. This algorithm is inspired by the PNE algorithm proposed in [32], and it can solve the PNE without weights. The preference relations are redefined in the PNE-PRTC algorithm. Furthermore, weak partial binary relation constraints and negative threshold constraints are added to construct a five-level domination criterion. A decisive step toward the optimal solution is provided for the PNE solution sets. The general framework of the PNE-PRTC algorithm is shown in Fig. 2.

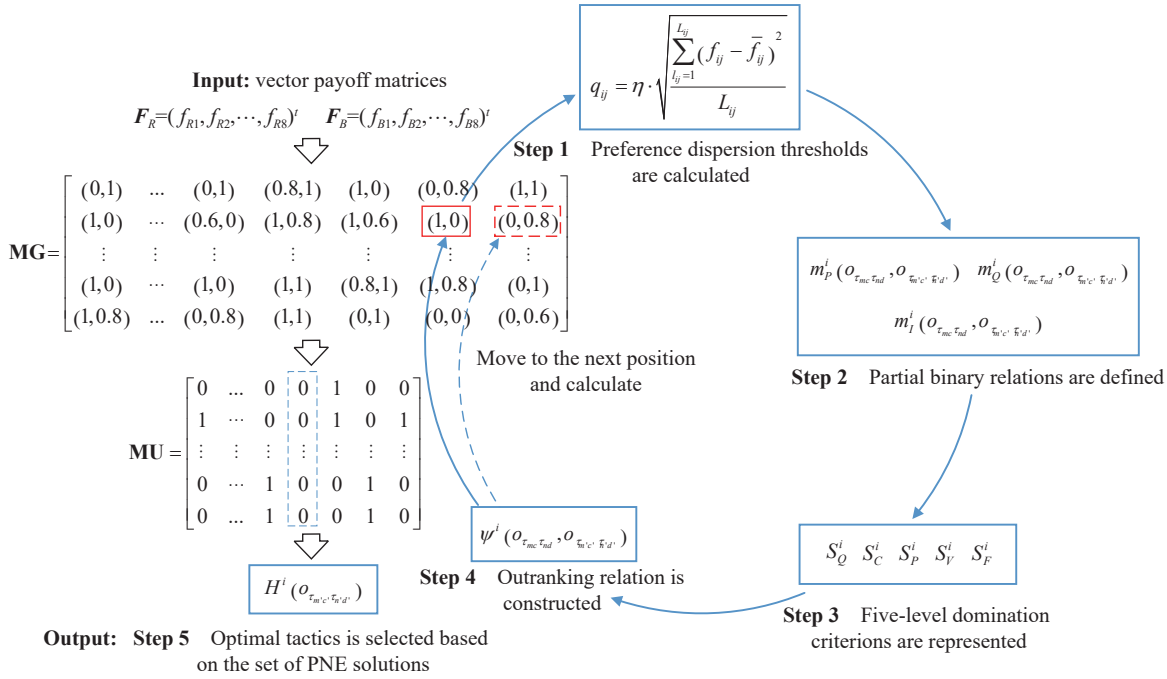


Fig. 2 General framework of the PNE-PRTC algorithm

First, the preference dispersion thresholds are calculated by extracting the discrete characteristics of the vector payoff matrices. Second, the partial binary relations are defined. Then, a five-level domination criterion is presented according to the combination of partial binary relations. Next, the outranking relation is constructed by the level of preferences restricted to the domination criterion, and the preference relations, as outranking relations corresponding to all strategy profiles, are represented by the values 0, 1, etc. Finally, the set of equilibrium solutions that are solved according to the strategies of each player are the best responses, and the optimal tactics are selected by a decision rule. The key steps in the PNE-PRTC algorithm are designed as follows:

Step 1 The preference dispersion thresholds are calculated by

$$q_{ij}=\eta\cdot\sqrt{\sum_{X_{ij}=1}^{X_{ij}}\left(f_{ij}-\bar{f}_{ij}\right)^2/X_{ij}} \tag{6}$$

where X_{ij} represents the number of samples selected for the jth objective function of i, $i=R,B$. η is defined as the threshold scaling coefficient, which ranges from 0.5 to 1.5. By multiplying the standard deviation with the threshold scaling coefficient, the dispersion degree characteristic of the vector payoff matrices is extracted as the preference dispersion threshold, which is used to measure i's sensitivity of the degree of outranking the vector payoffs in each objective.

Step 2 In Table 2, the partial binary relations for noncooperative situations are constructed. The partial binary relations consist of P_{ij}, Q_{ij} and I_{ij}, where $p_{ij}=2q_{ij}$ and $v_{ij}=3p_{ij}$ [38]. For simplicity, $o_{\tau_{mc}\tau_{nd}}=(r_{r_mw_c},b_{b_nv_d})$, $o_{\tau_{m'c'}\tau_{n'd'}}=(r_{r_{m'}w_{c'}},b_{b_{n'}v_{d'}})$, and $o_{\tau_{mc}\tau_{nd}},o_{\tau_{m'c'}\tau_{n'd'}}\in O$. Let $m_P^i(o_{\tau_{mc}\tau_{nd}},o_{\tau_{m'c'}\tau_{n'd'}})$ represent the number of partial binary relations for which $o_{\tau_{mc}\tau_{nd}}P_{ij}o_{\tau_{m'c'}\tau_{n'd'}}$. Similarly, let $m_Q^i(o_{\tau_{mc}\tau_{nd}},o_{\tau_{m'c'}\tau_{n'd'}})$ and $m_I^i(o_{\tau_{mc}\tau_{nd}},o_{\tau_{m'c'}\tau_{n'd'}})$ represent the numbers of partial binary relations for which $o_{\tau_{mc}\tau_{nd}}Q_{ij}o_{\tau_{m'c'}\tau_{n'd'}}$ and $o_{\tau_{mc}\tau_{nd}}I_{ij}o_{\tau_{m'c'}\tau_{n'd'}}$, respectively.

Table 2 Partial binary relations of noncooperative situations

Partial binary relation	Description	Constraint condition
$o_{\tau_{mc}\tau_{nd}}P_{ij}o_{\tau_{m'c'}\tau_{n'd'}}$	$o_{\tau_{mc}\tau_{nd}}$ is strongly preferred to $o_{\tau_{m'c'}\tau_{n'd'}}$	$f_{ij}\left(o_{\tau_{mc}\tau_{nd}}\right)>f_{ij}\left(o_{\tau_{m'c'}\tau_{n'd'}}\right)+p_{ij}$
$o_{\tau_{mc}\tau_{nd}}Q_{ij}o_{\tau_{m'c'}\tau_{n'd'}}$	$o_{\tau_{mc}\tau_{nd}}$ is weakly preferred to $o_{\tau_{m'c'}\tau_{n'd'}}$	$q_{ij}<f_{ij}\left(o_{\tau_{mc}\tau_{nd}}\right)-f_{ij}\left(o_{\tau_{m'c'}\tau_{n'd'}}\right)\leqslant p_{ij}$
$o_{\tau_{mc}\tau_{nd}}I_{ij}o_{\tau_{m'c'}\tau_{n'd'}}$	$o_{\tau_{mc}\tau_{nd}}$ is indifferent to $o_{\tau_{m'c'}\tau_{n'd'}}$	$-q_{ij}<f_{ij}\left(o_{\tau_{mc}\tau_{nd}}\right)-f_{ij}\left(o_{\tau_{m'c'}\tau_{n'd'}}\right)\leqslant q_{ij}$

Step 3 To establish an outranking relation, a five-level domination criterion is constructed.

Definition 1 First-level domination criterion. If $m_P^i(o_{\tau_{m'c'}\tau_{n'd'}},o_{\tau_{mc}\tau_{nd}})=0$ and $m_I^i(o_{\tau_{m'c'}\tau_{n'd'}},o_{\tau_{mc}\tau_{nd}})\leqslant 1+$

$m_I^i(o_{\tau_{mc}\tau_{nd}}, o_{\tau_{m'c'}\tau_{n'd'}}) + m_Q^i(o_{\tau_{mc}\tau_{nd}}, o_{\tau_{m'c'}\tau_{n'd'}}) + m_P^i(o_{\tau_{mc}\tau_{nd}}, o_{\tau_{m'c'}\tau_{n'd'}})$, then this criterion is established and is represented by $o_{\tau_{mc}\tau_{nd}} S_Q^i o_{\tau_{m'c'}\tau_{n'd'}}$.

Definition 2 Second-level domination criterion. If $m_P^i(o_{\tau_{m'c'}\tau_{n'd'}}, o_{\tau_{mc}\tau_{nd}}) = 0$, $m_Q^i(o_{\tau_{m'c'}\tau_{n'd'}}, o_{\tau_{mc}\tau_{nd}}) \leqslant 1 + J/5$, $m_Q^i(o_{\tau_{m'c'}\tau_{n'd'}}, o_{\tau_{mc}\tau_{nd}}) \leqslant m_P^i(o_{\tau_{m'c'}\tau_{n'd'}}, o_{\tau_{mc}\tau_{nd}})$, and $m_Q^i(o_{\tau_{m'c'}\tau_{n'd'}}, o_{\tau_{mc}\tau_{nd}}) + m_I^i(o_{\tau_{m'c'}\tau_{n'd'}}, o_{\tau_{mc}\tau_{nd}}) \leqslant 1 + m_I^i(o_{\tau_{mc}\tau_{nd}}, o_{\tau_{m'c'}\tau_{n'd'}}) + m_Q^i(o_{\tau_{mc}\tau_{nd}}, o_{\tau_{m'c'}\tau_{n'd'}}) + m_P^i(o_{\tau_{mc}\tau_{nd}}, o_{\tau_{m'c'}\tau_{n'd'}})$, then this criterion is established and is represented by $o_{\tau_{mc}\tau_{nd}} S_C^i o_{\tau_{m'c'}\tau_{n'd'}}$.

Definition 3 Third-level domination criterion. If $m_P^i(o_{\tau_{m'c'}\tau_{n'd'}}, o_{\tau_{mc}\tau_{nd}}) = 0$, $m_Q^i(o_{\tau_{m'c'}\tau_{n'd'}}, o_{\tau_{mc}\tau_{nd}}) \leqslant 1 + J/4$, $m_Q^i(o_{\tau_{m'c'}\tau_{n'd'}}, o_{\tau_{mc}\tau_{nd}}) \leqslant m_P^i(o_{\tau_{mc}\tau_{nd}}, o_{\tau_{m'c'}\tau_{n'd'}})$, and $m_Q^i(o_{\tau_{m'c'}\tau_{n'd'}}, o_{\tau_{mc}\tau_{nd}}) \leqslant 1 + m_Q^i(o_{\tau_{mc}\tau_{nd}}, o_{\tau_{m'c'}\tau_{n'd'}}) + m_P^i(o_{\tau_{mc}\tau_{nd}}, o_{\tau_{m'c'}\tau_{n'd'}})$ then this criterion is established and is represented by $o_{\tau_{mc}\tau_{nd}} S_P^i o_{\tau_{m'c'}\tau_{n'd'}}$.

