

Cooperative mission planning based on game theory for UAVs and USVs heterogeneous system in dynamic scenario

Hong Long and Haibin Duan

School of Automation Science and Electrical Engineering, Beihang University (BUAA), Beijing, China

Abstract

Purpose – The purpose of this paper is to present and implement a task allocation method based on game theory for reconnaissance mission planning of UAVs and USVs system.

Design/methodology/approach – In this paper, the decision-making framework via game theory of mission planning is constructed. The mission planning of UAVs–USVs is transformed into a potential game optimization problem by introducing a minimum weight vertex cover model. The modified population-based game-theoretic optimizer (MPGTO) is used to improve the efficiency of solving this complex multi-constraint assignment problem.

Findings – Several simulations are carried out to exhibit that the proposed algorithm obtains the superiority on quality and efficiency of mission planning solutions to some existing approaches.

Research limitations/implications – Several simulations are carried out to exhibit that the proposed algorithm obtains the superiority on quality and efficiency of mission planning solutions to some existing approaches.

Practical implications – The proposed framework and algorithm are expected to be applied to complex real scenarios with uncertain targets and heterogeneity.

Originality/value – The decision framework via game theory is proposed for the mission planning problem of UAVs–USVs and a MPGTO with swarm evolution, and the adaptive iteration mechanism is presented for ensuring the efficiency and quality of the solution.

Keywords Heterogeneous system, Mission planning, Modified population-based game-theoretic optimizer, Minimum weight vertex cover, Game theory

Paper type Research paper

Introduction

An unmanned system (US) will be beneficial in many aspects, especially when the system is composed of multiple entities distributed in capabilities or space (Yoon *et al.*, 2020). Cooperation enables the UAVs–USVs to work as a team and complete activities that they would be tough to accomplish individually (Yuan *et al.*, 2022; Wu *et al.*, 2021). To achieve collaborative cooperation, task allocation is deemed indispensable, as it allows for the adoption of the optimal task sequencing and the maximization of collective benefits within USs (Sheng *et al.*, 2022). This necessary behavior ensures that tasks are executed in the most favorable order and that the overall performance of the system is maximized. In such missions, with the increased complexity of missions and multiple expansion of UAVs–USVs, the US requires to achieve flexible and autonomous deployment optimization (Wu *et al.*, 2018; Gao and Bhattacharya, 2019). Many classical solutions require an unaltered communication interaction between the nuclear unit and other UAVs–USVs, which is directly corresponding to the computational cost (Nemer *et al.*, 2020). Hence, that is essential to supply

to UAVs–USVs more autonomy in decision-making for fulfilling dynamic control and management.

Compared with other traditional methods, game theory consists of series of rational strategies in which each player interacts with the other players in a competitive and cooperative manner, simplifying the problems of search, task allocation and path planning, making it easier to understand and analyze (Kusyk *et al.*, 2021; Semasinghe *et al.*, 2017). More importantly, under the Nash equilibrium theory and specific learning rules, satisfactory collective strategies or even optimal solutions from local interactions can be obtained (Sun *et al.*, 2019). According to the variations of scenario and resources, UAVs–USVs expect to complete all multi-robot tasks with the least time, which is involving with internal competition and cooperation. Therefore, the application of game theory in mission planning has received considerable attention.

In Mebrek and Yassine (2021), a decentralized task allocation algorithm that can learn and self-update according to the actions of other players is applied to the optimal allocation of energy consumption missions in internet of things (IoT). Although the

The current issue and full text archive of this journal is available on Emerald Insight at: <https://www.emerald.com/insight/1748-8842.htm>



Aircraft Engineering and Aerospace Technology
96/9 (2024) 1128–1138
© Emerald Publishing Limited [ISSN 1748-8842]
[DOI 10.1108/AEAT-02-2023-0057]

This work was partially supported by Science and Technology Innovation 2030-Key Project of “New Generation Artificial Intelligence” under grant #2018AAA0102403, National Natural Science Foundation of China under grant #U1913602, #U20B2071 and #U19B2033.

Received 5 March 2023
Revised 16 August 2023
22 December 2023
Accepted 29 January 2024