Definition 4 Fourth-level domination criterion. If $m_P^i(o_{\tau_{m'c'}\tau_{n'd'}}, o_{\tau_{mc}\tau_{nd}}) = 1$, $m_Q^i(o_{\tau_{m'c'}\tau_{n'd'}}, o_{\tau_{mc}\tau_{nd}}) \leqslant 1 + J/2$, and negative threshold constraints $f_{ij}(o_{\tau_{m'c'}\tau_{n'd'}}) \geqslant f_{ij}(o_{\tau_{mc}\tau_{nd}}) + v_{ij}$ do not exist, then this criterion is established and is represented by $o_{\tau_{mc}\tau_{nd}} S_V^i o_{\tau_{m'c'}\tau_{n'd'}}$.

Definition 5 Fifth-level domination criterion. If $m_P^i\left(o_{\tau'_m c'\tau'_n d'}, o_{\tau_{mc}\tau_{nd}}\right) \leqslant 2$, $m_Q^i\left(o_{\tau_{m'c'}\tau_{n'd'}}, o_{\tau_{mc}\tau_{nd}}\right) \geqslant 1 + J/2$, and negative threshold constraints $f_{ij}\left(o_{\tau'_m c'\tau'_n d'}\right) \geqslant f_{ij}(o_{\tau_{mc}\tau_{nd}}) + v_{ij}$ do not exist, then this criterion is established and is represented by $o_{\tau_{mc}\tau_{nd}} S_F^i o_{\tau'_m c'\tau'_n d'}$.

In this paper, a five-level domination criterion is constructed by adding or reducing the constraints of partial binary relations based on the classical multi-attribute decision making theory [39]. Negative threshold constraints are added to the fourth-level and fifth-level domination criteria to prevent outranking relations with gaps that are too large from being screened out.

Step 4 The outranking relation S^i includes five levels of domination criterion S_Q^i, S_C^i, S_P^i, S_V^i and S_F^i, $S^i = \left\{S_Q^i, S_C^i, S_P^i, S_V^i, S_F^i\right\}$. The preference level $\psi^i(o_{\tau_{mc}\tau_{nd}}, o_{\tau_{m'c'}\tau_{n'd'}})$ of the vector payoff matrices is constructed as follow:

$$\psi^i(o_{\tau_{mc}\tau_{nd}}, o_{\tau_{m'c'}\tau_{n'd'}}) = \begin{cases} 1, & o_{\tau_{mc}\tau_{nd}} S_Q^i o_{\tau_{m'c'}\tau_{n'd'}} \\ 0.8, & o_{\tau_{mc}\tau_{nd}} S_C^i o_{\tau_{m'c'}\tau_{n'd'}} \\ 0.6, & o_{\tau_{mc}\tau_{nd}} S_P^i o_{\tau_{m'c'}\tau_{n'd'}} \\ 0.4, & o_{\tau_{mc}\tau_{nd}} S_V^i o_{\tau_{m'c'}\tau_{n'd'}} \\ 0.2, & o_{\tau_{mc}\tau_{nd}} S_F^i o_{\tau_{m'c'}\tau_{n'd'}} \\ 0, & \text{otherwise} \end{cases}. \quad (7)$$

In (8), a matrix of global information **MG**, which is composed of the values 0, 0.2, 0.4, 0.6, 0.8, and 1, is formed by gathering all the outranking relations corresponding to the links between every pair of strategy profiles. Then, a matrix of useful information **MU**, which consists of refined outranking relations, is constructed to extract the noncooperative equilibrium (NCE) solution by (9), according to the formula for searching for NCE solutions in [32].

$$\mathbf{MG}_{(o_{\tau_{mc}\tau_{nd}},\ o_{\tau_{m'c'}\tau_{n'd'}})} = \begin{cases} \left(\psi^R, \psi^B\right), & o_{\tau_{mc}\tau_{nd}} S^R o_{\tau_{m'c'}\tau_{n'd'}}, o_{\tau_{mc}\tau_{nd}} S^B o_{\tau_{m'c'}\tau_{n'd'}} \\ \left(0, \psi^B\right), & o_{\tau_{mc}\tau_{nd}} \bar{S}^R o_{\tau_{m'c'}\tau_{n'd'}}, o_{\tau_{mc}\tau_{nd}} S^B o_{\tau_{m'c'}\tau_{n'd'}} \\ \left(\psi^R, 0\right), & o_{\tau_{mc}\tau_{nd}} S^R o_{\tau_{m'c'}\tau_{n'd'}}, o_{\tau_{mc}\tau_{nd}} \overline{S}^B o_{\tau_{m'c'}\tau_{n'd'}} \\ (0,0), & \text{otherwise} \end{cases} \quad (8)$$

$$\mathbf{MU}_{(o_{\tau_{mc}\tau_{nd}},\ o_{\tau_{m'c'}\tau_{n'd'}})} = \begin{cases} 1, & r_{r_m w_c} \neq r_{r'_m w'_c}; b_{b_n v_d} = b_{b'_n v'_d}; \\ & \mathbf{MG}^R_{(o_{\tau_{mc}\tau_{nd}},\ o_{\tau_{m'c'}\tau_{n'd'}})} \geqslant \xi \\ & \text{or } r_{r_m w_c} = r_{r'_m w'_c}; b_{b_n v_d} \neq b_{b'_n v'_d}; \\ & \mathbf{MG}^B_{(o_{\tau_{mc}\tau_{nd}},\ o_{\tau_{m'c'}\tau_{n'd'}})} \geqslant \xi \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

where $\mathbf{MG}_{(o_{\tau_{mc}\tau_{nd}}, o_{\tau_{m'c'}\tau_{n'd'}})}$ is a cell of **MG** which is the vector payoffs corresponding to the intersection of the row $o_{\tau_{mc}\tau_{nd}}$ and the column $o_{\tau_{m'c'}\tau_{n'd'}}$, and $\mathbf{MG}_{(o_{\tau_{mc}\tau_{nd}}, o_{\tau_{m'c'}\tau_{n'd'}})} = \left(\mathbf{MG}^R_{(o_{\tau_{mc}\tau_{nd}}, o_{\tau_{m'c'}\tau_{n'd'}})}, \mathbf{MG}^B_{(o_{\tau_{mc}\tau_{nd}}, o_{\tau_{m'c'}\tau_{n'd'}})}\right)$.

Similarly, $\mathbf{MU}_{(o_{\tau_{mc}\tau_{nd}}, o_{\tau_{m'c'}\tau_{n'd'}})}$ is a cell of **MU**.

$o_{\tau_{mc}\tau_{nd}} S^i o_{\tau_{m'c'}\tau_{n'd'}}$ denotes that $o_{\tau_{mc}\tau_{nd}}$ is preferred to $o_{\tau_{m'c'}\tau_{n'd'}}$ by i, while $o_{\tau_{mc}\tau_{nd}} \bar{S}^i o_{\tau_{m'c'}\tau_{n'd'}}$ means that $o_{\tau_{mc}\tau_{nd}}$ is not preferred to $o_{\tau_{m'c'}\tau_{n'd'}}$ by i. ξ is defined as the maximum value of the preference level that is allowed, $\xi \in \{0.2, 0.4, 0.6, 0.8, 1\}$; the larger the value of ξ is, the greater the number of NCE solutions, and to be conservative, we set ξ=0.6. When all elements of a column in **MU** are equal to 0, the strategy profile corresponding to that column is an NCE solution of the multicriteria game.

Step 5 If the number of NCE solutions calculated in Step 4 is more than one, only one NCE solution will be selected according to the maximum hypervolume rule:

$$H^i(o_{\tau_{m'c'}\tau_{n'd'}}) = f_{i1}(o_{\tau_{m'c'}\tau_{n'd'}}) \times f_{i2}(o_{\tau_{m'c'}\tau_{n'd'}}) \times \cdots \times f_{iJ}(o_{\tau_{m'c'}\tau_{n'd'}}). \quad (10)$$

On the basis of the definition of a hypervolume in multiobjective optimization theory, the $H^i(o_{\tau_{m'c'}\tau_{n'd'}})$ value in (10) reflects the solution quality, which includes the convergence and diversity of solutions. Finally, a pair of strategy profiles $o_{\tau_{m'c'}\tau_{n'd'}}$ with the maximum value of $H(o_{\tau_{m'c'}\tau_{n'd'}}) = H^R(o_{\tau_{m'c'}\tau_{n'd'}}) + H^B(o_{\tau_{m'c'}\tau_{n'd'}})$ is selected, and the optimal tactics $r_{m'}$ corresponding to the pair of strategy profiles are executed by R.

In summary, Step 1 to Step 4 demonstrate the process of the PNE-PRTC algorithm. The role of Step 5 is to make the final decision from the PNE solution sets, and

this step can also be flexibly replaced by other decision approaches based on the decision maker's preferences.

4.2 Proof of the PNE solution

The equilibrium solutions obtained by the PNE-PRTC algorithm are the NCE solutions in Subsection 4.1. In this section, we need to prove that the NCE solutions are refinements of the PNE solution. First, the definitions involved are summarized [32,40].

Definition 6 For $\boldsymbol{a},\boldsymbol{b}\in\mathbf{R}^n$, the notation of vector inequalities is defined as follows:

$$\boldsymbol{a}\geqq\boldsymbol{b}\iff a_i\geqq b_i \text{ for all } i\in\{1,2,\cdots,n\}.$$

$$\boldsymbol{a}\geqslant\boldsymbol{b}\iff a_i\geqq b_i \text{ and } a\neq b.$$

$$\boldsymbol{a}>\boldsymbol{b}\iff a_i>b_i \text{ for all } i\in\{1,2,\cdots,n\}.$$

Definition 7 A pair of strategy profiles $(x^*,y^*)\in X\times Y$ is called a PNE solution for a multicriteria game, and it needs to meet:

$$\begin{cases}\nexists x\in X,\ F_1(x,y^*)\geqslant F_1(x^*,y^*)\\ \nexists y\in Y,\ F_2(x^*,y)\geqslant F_2(x^*,y^*)\end{cases}\tag{11}$$

where $F_i(x,y)=(f_{i1}(x,y),f_{i2}(x,y),\cdots,f_{iN_i}(x,y))^t$ denotes vector payoff functions and N_i is the number of objectives, $i=1,2$.

Definition 8 $F_i(x,y)\geqslant F_i(x',y')$ denotes that a pair of strategy profiles $(x,y)\in X\times Y$ dominates another pair of strategy profiles $(x',y')\in X\times Y$ for player P_i $(i=1,2)$. $F_i(x,y)\geqslant F_i(x',y')$ can also be represented as $(x,y)D^i(x',y')$. D^i represents a dominance relation.

Definition 9 A pair of strategy profiles $(x^*,y^*)\in X\times Y$ is called an NCE solution for a multicriteria game, and it needs to meet

$$\begin{cases}\nexists x\in X,\ (x,y^*)S^1(x^*,y^*)\\ \nexists y\in Y,\ (x^*,y)S^2(x^*,y^*)\end{cases}\tag{12}$$

where S^i represents a binary relation.

Lemma 1 If $D^i\subseteq S^i$, the set of NCE solutions for a multicriteria game is included by the PNE solution sets and can be expressed as NCE⊆PNE.

This is proved in detail in [32].

Lemma 2 For the PNE-PRTC algorithm, $D^i\subseteq S^i$ $(i=1,2)$.

Proof We proceed by contradiction. Assume that there exist two pairs of strategy profiles $o_{x,y}=(x,y)$ and $o_{x',y'}=(x',y')$, where $o_{x,y},o_{x',y'}\in X\times Y$, such that "$o_{x,y}D^io_{x',y'}$ exists" and "$o_{x,y}S^io_{x',y'}$ does not exist".

Suppose that $o_{x,y}D^io_{x',y'}$, we can infer that $F_i(o_{x,y})\geqslant F_i(o_{x',y'})$ from Definition 7 and Definition 8. Thus, $\forall j\in N_i$, $f_{ij}(o_{x,y})\geqq f_{ij}(o_{x',y'})$, and $\exists j_0\in N_i$ such that $f_{ij_0}(o_{x,y})>f_{ij_0}(o_{x',y'})$ from Definition 6. We can also deduce that $\nexists j\in N_i$ such that $f_{ij}(o_{x,y})<f_{ij}(o_{x',y'})$, which implies that $m_P^i(o_{x',y'},o_{x,y})=0$ and $m_Q^i(o_{x',y'},o_{x,y})=0$. The details are discussed in different value ranges for $f_{ij}(o_{x,y})$ and $f_{ij}(o_{x',y'})$.

(i) $\exists j_0\in N_i$ such that $0<f_{ij_0}(o_{x,y})-f_{ij_0}(o_{x',y'})\leqslant q_{ij_0}$, this implies that $m_I^i\left(o_{x,y},o_{x',y'}\right)\geqslant 1$, $m_I^i\left(o_{x,y},o_{x',y'}\right)=m_I^i(o_{x',y'},o_{x,y})$, $m_Q^i\left(o_{x,y},o_{x',y'}\right)=0$ and $m_P^i\left(o_{x,y},o_{x',y'}\right)=0$. Thus, $m_I^i(o_{x',y'},o_{x,y})\leqslant 1+m_I^i\left(o_{x,y},o_{x',y'}\right)$, and it follows that $o_{x,y}S_Q^io_{x',y'}$.

(ii) $\exists j_0\in N_i$ such that $q_{ij_0}<f_{ij_0}(o_{x,y})-f_{ij_0}(o_{x',y'})\leqslant p_{ij_0}$, this implies that $m_Q^i\left(o_{x,y},o_{x',y'}\right)\geqslant 1$, $m_I^i\left(o_{x,y},o_{x',y'}\right)=m_I^i(o_{x',y'},o_{x,y})\geqslant 0$ and $m_P^i\left(o_{x,y},o_{x',y'}\right)=0$. Thus, $m_I^i(o_{x',y'},o_{x,y})\leqslant 1+m_I^i\left(o_{x,y},o_{x',y'}\right)+m_Q^i\left(o_{x,y},o_{x',y'}\right)$, and it follows that $o_{x,y}\cdot S_Q^io_{x',y'}$.

(iii) $\exists j_0\in N_i$ such that $f_{ij_0}(o_{x,y})>f_{ij_0}(o_{x',y'})+q_{ij_0}$, this implies that $m_P^i\left(o_{x,y},o_{x',y'}\right)\geqslant 1$, $m_I^i\left(o_{x,y},o_{x',y'}\right)=m_I^i(o_{x',y'},o_{x,y})\geqslant 0$ and $m_Q^i\left(o_{x,y},o_{x',y'}\right)=0$. Thus, $m_I^i(o_{x',y'},o_{x,y})\leqslant 1+m_I^i\left(o_{x,y},o_{x',y'}\right)+m_P^i\left(o_{x,y},o_{x',y'}\right)$, and it follows that $o_{x,y}\cdot S_Q^io_{x',y'}$.

(iv) $\exists j_0\in N_i$ such that $0<f_{ij_0}(o_{x,y})-f_{ij_0}(o_{x',y'})\leqslant q_{ij_0}$ and $\exists j_1\in N_i$ such that $q_{ij_1}<f_{ij_1}(o_{x,y})-f_{ij_1}(o_{x',y'})\leqslant p_{ij_1}$, this implies that $m_I^i\left(o_{x,y},o_{x',y'}\right)\geqslant 1$, $m_I^i\left(o_{x,y},o_{x',y'}\right)\geqslant m_I^i(o_{x',y'},o_{x,y})$, $m_Q^i\left(o_{x,y},o_{x',y'}\right)\geqslant 1$ and $m_P^i\left(o_{x,y},o_{x',y'}\right)=0$. Thus, $m_I^i(o_{x',y'},o_{x,y})\leqslant 1+m_I^i\left(o_{x,y},o_{x',y'}\right)+m_Q^i\left(o_{x,y},o_{x',y'}\right)$, and it follows that $o_{x,y}S_Q^io_{x',y'}$.

(v) $\exists j_0\in N_i$ such that $0<f_{ij_0}(o_{x,y})-f_{ij_0}(o_{x',y'})\leqslant q_{ij_0}$ and $\exists j_1\in N_i$ such that $f_{ij_1}(o_{x,y})>f_{ij_1}(o_{x',y'})+p_{ij_1}$, this implies that $m_I^i\left(o_{x,y},o_{x',y'}\right)\geqslant 1$, $m_I^i\left(o_{x,y},o_{x',y'}\right)\geqslant m_I^i(o_{x',y'},o_{x,y})$, $m_P^i\left(o_{x,y},o_{x',y'}\right)\geqslant 1$ and $m_Q^i\left(o_{x,y},o_{x',y'}\right)=0$. Thus, $m_I^i(o_{x',y'},o_{x,y})\leqslant 1+m_I^i\left(o_{x,y},o_{x',y'}\right)+m_Q^i\left(o_{x,y},o_{x',y'}\right)$, and it follows that $o_{x,y}S_Q^io_{x',y'}$.

(vi) $\exists j_0\in N_i$ such that $q_{ij_0}<f_{ij_0}(o_{x,y})-f_{ij_0}(o_{x',y'})\leqslant p_{ij_0}$ and $\exists j_1\in N_i$ such that $f_{ij_1}(o_{x,y})>f_{ij_1}(o_{x',y'})+p_{ij_1}$, this implies that $m_P^i\left(o_{x,y},o_{x',y'}\right)\geqslant 1$, $m_Q^i\left(o_{x,y},o_{x',y'}\right)\geqslant 1$ and $m_I^i\left(o_{x,y},o_{x',y'}\right)=m_I^i(o_{x',y'},o_{x,y})\geqslant 0$. Thus, $m_I^i(o_{x',y'},o_{x,y})\leqslant 1+m_I^i\left(o_{x,y},o_{x',y'}\right)+m_Q^i\left(o_{x,y},o_{x',y'}\right)+m_P^i\left(o_{x,y},o_{x',y'}\right)$, and it follows that $o_{x,y}S_Q^io_{x',y'}$.

(vii) $\exists j_0\in N_i$ such that $0<f_{ij_0}(o_{x,y})-f_{ij_0}(o_{x',y'})\leqslant q_{ij_0}$, $\exists j_1\in N_i$ such that $q_{ij_1}<f_{ij_1}(o_{x,y})-f_{ij_1}(o_{x',y'})\leqslant p_{ij_1}$ and $\exists j_2\in N_i$ such that $f_{ij_2}(o_{x,y})>f_{ij_2}(o_{x',y'})+p_{ij_2}$, this implies that $m_I^i\left(o_{x,y},o_{x',y'}\right)\geqslant 1$, $m_I^i\left(o_{x,y},o_{x',y'}\right)=m_I^i(o_{x',y'},o_{x,y})$, $m_P^i\left(o_{x,y},o_{x',y'}\right)\geqslant 1$ and $m_Q^i\left(o_{x,y},o_{x',y'}\right)\geqslant 1$. Thus, $m_I^i(o_{x',y'},o_{x,y})\leqslant 1+m_I^i\left(o_{x,y},o_{x',y'}\right)+m_Q^i\left(o_{x,y},o_{x',y'}\right)+m_P^i\left(o_{x,y},o_{x',y'}\right)$, and it follows that $o_{x,y}S_Q^io_{x',y'}$.

There exist $S_Q^i\subseteq S^i$, which implies that $o_{x,y}S^io_{x',y'}$. This contradicts the fact that $o_{x,y}S^io_{x',y'}$ does not exist. □

5. Experiments

To verify the generality and efficiency of the PNE-PRTC algorithm, numerical experiments are carried out in Subsection 5.1. Simulation experiments are analyzed in Subsection 5.2 to verify the effectiveness and correctness of the multicriteria game model. All experiments are run on computers with an Intel i7-7700 3.60 GHz CPU and 32 GB RAM.

5.1 Numerical experiments

The parametric analysis of the PNE-PRTC algorithm is performed first. Then, the performance metrics of the PNE-PRTC algorithm considering the optimal parameters are employed. Three different datasets represented by Set 1, Set 2, and Set 3 are generated by the leading game generator GAMUT [41]. Each dataset includes 15 different sizes of instances, and each size includes 20 instances of a multicriteria game, as shown in Table 3. The vector-valued payoffs for the multicriteria games range from 0 to 1. $\alpha\times\beta\times\chi$ represents multicriteria games with the same size. α and β denote the numbers of R's strategies and B's strategies. χ represents the objective number for each side.

5.1.1 Performance metrics

Unlike matrix games, which evaluate the quality of Nash equilibrium solutions by comparing payoffs with the same dimension, it is difficult to measure the quality of the solutions with a single performance metric for multicriteria games because of payoffs that are vector valued. Similar to the performance metrics of multiobjective optimization, in this paper, the quality of PNE solutions is measured in three respects: convergence, uniformity and spread. The performance metrics of the multicriteria game are described as shown in (13), (14), and (15).

Definition 10 Average hypervolume for a multicriteria game [42] can be expressed as

$$W_{\mathrm{HV}}=\frac{1}{S_{\mathrm{set}}}\sum_{s_{\mathrm{set}}=1}^{S_{\mathrm{set}}}\bigcup_{n_{\mathrm{PNE}}\in N_{\mathrm{PNE}}}^{N_{\mathrm{PNE}}}\mathrm{hv}\left(n_{\mathrm{PNE}},p_{\mathrm{PNE}}\right)\tag{13}$$

where S_{set} denotes vector payoff matrices with the same size, $S_{\mathrm{set}}=20$. N_{PNE} represents the PNE solution set obtained by the PNE-PRTC algorithm, and n_{PNE} is a PNE solution, $n_{\mathrm{PNE}}\in N_{\mathrm{PNE}}$. p_{PNE} denotes a reference point. hv indicates the hypervolume value enclosed by N_{PNE} and p_{PNE}. The quality of the PNE solution set in terms of the convergence and diversity of the vector payoffs is evaluated in (13). A larger value of W_{HV} means that the PNE solution sets are more convergent and diverse.