game-theoretic approach exhibits promising potential for intelligent resource allocation in IoT application, its computational complexity necessitates considerations regarding system performance. Sun *et al.* (2021) addressed self-organized task allocation in multi-satellite system by formulating it as a potential game, and each Nash equilibrium of the game constitutes a task cover that maximizes the number of executed tasks is proved. The multi-objective optimization of task allocation is facilitated by the game-theoretic framework used in the approach, leading to the improvement of overall system efficiency. However, the feasibility of information requirements and exchanges needs to be carefully considered during the design process. In Wu and Shang (2020), a novel multiagent dynamic task allocation method based on potential game is used to effectively solve each agent constraints of the selectable action set in the communication range. In this approach, the task allocation strategies of the intelligent agents are continuously improved through learning and optimization during the game process, with the goal of maximizing the overall system performance. Specifically, it should be noted that the game process may be characterized by slow convergence rates and the potential for suboptimal solutions to be reached. In Hassija *et al.* (2020), the cost optimization model of energy trading between UAV and charging station is established by using game theory. Through strategy selection and game equilibrium, effective allocation within multi-drone networks can be achieved by this method. Nevertheless, a careful consideration should be given to the computational complexity and real-time feasibility of this approach in practical systems, given the involvement of large-scale game and decision-making processes. Gao and Bhattacharya (2019) used an algorithm to obtain the pure strategy Nash equilibrium of noncooperative game and showed its superior performance of assigning multiple robots to charge stations. In contrast, the game process faces the challenge of slow convergence rates, particularly in complex and large-scale systems. This can lead to prolonged optimization and improvement processes for task allocation strategies, ultimately impacting the system's performance.

In this paper, to introduce game theory into the real mission planning scenarios, the problem is transformed into a mathematical model of minimum weight vertex cover (MWVC). As the important theory of decision-making, the MWVC has been studied and concerned in many practical applications (Chen and Li, 2023; Sun *et al.*, 2019; Safar *et al.*, 2007). Thus, the mission planning problem is considered to transform into the MWVC problem by treating all mission schemes as autonomous players with theoretical proof of potential game theory.

Population-based game-theoretic optimizer (PGTO) is a novel optimal algorithm, which possesses the superiority of some existing methodologies for resolving the MWVC game problem, was proposed by Qiu *et al.* (2022). However, the PGTO algorithm's population-based nature and incorporation of game theory may introduce additional computational complexity. This may result in increased time and resource requirements, particularly for large-scale instances of the problem. Pigeon-inspired optimization (PIO) is inspired by the homing of a flock of pigeons, its procedures mainly include the map, compass and landmark operators, which includes the position and velocity information, and there is an exponential

convergence (Yuan and Duan, 2022; Duan and Qiao, 2014). Due to the simply theoretical principle, fast convergence and high search efficiency, PIO has been widely applied (Tong *et al.*, 2022; Duan *et al.*, 2021; Alazzam *et al.*, 2020). However, PIO primarily operates on continuous functions, necessitating its discretization for task allocation purposes. Thus, we consider the evolutionary strategy of discrete PIO was introduced into the PGTO to construct the discrete map and compass swarm evolution (DMC-SE) strategy, then a novel modified population-based game-theoretic optimizer (MPGTO), which can complete missions in the hope of a shorter time and ensure optimization effect. In addition, a decision-making framework for the least time to complete all tasks in dynamic scenario is designed by mission planning stage and status update stage. Moreover, information entropy evaluation and task completion rate indicators to evaluate the orderliness and quality of tasks completed is also founded. Finally, the feasibility and superiority of decision-making framework are demonstrated through several simulation experiments.

The remainder of this paper is organized as follows. In the section "Problem description and model formulation", the model of UAV, USV and target are built and the mission planning problem is transformed into a MWVC problem. The MPGTO is proposed in the section "Algorithm design and decision-making framework", a decision-making framework that can handle dynamic reconnaissance process is built and the information entropy evaluation indicator is adopted. Four situations and a series of detailed simulation analyses are used in the section "Simulation" to demonstrate the superiority of the proposed methodology. Finally, some conclusions and future plans are presented in the section "Conclusions and future plans".

Problem description and model formulation

Problem formulation

For illustrating the formulation of UAVs–USVs heterogeneous system, we assume that n UAVs–USVs $U = \{U_1, U_2, \dots, U_n\}$ and m dynamic targets $T = \{T_1, T_2, \dots, T_m\}$. The UAVs–USVs include n_f faster UAVs, n_r normal UAVs and n_s slowest USVs that are equipped with different detection capabilities (Stolfi *et al.*, 2021). In our simulation, the detection speed and range of USs will be based on the scenarios presented herein. This consideration will serve as a reference for determining the parameters of the USs in our study.

The fuel quantity $F = \{F_1, F_2, \dots, F_n\}$ of UAVs–USVs decrease ΔF with unit time until minimum value F_{\min} . A reconnaissance operation is typically considered to be in close proximity to the target at a safe distance range using radar or optical sensors to track the target for a period of time t_b (Suresh and Ghose, 2012). For the objects of reconnaissance, they could contain known targets T^k , unknown targets T^u and deceptive targets T^d (Berger *et al.*, 2007).