Definition 11 Average spacing metric for the multicriteria game [43] can be expressed as

$$W_{\mathrm{SP}}=\frac{1}{S_{\mathrm{set}}}\sum_{s_{\mathrm{set}}=1}^{S_{\mathrm{set}}}\sqrt{\frac{1}{|N_{\mathrm{PNE}}|}\sum_{n_{\mathrm{PNE}}=1}^{|N_{\mathrm{PNE}}|}\left(d'_{n_{\mathrm{PNE}}}-\bar{d}\right)^2}\tag{14}$$

where $d_{n'_{\mathrm{PNE}}}=\min\limits_{n_{\mathrm{PNE}}\in N_{\mathrm{PNE}},n_{\mathrm{PNE}}\neq n'_{\mathrm{PNE}}}\sum_{j=1}^{J}\left|f_j^{n_{\mathrm{PNE}}}-f_j^{n'_{\mathrm{PNE}}}\right|$ denotes the Euclidean distance between the n_{PNE}th PNE solutions. $\bar{d}=\sum_{n_{\mathrm{PNE}}=1}^{|N_{\mathrm{PNE}}|}d_{n'_{\mathrm{PNE}}}/|N_{\mathrm{PNE}}|$ represents the average value of $d_{n'_{\mathrm{PNE}}}$. $f_j^{n_{\mathrm{PNE}}}$ and $f_j^{n'_{\mathrm{PNE}}}$ denote the payoffs of the n_{PNE}th and n'_{PNE}th PNE solutions for the jth objective, respectively. J is the number of objectives. W_{SP} is used to measure the degree of distribution of the PNE solution set. A lower value of W_{SP} means that the distribution of the PNE solution set is more uniform.

Definition 12 Average diversity measure for the multicriteria game [44] can be expressed as

$$W_D=\frac{1}{S_{\mathrm{set}}}\sum_{s_{\mathrm{set}}=1}^{S_{\mathrm{set}}}\sqrt{\sum_{j=1}^{J}\left(\max_{n_{\mathrm{PNE}}=1}^{|N_{\mathrm{PNE}}|}f_j^{n_{\mathrm{PNE}}}-\min_{n_{\mathrm{PNE}}=1}^{|N_{\mathrm{PNE}}|}f_j^{n_{\mathrm{PNE}}}\right)^2}.\tag{15}$$

W_D is used to measure the diagonal length of the hypercube formed by the extreme objective values in a PNE solution set. It is a basic indicator that can reflect the diversity of a PNE solution set. Clearly, the higher the value of W_D is, the better the distribution of the PNE solution set.

5.1.2 Parametric analysis

The threshold scaling coefficient η is a key parameter that strongly affects the performance of the PNE-PRTC algorithm, and parametric analyses must be carried out to optimize its performance. We change η from 0.5 to 1.5 with a step size of 0.1. A total of 9 900 parametric analyses are conducted, with 20 multicriteria games conducted for each η in different sizes of instances. The quality of the PNE solution set is evaluated by three metrics: the average hypervolume W_{HV}, average spacing metric W_{SP} and average diversity measure W_D. Fig. 3 shows the results of numerical experiments for the PNE-PRTC algorithm with different η. Specifically, Fig. 3 (a), Fig. 3 (b), and Fig. 3 (c) represent the evaluation of R's vector payoffs corresponding to the PNE solutions for Set 1. Similarly, Fig. 3(d), Fig. 3(e), Fig. 3(f) and Fig. 3(g), Fig. 3(h), Fig. 3(i) represent the evaluation of R's vector payoffs corresponding to the PNE solutions for Set 2 and Set 3, respectively.

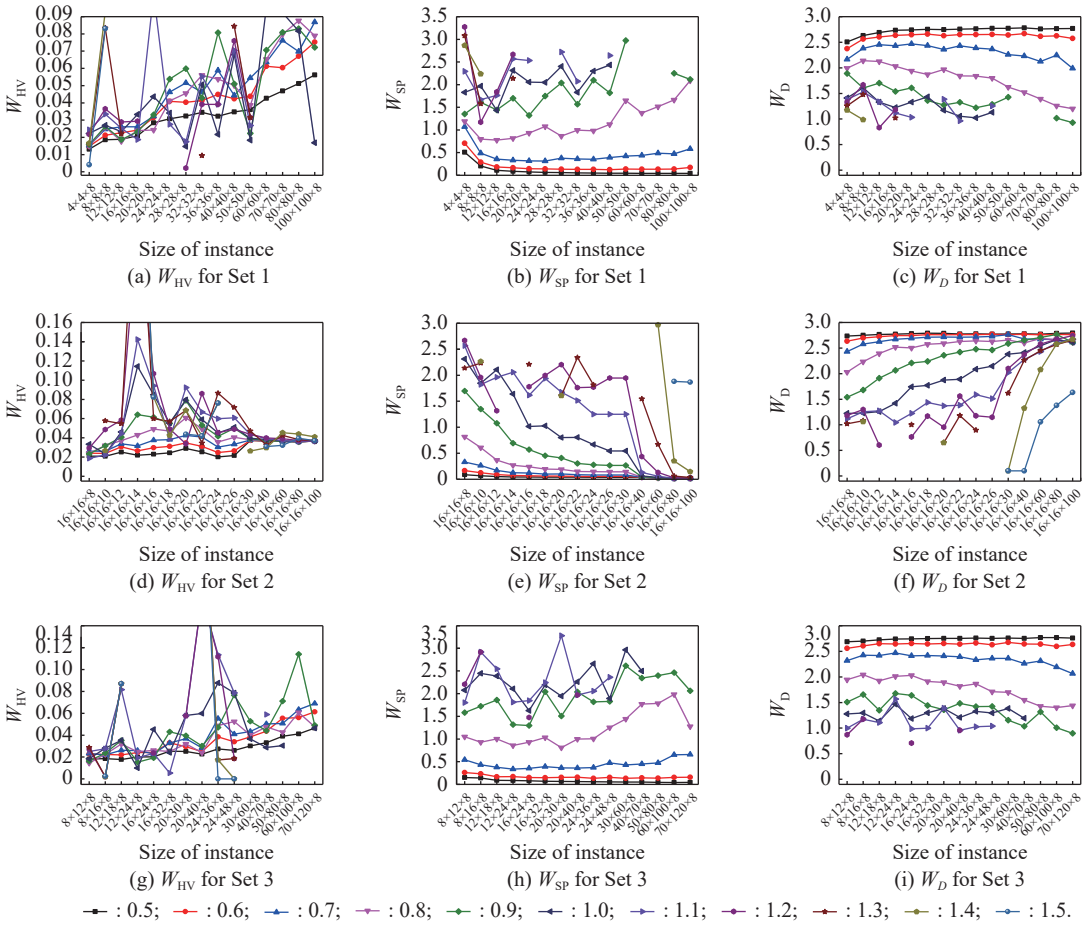


Fig. 3 Comparison of the results of numerical experiments for the PNE-PRTC algorithm with different η

Generally, in all subplots of Fig. 3, as the η value increases, the W_{HV} and W_{SP} values for multicriteria games obtained by using the PNE-PRTC algorithm gradually increase, while W_D gradually decreases. There is a certain positive correlation between the optimal value of the W_{SP} curve and the W_D curve. Nevertheless, there is a certain negative correlation between the optimal value of the W_{HV} curve and the W_{SP} curve. The correlation among the three performance metrics makes it difficult to achieve the optimal value for the metrics at the same time by adjusting η, but η can be adjusted to a compromise value according to the decision maker's preference on different performance metrics. In addition, when the value of η is greater than 0.9, as η increases, the variation ranges of W_{HV}, W_{SP} and W_D are significant. This finding implies that the larger the value of η, the greater the preference dispersion threshold q_{ij}, the more difficult it is to construct a binary relation S^i, and the smaller the number of PNE solutions calculated by the PNE-PRTC algorithm. When the number of PNE solutions drops to 0, W_{HV} suddenly falls to 0, and W_{SP} and W_D cannot be calculated, leading to unstable solution results. This eventually leads to the fluctuation and fracture of the curve.

For three different datasets, the performance metrics of the PNE-PRTC algorithm are analyzed in detail as follows:

(i) The W_{HV} curves corresponding to the datasets are shown in Fig. 3 (a), Fig. 3 (d), and Fig. 3 (g). When η is fixed, W_{HV} for Set 1 and Set 3 increases significantly as the number of strategies increases, while W_{HV} for Set 2 does not grow sharply with the increase in the number of objectives. The reason is that as the number of strategies increases, the strategy profiles increase significantly, the number of partial binary relations in Step 2 of the PNE-PRTC algorithm rises sharply, and the PNE solution sets obtained by the algorithm become larger. This makes it easier for the algorithm to balance the convergence and diversity of the PNE solution sets. The increase in the objectives does not increase the strategy profiles but increases the complexity of constructing preference relations for all strategy profiles. In addition, with the increase in the objects of instances in Set 2, the fluctuation range of the W_{HV} value of the PNE-PRTC algorithm gradually decreases, especially for large-scale instances.

The reason for this is that the PNE solution sets calculated by the PNE-PRTC algorithm are larger and more uniform for solving large-scale instances.

(ii) The W_{SP} curves corresponding to the datasets are shown in Fig. 3 (b), Fig. 3 (e) and Fig. 3 (h). Similarly, the W_D curves are shown in Fig. 3 (c), Fig. 3 (f) and Fig. 3 (i). When η is fixed, compared with Set 1 and Set 3, as the number of objectives in Set 2 increases, the trend of decreasing W_{SP} and increasing W_D becomes more obvious, especially when solving the large-scale instances. The reason is that as the number of objectives increases, it becomes more difficult to construct binary relations for each pair of vector payoffs by using the PNE-PRTC algorithm, the PNE solutions become more harmonious across vector payoffs, and the distribution of PNE solutions becomes more uniform and diverse.

Considering the positive and negative correlations among the W_{HV}, W_{SP} and W_D curves in Fig. 3, the η values of the PNE-PRTC algorithm for different datasets are shown in Table 3.

Table 3 Threshold scaling coefficient settings of the PNE-PRTC algorithm for different datasets

Number	Set 1	η	Set 2	η	Set 3	η
1	4×4×8	0.6	16×16×8	0.6	8×12×8	0.7
2	8×8×8	0.7	16×16×10	0.7	8×16×8	0.7
3	12×12×8	0.7	16×16×12	0.8	12×18×8	0.7
4	16×16×8	0.7	16×16×14	0.8	12×24×8	0.7
5	20×20×8	0.7	16×16×16	0.8	16×24×8	0.7
6	24×24×8	0.7	16×16×18	0.8	16×32×8	0.7
7	28×28×8	0.7	16×16×20	0.8	20×30×8	0.7
8	32×32×8	0.7	16×16×22	0.9	20×40×8	0.7
9	36×36×8	0.7	16×16×24	0.9	24×36×8	0.7
10	40×40×8	0.7	16×16×26	0.9	24×48×8	0.7
11	50×50×8	0.7	16×16×30	1.0	30×60×8	0.7
12	60×60×8	0.7	16×16×40	1.0	40×70×8	0.7
13	70×70×8	0.7	16×16×60	1.0	50×80×8	0.7
14	80×80×8	0.7	16×16×80	1.0	60×100×8	0.6
15	100×100×8	0.7	16×16×100	1.0	70×120×8	0.6

5.1.3 Comparison experiments

In this subsection, the proposed algorithm is compared with the algorithm of PNE based on the preference relation (PNE-E, PNE-P) proposed in [32] and the algorithm of PNE based on the scalarized single criterion (PNE-S) presented in [24]. The weights of the PNE-E, PNE-P, and PNE-S algorithms are set to the same value according to the principle of maximum entropy [45]. A total of 45 sizes of instances are solved on three datasets. For each size of instance, 20 multicriteria game instances are solved when the value is chosen according to Table 3. The average computational times (ACTs) of the PNE-S, PNE-E, PNE-P, and PNE-PRTC algorithms are shown in Fig. 4. Three performance metrics, W_{HV}, W_{SP} and W_D, of these algorithms are shown in Fig. 5.

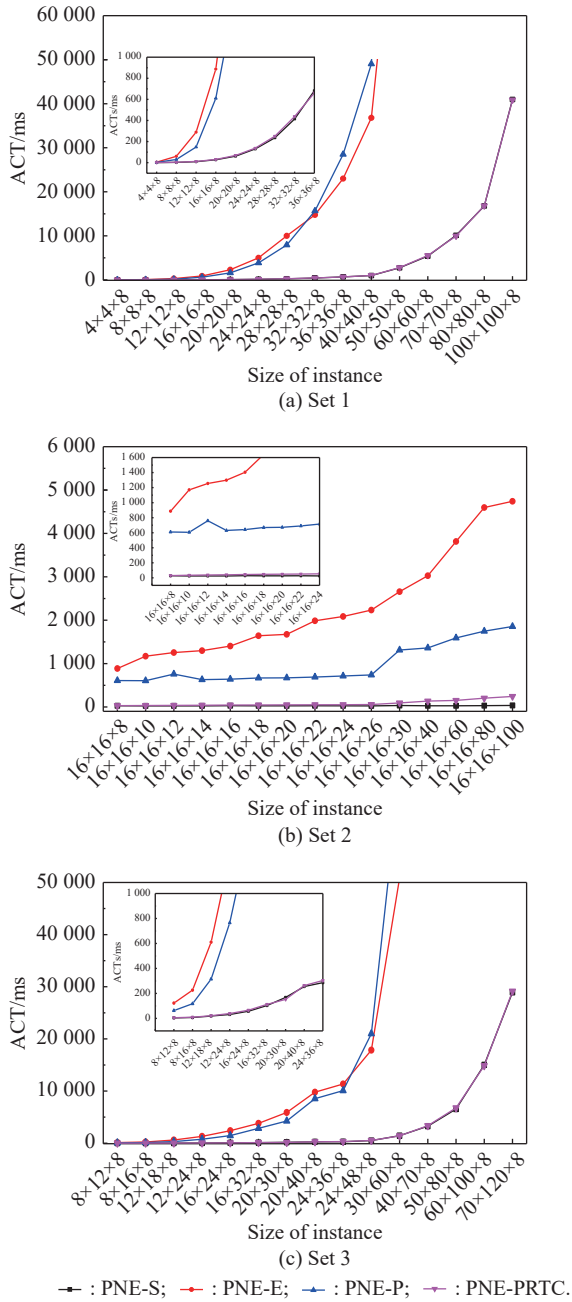


Fig. 4 ACTs of the PNE-S, PNE-E, PNE-P, and PNE-PRTC algorithms

From the overall trends shown in Fig. 4, it is clear that the ACT of the PNE-PRTC algorithm is less than those of the PNE-E and PNE-P algorithms, and the ACT curves of the PNE-PRTC and PNE-S algorithms almost overlap. To

solve the large-scale instances, the solution time of PNE-E and PNE-P algorithms increases exponentially with increasing instance size. When the size of instances in Set 1 is greater than 60×60×8, or the size of instances in Set 3 is greater than 50×80×8, the PNE solution sets cannot be calculated within six hours by using the above baseline algorithms. However, the effective PNE solution sets of the large-scale instances can be solved stably by the PNE-PRTC and PNE-E algorithms within seconds.

The reason for this is that it is necessary for the PNE-E and PNE-P algorithms to calculate the deviation distance for every objective and sort them when constructing the binary relation for a pair of strategy profiles. The original sorting needs to be updated for each binary relation. This repetitive sorting can cost much time. For the PNE-PRTC algorithm, the binary relations are constructed by the preference dispersion thresholds, which can reflect the discrete characteristics of the vector payoffs for each objective function. Repetitive sorting steps are not needed. Although the PNE-E, PNE-P, and PNE-PRTC algorithms can all solve multicriteria games based on preference relations, the PNE-PRTC algorithm can be even more efficient than the PNE-E and PNE-P algorithms.

The ACT of the PNE-PRTC algorithm on the three datasets is analyzed as follows: In Fig. 4 (a) and Fig. 4 (c), the ACT on Set 1 and Set 3 gradually increases as the number of strategies increases, while the ACT on Set 2 increases slightly as the number of objectives increases in Fig. 4 (b). The reason is that as the number of strategies increases in Set 1 and Set 3, the search space of the PNE solutions rises rapidly. For Set 2, as the number of objectives increases, the search space of PNE solutions does not change, but it is more difficult to construct preference relations for the strategy profiles. The slight increase in the ACT on Set 2 illustrates that the PNE-PRTC algorithm is able to construct binary relations efficiently for a pair of strategy profiles even in high-dimensional objectives. The ACT of the PNE-PRTC algorithm is more sensitive to the increase in the number of strategies than the number of objectives at the sizes selected in this paper.

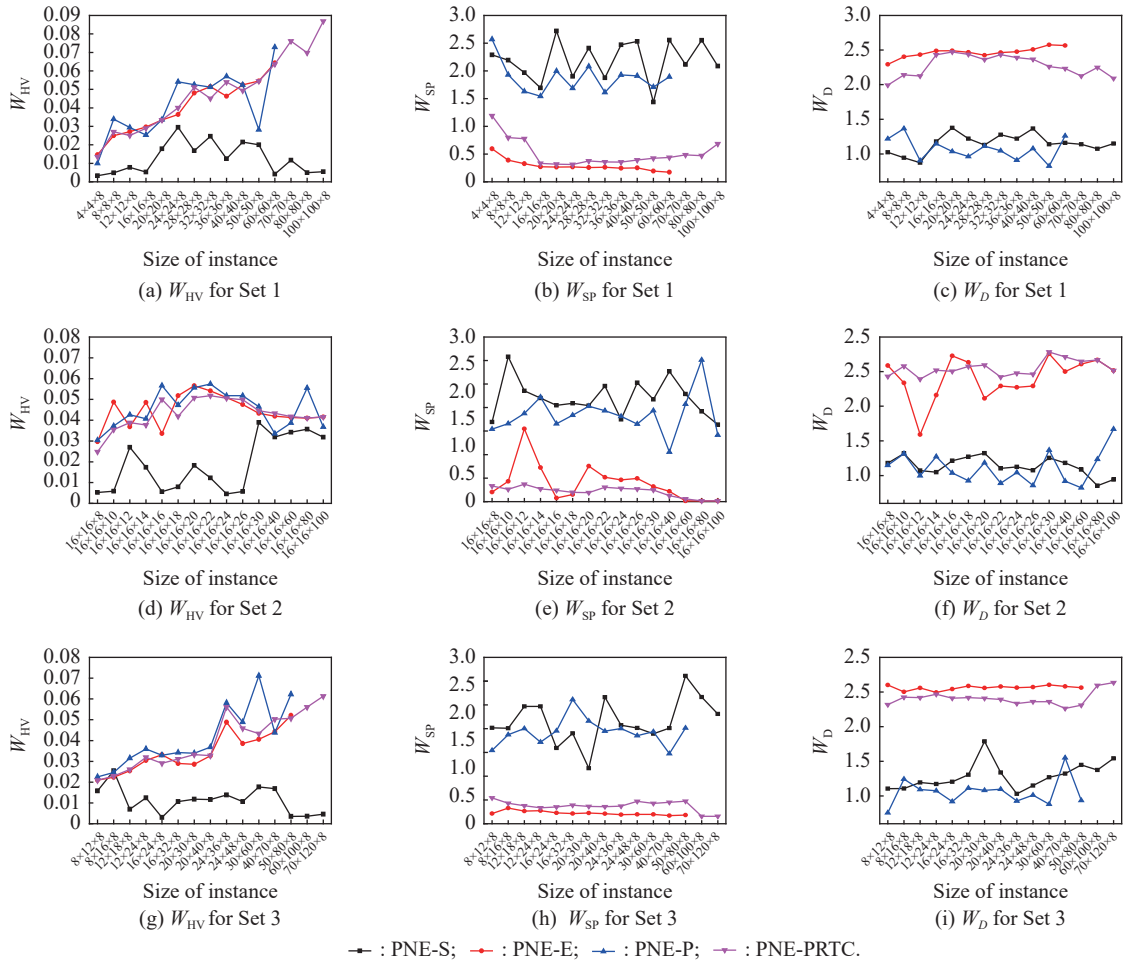


Fig. 5 Comparison of the results of the performance experiments for the PNE-S, PNE-E, PNE-P, and PNE-PRTC algorithms

Generally, in Fig. 5 (a), Fig. 5 (d), and Fig. 5 (g), the W_{HV} curves of the PNE-E, PNE-P, and PNE-PRTC algorithms are almost higher than the curves of the PNE-S algorithm. In Fig. 5 (b), Fig. 5(e), Fig. 5(h) and Fig. 5 (c), Fig. 5(f), Fig. 5(i), the overall trends of the W_{SP} and W_D curves of the PNE-S and PNE-P algorithms are roughly the same. In all subplots of Fig. 5, the W_{HV}, W_{SP} and W_D curves of the PNE-E and PNE-PRTC algorithms have similar trends and even overlap with each other. To solve the large-scale instances, high-quality PNE solution sets can be calculated stably by the PNE-PRTC algorithm. We can conclude that the performances of the PNE-PRTC algorithm and PNE-E algorithm are very close, and the PNE-PRTC algorithm outperforms the PNE-S algorithm in all three metrics W_{HV}, W_{SP}, and W_D.

For the instances that can be solved by the baseline algorithms, the three performance metrics W_{HV}, W_{SP}, and W_D of the PNE-E, PNE-P, and PNE-PRTC algorithms are analyzed in detail below.

(i) For the performance metric W_{HV} shown in Fig. 5 (a), Fig. 5(d) and Fig. 5(g), the W_{HV} of the PNE-PRTC algorithm improves by 319.9% on average compared with that of the PNE-S algorithm and by 5.6% on average compared with that of the PNE-E algorithm, and it is slightly lower than that of the PNE-P algorithm. However, the variation ranges of W_{HV} for the PNE-PRTC algorithm are narrower than those for the PNE-P algorithm, which indicates that the PNE-PRTC algorithm is more stable.

(ii) For the performance metric W_{SP} shown in Fig. 5 (b), Fig. 5(e) and Fig. 5(h), with different sizes of instances, the W_{SP} of the PNE-PRTC algorithm decreases by 82.1% on average compared with that of the PNE-S algorithm and by 79.8% on average compared with that of the PNE-P algorithm. The W_{SP} curves for the PNE-PRTC algorithm are smoother than the other curves, indicating that the PNE-PRTC algorithm has a more uniform distribution of PNE solution sets than the PNE-S algorithm and the PNE-P algorithm.

(iii) For the performance metric W_D shown in Fig. 5 (c), Fig. 5(f), Fig. 5(i), the W_D of the PNE-PRTC algorithm improves by 106.4% on average compared with that of the PNE-S algorithm and by 129.7% on average compared with that of the PNE-P algorithm. The small variation ranges of W_D for the PNE-PRTC algorithm indicate that the PNE solution sets obtained by the PNE-PRTC algorithm are diverse.

5.2 Simulation experiments

A multiple-UAV cooperative confrontation simulation system is constructed based on VR-Forces, which is distributed computer force generation software. In this simulation system, red and blue UAV entities can be constructed, and flight, missile launch and damage can be simulated for these UAVs. Fig. 6 shows typical two-on-two air combat tactical decision scenarios designed in this simulation system. The red and blue roles, involving multiple UAVs, and the white role, with a global perspective, are established. This deployment can be used to simulate air-to-air combat tactical decisions. The multicriteria game approach of ACTDMU (MG-ACTDMU), composed of a multicriteria game model and the PNE-PRTC algorithm, is programmed in the Visual C++ language. The software plugin used to load this approach is developed for embedding in this simulation system. The simulation data in air combat tactical decision scenarios can be obtained by the data logger function in VR-Forces. MySQL workbench software is used to extract simulation information such as the UAV position and state, missile launches and other information by setting time stamp constraints.

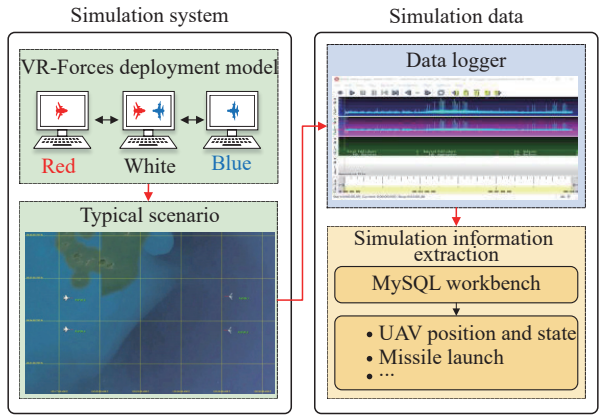


Fig. 6 Deployment mode and process of obtaining simulation information

5.2.1 Simulation scenarios

In a multiple-UAV cooperative confrontation simulation system, three simulation scenarios are designed according to the foe's tactical styles: offensive, neutral, and defensive. The possible tactics chosen for B are shown in Table 4. R is able to choose one of five tactics, including the Head-on Bracket tactic (a), the Turn-away Drag tactic (b), Loose Deuce Engagement tactic (c), Tac-Turn Disengagement tatic (d), and Defensive split tactic (e) [46].

Table 4 Three foe tactical style settings

Number	Foe's tactical style	Opponent tactics set
1	Offensive	a, c, d
2	Neutral	b, d, e
3	Defensive	c, d, e

B, as the baseline opponent, selects the best responses strategies to R's strategy. R selects the strategies based

on the calculation results of the different approaches. The vector payoffs of R are compared in this simulation. We introduce two typical decision-making approaches and compare them with the MG-ACTDMU approach proposed in this paper. The detailed settings of the three approaches are as follows:

(i) MG-ACTDMU: R selects the tactics corresponding to the PNE solutions obtained by MG-ACTDMU.

(ii) MADM: R selects the tactics corresponding to the optimal solutions obtained by MADM.

(iii) Random: R selects tactics randomly based on uniformly distributed sampling.

5.2.2 Evaluation metrics

Three evaluation metrics are employed for the simulation experiments.

Definition 13 Hypervolume of R. The average hypervolume, given in Definition 10, is used to evaluate the quality of the vector payoffs corresponding to the selected tactic. Since only one tactic is selected for each simulation, Definition 10 is simplified to

$$W_{\mathrm{HV}} = \bigcup \mathrm{hv}(n_{\mathrm{PNE}}, p_{\mathrm{PNE}}) \tag{16}$$

where n_{PNE} denotes a PNE solution. p_{PNE} denotes a reference point. hv indicates the hypervolume enclosed by n_{PNE} and p_{PNE}.

Definition 14 Winning percentage of R [47]. After a simulation, if the number of surviving UAVs of R is greater than that of B, then R wins. The winning percentage is calculated by

$$P_w = \frac{n_{\mathrm{win}}}{n_{\mathrm{all}}} \times 100\% \tag{17}$$

where n_{win} represents the number of wins of R and n_{all} is the total number of simulation experiments.

Definition 15 Kill ratio of R [48]. The kill ratio is calculated by

$$P_{\mathrm{CER}} = \frac{L_R}{L_B} \tag{18}$$

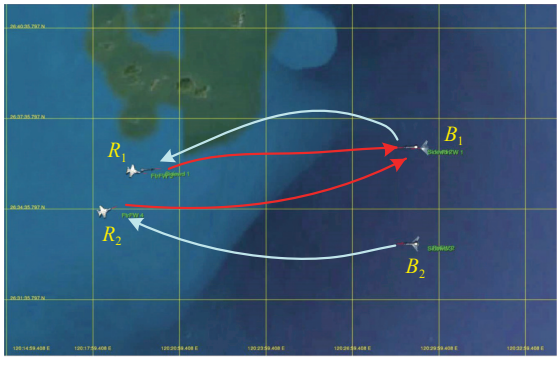
where L_R represents the amount of damage to R's UAVs and L_B represents the amount of damage to B's UAVs in all simulation experiments. Clearly, the smaller the kill ratio is, the greater the advantage of R in the simulation experiments.

5.2.3 Simulation results

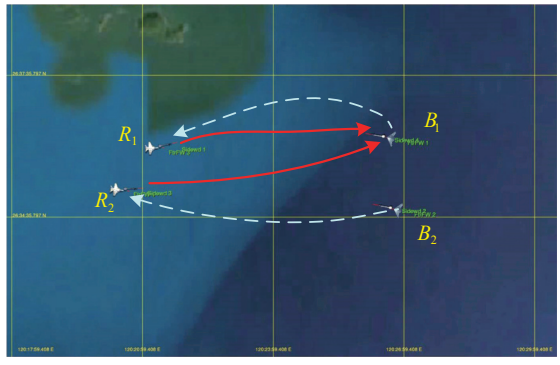
In this section, a case study is first analyzed to illustrate the rationality and stability of the solution obtained by the MG-ACTDMU approach. Then, the effectiveness and efficiency of the approach proposed in this paper is verified by simulation comparison experiments.

(i) Case study

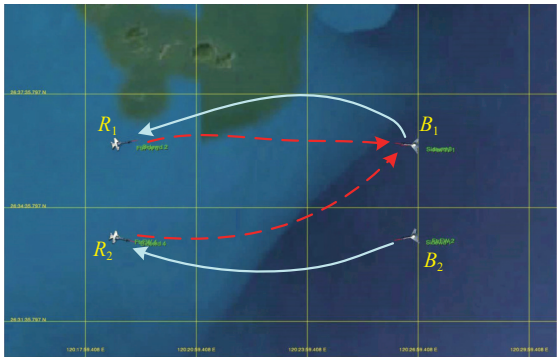
A typical case in which R fights with B, who has an offensive tactical style, is analyzed. According to Table 4, the sets of R's tactics and B's tactics are represented by $R_{\mathrm{tac}} = (r_1, r_2, r_3, r_4, r_5)$ and $B_{\mathrm{tac}} = (b_1, b_3, b_4)$, respectively. On this basis, sets of air-to-air combat tactics incorporating WTA strategies $\bar{R}_{\mathrm{tac}} = (r_{r_1w_1}, \cdots, r_{r_1w_4}, r_{r_2w_1}, \cdots, r_{r_2w_4}, \cdots, r_{r_5w_1}, \cdots, r_{r_5w_4})$ and $\bar{B}_{\mathrm{tac}} = (b_{b_1v_1}, \cdots, b_{b_1v_4}, b_{b_2v_1}, \cdots, b_{b_2v_1}, \cdots, b_{b_4v_1}, \cdots, b_{b_4v_4})$ are constructed. Then, a PNE solution $(r_{r_2w_2}, b_{b_1v_1})$ is calculated according to the MG-ACTDMU approach, indicating that R chooses the $r_{r_2w_2}$ strategy and B chooses the $b_{b_1v_1}$ strategy. The execution process of the tactics corresponding to the PNE solutions is shown in Fig. 7 (a). If R and B choose other strategies alone, they will not obtain better vector payoffs according to the definition of PNE. To illustrate this point, for example, the situation in which R chooses strategy $r_{r_1w_2}$ alone and B chooses strategy $b_{b_4v_1}$ alone is analyzed. Fig. 7 shows the execution of tactics corresponding to PNE solutions and non-PNE solutions.



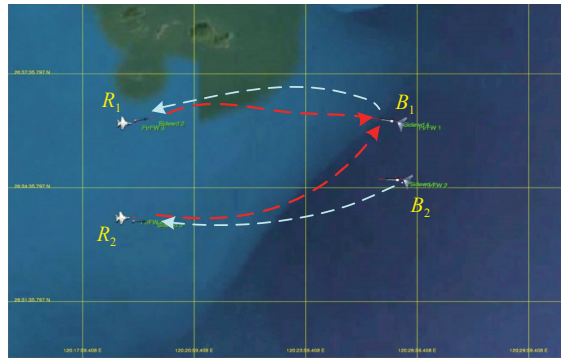
(a) R, who chooses tactic 2, fights with B, who chooses tactic 1



(b) R, who chooses tactic 2, fights with B, who chooses tactic 4



(c) R, who chooses tactic 1, fights with B, who chooses tactic 1



(d) R, who chooses tactic 1, fights with B, who chooses tactic 4

Fig. 7 Executions of tactics corresponding to PNE solutions and non-PNE solutions

Table 5 shows the UAV position and state when R chooses $r_{r_2w_2}$, $r_{r_1w_2}$ and B chooses $b_{b_1v_1}$, $b_{b_4v_1}$. Table 6 indicates the vector payoff matrices calculated by using the MG-ACTDMU approach and the corresponding executions of tactics for the selected strategy profiles. Parameters and vector payoffs associated with PNE are marked in bold.

Table 5 UAV position and state corresponding to the selected tactics

R'tactics	$(x_R/\text{km}, y_R/\text{km}, z_R/\text{km}, v_R/(\text{m/s}), \theta_R/\text{rad}, \phi_R/\text{rad})$	B'tactics	$(x_B/\text{km}, y_B/\text{km}, z_B/\text{km}, v_B/(\text{m/s}), \theta_B/\text{rad}, \phi_B/\text{rad})$
$r_{r_2w_2}$	$p_{r_2}^1$ **=(−94.38, 12.27, 8.341, 314.5, 0.014, 0.181),** $p_{r_2}^2$ **=(−73.44, −11.72, 8.633, 222.1, −0.023, 0.178)**	$b_{b_1v_1}$	$p_{b_1}^1$ **=(103.82, 17.82, 8.24, 275.31, 0.087, 3.181),** $p_{b_1}^2$ **= (102.19, −15.42, 8.57, 255.82, −0.041, 3.078)**
$r_{r_1w_2}$	$p_{r_1}^1$ =(−94.26, 22.53, 8.213, 262.2, 0.005, −0.117), $p_{r_1}^2$ =(−101.92, −21.42, 8.335, 241.4, −0.031, 0.168)	$b_{b_4v_1}$	$p_{b_4}^1$ = (103.82, 15.82, 8.24, 285.47, −0.037, 3.219), $p_{b_4}^2$ = (95.19, −15.42, 8.37, 225.82, −0.052, 3.183)

Table 6 Vector payoff matrices and corresponding executions of tactics for the selected strategy profiles

Tactic	$b_{b_1v_1}$	$b_{b_4v_1}$
$r_{r_2w_2}$	$\boldsymbol{F}_R(r_{r_2w_2}, b_{b_1v_1})$**=(0.498 25,0.409 16,0.410 481,0.285 469, 0.268 969,0.281 608,0.290 261,0.455 265)** $\boldsymbol{F}_B(r_{r_2w_2}, b_{b_1v_1})$**= (0.002 112,0.061 714 6,0.059 379,0.251 632, 0.268 914,0.255 821,0.246 304,0.003 782 54)**	$\boldsymbol{F}_R(r_{r_2w_2}, b_{b_4v_1})$= (0.5,0.491 106,0.485 832,0.288 779, 0.499 388,0.499 952,0.499 92,0.499 755) $\boldsymbol{F}_B(r_{r_2w_2}, b_{b_4v_1})$= (0,6.22E−13,2.18E−08,0.117 967, 0,0,0,0)
$r_{r_1w_2}$	$\boldsymbol{F}_R(r_{r_1w_2}, b_{b_1v_1})$= (0.498 25,0.420 225,0.120 395,0.200 223, 0.268 969,0.282 472,0.227 1,0.473 773) $\boldsymbol{F}_B(r_{r_1w_2}, b_{b_1v_1})$= (0,0.042 818 2,0.376 873,0.234 746, 0.268 914,0.254 891,0.306 51,0.000 073 1)	$\boldsymbol{F}_R(r_{r_1w_2}, b_{b_4v_1})$= (0.5,0.492 209,0.422 766,0.303 366, 0.499 388,0.499 953,0.499 891,0.499 857) $\boldsymbol{F}_B(r_{r_1w_2}, b_{b_4v_1})$= (0,1.16E−14,0.038 753 5,0.230 959, 0,0,0,0)

When B chooses a fixed PNE strategy, the vector payoffs when both R and B choose the PNE strategies are compared with those when R chooses the non-PNE strategy $r_{r_1w_2}$ and B chooses the PNE strategy $b_{b_1v_1}$; that is, $\boldsymbol{F}_R(r_{r_2w_2}, b_{b_1v_1})$ is compared with $\boldsymbol{F}_R(r_{r_1w_2}, b_{b_1v_1})$. For simplicity, $\boldsymbol{F}_R(o_{\tau_{22}\tau_{11}}) = \boldsymbol{F}_R(r_{r_2w_2}, b_{b_1v_1})$ and $\boldsymbol{F}_R(o_{\tau_{12}\tau_{11}}) = \boldsymbol{F}_R(r_{r_1w_2}, b_{b_1v_1})$. Partial binary relations are constructed by using the MG-ACTDMU approach, and they are expressed as $m_P^R(o_{\tau_{12}\tau_{11}}, o_{\tau_{22}\tau_{11}}) = 3$, $m_I^R(o_{\tau_{12}\tau_{11}}, o_{\tau_{22}\tau_{11}}) = 2$, $m_Q^R(o_{\tau_{22}\tau_{11}}, o_{\tau_{12}\tau_{11}}) = 2$ and $m_P^R(o_{\tau_{22}\tau_{11}}, o_{\tau_{12}\tau_{11}}) = 1$. The outranking relation $o_{\tau_{22}\tau_{11}} S_V^R o_{\tau_{12}\tau_{11}}$ is obtained by Definition 4 proposed in this paper. This indicates that the PNE strategy $r_{r_2w_2}$ is better than the non-PNE strategy $r_{r_1w_2}$ for R. However, the average hypervolumes W_{HV} of R corresponding to strategy $r_{r_2w_2}$ and strategy $r_{r_1w_2}$ are 2.391e−4 and 4.126e−5, respectively. This also shows that strategy $r_{r_2w_2}$ is better than strategy $r_{r_1w_2}$.

When R chooses a fixed PNE strategy, the vector payoffs when both R and B choose the PNE strategies are compared with those when R chooses the PNE strategy $r_{r_2w_2}$ and B chooses the non-PNE strategy $b_{b_4v_1}$; that is, $\boldsymbol{F}_B(o_{\tau_{22}\tau_{11}})$ is compared with $\boldsymbol{F}_B(o_{\tau_{22}\tau_{41}})$. The partial binary relations constructed by using the MG-ACTDMU approach are expressed as $m_P^B(o_{\tau_{22}\tau_{11}}, o_{\tau_{22}\tau_{41}}) = 6$ and

$m_Q^B(o_{\tau_{22}\tau_{11}}, o_{\tau_{22}\tau_{41}}) = 2$. The outranking relation $o_{\tau_{22}\tau_{11}} S_C^B o_{\tau_{22}\tau_{41}}$ is obtained by Definition 2. This means that the PNE strategy $b_{b_1v_1}$ is better than the non-PNE strategy $b_{b_4v_1}$ for B. However, the average hypervolumes W_{HV} of B corresponding to strategy $b_{b_1v_1}$ and strategy $b_{b_4v_1}$ are 1.248e−10 and 0, respectively. This also confirms that strategy $b_{b_1v_1}$ is better than strategy $b_{b_4v_1}$.

When both R and B choose non-PNE strategies, the vector payoffs when both R and B choose the PNE strategies are compared with those when R chooses the non-PNE strategy $r_{r_1w_2}$ and B chooses the non-PNE strategy $b_{b_4v_1}$; that is, $\boldsymbol{F}_R(o_{\tau_{22}\tau_{11}})$ is compared with $\boldsymbol{F}_R(o_{\tau_{12}\tau_{41}})$, and $\boldsymbol{F}_B(o_{\tau_{22}\tau_{11}})$ is compared with $\boldsymbol{F}_B(o_{\tau_{12}\tau_{41}})$. The outranking relations $o_{\tau_{12}\tau_{41}} S_Q^R o_{\tau_{22}\tau_{11}}$ and $o_{\tau_{22}\tau_{11}} S_Q^B o_{\tau_{12}\tau_{41}}$ can be obtained by using the MG-ACTDMU approach. They indicate that $r_{r_2w_2}$ is better than $r_{r_1w_2}$ for R and that $b_{b_1v_1}$ is better than $b_{b_4v_1}$ for B. On the other hand, for the PNE solution $o_{\tau_{22}\tau_{11}}$, the W_{HV} of R is 1.969e−3 and the W_{HV} of B is 0. Although R's W_{HV} for $o_{\tau_{12}\tau_{41}}$ is greater than that for $o_{\tau_{22}\tau_{11}}$, B will choose $b_{b_4v_1}$ for higher payoffs instead of $b_{b_1v_1}$ when R chooses $r_{r_1w_2}$. In summary, the pair of strategy profiles $o_{\tau_{12}\tau_{41}}$ is obviously unstable, while the PNE strategy is stable. The analysis of the other pairs of strategy profiles can also be done in this way.

(ii) Simulation results

Three scenarios are designed according to B's three tactical styles. In each scenario, the MG-ACTDMU approach, MADM approach and random approach are used to solve the ACTDMU problem for R, and 100 simulation experiments are carried out for each approach. The three evaluation metrics in Subsection 5.2.2 are calculated based on the simulation information obtained from the data logger function in VR-Forces. The three evaluation metrics for the MG-ACTDMU, MADM and random approaches are shown in Table 7 where the results of our MG-ACTDMU are marked in bold.

Table 7 Three evaluation metrics for the MG-ACTDMU, MADM and random approaches

Tactics	Approaches	Offensive	Neutral	Defensive
W_{HV}	MG-ACTDMU	**0.01721**	**0.01836**	**0.01912**
	MADM	0.0154	0.01742	0.01787
	Random	0.00669	0.00777	0.00725
P_w	MG-ACTDMU	**0.36**	**0.42**	**0.38**
	MADM	0.22	0.19	0.27
	Random	0.11	0.13	0.07
P_{CER}	MG-ACTDMU	**0.8461**	**0.9469**	**0.8983**
	MADM	1.2978	1.3804	1.1800
	Random	1.6296	1.7625	1.8243

The simulation results show that the MG-ACTDMU approach proposed in this paper is obviously better than the MADM approach and random approach for the above three scenarios.

Compared with the MADM approach, the hypervolume W_{HV} of the MG-ACTDMU approach is greater, the winning percentage P_w is increased by 70.6% on average and the kill ratio P_{CER} is decreased by 30.2% on average. Compared with the random approach, the W_{HV} of the MG-ACTDMU approach is significantly greater, P_w is increased by 274.2% on average and P_{CER} is decreased by 48.0% on average. The simulation analysis shows that the MG-ACTDMU approach can effectively improve the winning percentage and reduce the kill ratio of UAVs. The reason is that the optimal strategy for R is selected by the MG-ACTDMU approach when the most favorable situation for B is reached. However, the MADM approach and random approach only make decisions from their own perspective, and ignoring the impact of the opponent's decision on our decision leads to a failure to achieve the desired results.

6. Conclusions

In the process of beyond-visual-range combat, ACTDMU is a key step. This study proposes a MG-ACTDMU. Our approach includes a multicriteria game model and a PNE-PRTC. For the model, sets of air-to-air combat tactics incorporating WTA strategies are constructed. In addition, the vector payoff functions of the multicriteria game are designed based on four predominance factors. For the algorithm, weak partial binary relation constraints and negative threshold constraints are added to construct a five-level domination criterion. Then, the preference relations are redefined. It is proven that the solutions of the PNE-PRTC algorithm are refinements of PNE solutions.

A leading game generator, GAMUT, is used to generate three datasets with 45 sizes of instances. These datasets can provide a benchmark for research in the field of multicriteria games. The numerical experiments show that the PNE-PRTC algorithm has a high speed, and its performance is better in terms of the average spacing metric and average diversity measure than the baseline algorithms. In particular, the PNE-PRTC algorithm has more advantages for solving large-scale instances. This algorithm can be applied not only to the ACTDMU problem in this paper but also to general multicriteria games.

A multiple-UAV cooperative air combat simulation system is constructed based on VR-Forces software. This simulation system lays the foundation for testing the effect of this approach and the development of future works. The simulation experiments show that the MG-ACTDMU approach is more effective than the multiple-attribute decision-making approach and the randomly chosen decision approach for solving the ACTDMU

problem. A case study is used to analyze the stability and effectiveness of the solutions obtained by this approach.

In future work, uncertainty regarding the capability of foe UAVs should be taken into account in the ACTDMU problem. In addition, improving the robustness of the multicriteria game algorithm and refining the PNE solution sets are important research directions.

References

[1] SUN Z X, PIAO H Y, YANG Z, et al. Multi-agent hierarchical policy gradient for air combat tactics emergence via self-play. Engineering Applications of Artificial Intelligence, 2022, 98: 104112.

[2] XU X M, YANG R N, FU Y. Situation assessment for air combat based on novel semi-supervised naive Bayes. Journal of Systems Engineering and Electronics, 2018, 29(4): 768–779.

[3] JIANG Y Y, GAO Y, SONG W Q, et al. Bibliometric analysis of UAV swarms. Journal of Systems Engineering and Electronics, 2022, 33(2): 406–425.

[4] JEREMY J, FREDERICK L, DENNIS E. Deployment and flight operations of a large scale UAS combat swarm: results from DARPA service academies swarm challenge. Proc. of the International Conference on Unmanned Aircraft Systems, 2018: 12−15.

[5] JOAO P, MARCOS R, ANDRE N, et al. Machine learning to improve situational awareness in beyond visual range air combat. IEEE Latin America Transactions, 2022, 20(8): 2039–2045.

[6] YANG Z, SUN Z X, PIAO H Y, et al. Online hierarchical recognition method for target tactical intention in beyond-visual-range air combat. Defence Technology, 2022, 18: 1349–1361.

[7] SHIN H, LEE J, KIM H, et al. An autonomous aerial combat framework for two-on-two engagements based on basic fighter maneuvers. Aerospace Science and Technology, 2018, 72: 305–315.

[8] ISHWARYA M S, ASWANI K C. Quantum-inspired ensemble approach to multi-attributed and multi-agent decision-making. Applied Soft Computing, 2021, 106: 107283.

[9] ZHANG Q H, YANG C C, WANG G Y. A sequential three-way decision model with intuitionistic fuzzy numbers. IEEE Trans. on Systems, Man, and Cybernetics: System, 2021, 51(5): 2640–2652.

[10] SELVACHANDRAN G, QUEK S, HAN L, et al. A new design of mamdani complex fuzzy inference system for multiattribute decision making problems. IEEE Trans. on Fuzzy Systems, 2021, 29(4): 716–730.

[11] ALKAHER D, MOSHAIOV A. Dynamic-escape-zone to avoid energy-bleeding coasting missile. Journal of Guidance, Control, and Dynamics, 2015, 38(10): 1908–1921.

[12] LI S Y, CHEN M, WU Q X, et al. Threat sequencing of multiple UCAVs with incomplete information based on game theory. Journal of Systems Engineering and Electronics, 2022, 33(4): 986–996.

[13] ZHE Y. A coalitional extension of generalized fuzzy games. Fuzzy Sets and Systems, 2020, 383: 68–79.

[14] BACKWELL D. An analog of the minimax theorem for vector payoffs. Pacific Journal of Mathematics, 1956, 6(1): 1–8.

[15] RETTIEVA A. Dynamic multicriteria games with asymmetric players. Journal of Global Optimization, 2022, 83(3): 521–537.

[16] CLEMENTE M, FERNANDEZ F, PUERTO J. Pareto-optimal security strategies in matrix games with fuzzy payoffs. Fuzzy Sets and Systems, 2011, 176(1): 36–45.

[17] FERNANDEZ F, MONROY L, PUERTO J. Multicriteria goal games. Journal of Optimization Theory and Applications, 1998, 99(2): 403–421.

[18] RADJEF M S, FAHEM K. A note on ideal Nash equilibrium in multicriteria games. Applied Mathematics Letters, 2008, 21(11): 1105–1111.

[19] ZAPATA A, MARMOL A, MONROY L, et al. A maxmin approach for the equilibria of vector-valued games. Group Decision and Negotiation, 2019, 28(2): 415–432.

[20] QUANT M, BORM P, FIESTRAS-JANEIRO G. On properness and protectiveness in two-person multicriteria games. Journal of Optimization Theory and Applications, 2009, 140(3): 499–512.

[21] FAHEM K, RADJEF M S. Properly efficient Nash equilibrium in multicriteria noncooperative games. Mathematical Methods of Operations Research, 2015, 82(2): 175–193.

[22] ZELENY M. Games with multiple payoffs. International Journal of Game Theory, 1975, 4(4): 179–191.

[23] MARCO G, MORGAN J. A refinement concept for equilibria in multicriteria games via stable scalarizations. International Game Theory Review, 2007, 9(2): 169–181.

[24] NOVIKOVA N M, POSPELOVA I I. Scalarization method in multicriteria games. Computational Mathematics and Mathematical Physics, 2018, 58(2): 180–189.

[25] COOK W D. Zero-sum games with multiple goals. Naval Research Logistics, 1976, 23(4): 615–621.

[26] NISHIZAKI I, NOTSU T. Nondominated equilibrium solutions of a multiobjective two-person nonzero-sum game and corresponding mathematical programming problem. Journal of Optimization Theory and Applications, 2007, 135(2): 217–239.

[27] HONG J M, KIM M H, LEE G M. On linear vector program and vector matrix game equivalence. Optimization Letters, 2012, 6(2): 231–240.

[28] CHANDRA S, AGGARWAL A. On solving matrix games with pay-offs of triangular fuzzy numbers: certain observations and generalizations. European Journal of Operational Research, 2015, 246(2): 575–581.

[29] PENG H G, WANG X K, WANG T L, et al. Multi-criteria game model based on the pairwise comparisons of strategies with Z-numbers. Applied Soft Computing, 2019, 74: 451–465.

[30] ZYCHOWSKI A, GUPTA A, MANDZIUK J, et al. Addressing expensive multi-objective games with postponed preference articulation via memetic co-evolution. Knowledge-Based Systems, 2018, 154: 17–31.

[31] HAREL M, MOSHAIOV A, ALKAHER D. Rationalizable strategies for the navigator–target–missile game. Journal of Guidance, Control, and Dynamics, 2020, 43(6): 1129–1142.

[32] YOUSFI-HALIMI N, RADJEF M S, SLIMANI H. Refinement of pure Pareto Nash equilibria in finite multicriteria games using preference relations. Annals of Operations Research, 2018, 267(1): 607–628.

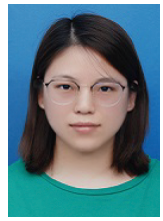
[33] HAMEL A H, LOHNE A. A set optimization approach to zero-sum matrix games with multi-dimensional payoffs. Mathematical Methods of Operations Research, 2018, 88(3): 369–397.

[34] MA Y Y, WANG G Q, HU X X, et al. Two-stage hybrid heuristic search algorithm for novel weapon target assignment problems. Computers & Industrial Engineering, 2021, 162: 107717.
[35] YAO Z X, MING L, CHEN Z J, et al. Mission decision-making method of multi-aircraft cooperatively attacking multi-target based on game theoretic framework. Chinese Journal of Aeronautics, 2016, 29(6): 1685–1694.
[36] ZHAO K Y, LI L, CHEN Z Q, et al. A survey: optimization and applications of evidence fusion algorithm based on Dempster-Shafer theory. Applied Soft Computing, 2022, 124: 109075.
[37] SU M C, LAI S C, LIN S C, et al. A new approach to multi-aircraft air combat assignments. Swarm and Evolutionary Computation, 2012, 6: 39–46.
[38] ZHOU L P, WAN S P, DONG J Y. A fermatean fuzzy ELECTRE method for multi-criteria group decision-making. Informatica, 2022, 33(1): 181–224.
[39] JOSE R F, SALVATORE G, BERNARD R. Electre-Score: a first outranking based method for scoring actions. European Journal of Operational Research, 2022, 297: 986–1005.
[40] LOTFI A, DJAAFAR Z, KAMAL A, et al. A novel epsilon-dominance Harris Hawks optimizer for multi-objective optimization in engineering design problems. Neural Computing and Applications, 2022, 34: 17007–17036.
[41] SANDHOLM T, GILPIN A, CONITZER V. Mixed-integer programming methods for finding Nash equilibria. Proc. of the 20th National Conference on Artificial Intelligence, 2005, 2: 495–501.
[42] SHANG K. A survey on the hypervolume indicator in evolutionary multiobjective optimization. IEEE Trans. on Evolutionary Computation, 2021, 25(1): 1–20.
[43] REDDY M J, KUMAR D N. Multiobjective differential evolution with application to reservoir system optimization. Journal of Computing in Civil Engineering, 2007, 21(2): 136–146.
[44] ZITZLER E, DEB K, THIELE L. Comparison of multiobjective evolutionary algorithms: empirical results. Evolutionary Computation, 2000, 8(2): 173–195.
[45] WU J H, ZHANG D Q, JIANG C, et al. On reliability analysis method through rotational sparse grid nodes. Mechanical Systems and Signal Processing, 2021, 147: 107106.
[46] SHAW R L. Fighter combat: tactics and maneuvering. Maryland: The United States Naval Institute Annapolis, 1985.
[47] HAN Y T, ZHOU Q B, DUAN F Q. A game strategy model in the digital curling system based on NFSP. Complex & Intelligent Systems, 2021, 8: 1857–1863.
[48] HU X W, LUO P C, ZHANG X N, et al. Improved ant colony optimization for weapon-target assignment. Mathematical Problems in Engineering, 2018, 2018: 6481635.

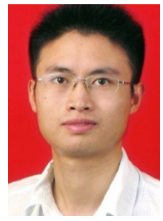
Biographies

JIANG Ruhao was born in 1993. He received his B.S. and M.S. degrees in mechanical engineering from Hefei University of Technology, in 2016 and 2019, respectively. He is currently pursuing his Ph.D. degree with the School of Management, HeFei University of Technology. His research interests include multi-aircrafts cooperative decision-making and game theory.
E-mail: jiangrh@mail.hfut.edu.cn

LUO He was born in 1982. He received his B.S. and Ph.D. degrees from Hefei University of Technology, in 2004 and 2009, respectively. He is currently a professor in HeFei University of Technology. His research interests include intelligent decision-making, multi-agent system, and the applications of un-manned aerial vehicle.
E-mail: luohe@hfut.edu.cn



MA Yingying was born in 1994. She received her B.S. and Ph.D. degrees from Hefei University of Technology, in 2016 and 2022, respectively. She is currently a post-doctoral in HeFei University of Technology. Her research interests include multi-UAVs cooperative target assignment in air combat and game theory.
E-mail: mayingying@mail.hfut.edu.cn



WANG Guoqiang was born in 1982. He received his B.S. and M.S. degrees from University of Science and Technology of China, in 2004 and 2007, respectively, and Ph.D. degree from Hefei University of Technology, in 2016. He is currently an associate professor in HeFei University of Technology. His research interests include management and intelligent decision-making of un-manned aerial vehicle formation.
E-mail: gqwang2017@hfut.edu.cn