Considering the design of complex UAVs–USVs waypoints at top-level control, the time-stamped model is used to generate trajectory on account of sailing at an invariant altitude and velocity, the i -th UAV–USV as follows (Zhang and Duan, 2018):

$$\begin{cases} \mathbf{P}_i(t) = \mathbf{P}_i(t-1) + \mathbf{v}_i(t) \\ \mathbf{v}_i(t) = \begin{bmatrix} \|\mathbf{v}_i(t)\| \cos(\psi_{i,t}) \\ \|\mathbf{v}_i(t)\| \sin(\psi_{i,t}) \end{bmatrix} \\ \psi_i(t) = \arctan \frac{E_i^y(t) - S_i^y(t)}{E_i^x(t) - S_i^x(t)} \end{cases} \quad (1)$$

In the time-stamp model, two components namely the velocity component $\mathbf{v}_i(t) = [v_i^x(t), v_i^y(t)]^T$ and position component $\mathbf{P}_i(t) = [P_i^x(t), P_i^y(t)]^T$ are designated. $\psi_i(t) = \arctan \frac{E_i^y(t) - S_i^y(t)}{E_i^x(t) - S_i^x(t)}$ denotes the angle between the line connecting the current point $\mathbf{S}_i(t)$ to the target point $\mathbf{E}_i(t)$ and the x -axis of the Cartesian coordinate system. Where $\mathbf{E}_i(t) = [E_i^x(t), E_i^y(t)]^T$ denotes the position of the target and $\mathbf{S}_i(t) = [S_i^x(t), S_i^y(t)]^T$ denotes the mission switching position of UAVs-USVs. In Figure 1, upon determining the location $\mathbf{E}_i(t)$, the schematic representation suggests the existence of a defined region of threat. UAVs-USVs commencing from their current positions and converging towards the target, engage in a reconnaissance maneuver within a time interval denoted as t_b , upon reaching the vicinity of the threat region. Upon the completion of the designated reconnaissance time, the UAVs-USVs promptly transitions to a new reconnaissance phase, commencing from its mission switching point $\mathbf{S}_i(t + t_b)$ and moving toward the subsequent target location $\mathbf{E}_i(t + t_b)$.

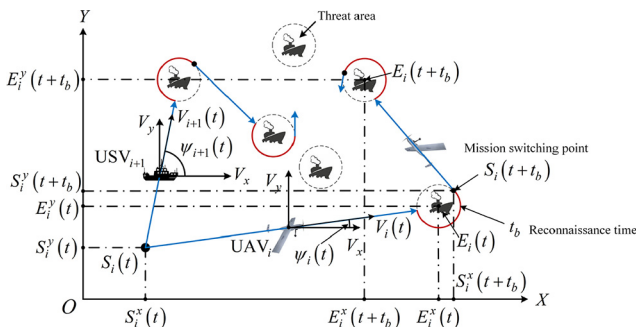
If the Euclidean distance D_{ij} between i -th UAV-USV and j -th target is equal to the reconnaissance range R_d , this target will be detected and its authenticity can be identified. For the moving vessel targets on the sea, two-dimensional plane motion of targets without altitude variation, such as the j -th target model is designed as:

$$\begin{cases} \dot{x}_j = r_1 \cdot u \\ \dot{y}_j = r_2 \cdot v \end{cases} \quad (2)$$

where $r_1, r_2 \in (-1, 1)$, u and v denote the constant rate of vessel velocity. We assume that the motion of the target vessel is a fully unpredictable random movement occurring in two dimensions, within a certain range of velocities. The velocity magnitudes along the two coordinate axes are determined by multiplying the constant values u and v with their respective corresponding random numbers, denoted as r_1 and r_2 .

To ensure safety during a joint reconnaissance operation, UAVs-USVs can only surround and follow the ships within a

Figure 1 Diagram of time-stamp model



Source: Author's own creation

specific range R_L and UAVs-USVs must have an angular distribution θ_{ik} between each other, which are formulated as:

$$\begin{cases} D_{ij} = R_L, \forall i \in U; \forall j \in T \\ \theta_k = 2\pi / \sum_{k=1}^n \sum_{j=1}^m X_{(k,j)}, \forall k \in U \\ X_{(k,j)} \in \{0, 1\} \end{cases} \quad (3)$$

where $X_{(k,j)} = 1$ means that the k -th UAV-USV to execute j -th target, whereas $X_{(k,j)} = 0$ otherwise.

In general, radar sensors are more capable of detection than optical sensors, the number n of payloads required and the duration t_b of the reconnaissance will need to change proportionally. For example, when performing a mission, it requires N_R radar payloads to complete t_R units of time or N_y optical payloads to complete t_y units of time.

For traditional constraints of multiple traveling salesman problem (Ghassemi and Chowdhury, 2022), it should ensure that every target must be executed once via n UAVs-USVs, expressed by:

$$\begin{cases} \sum_{k=1}^n \sum_{j=1}^m X_{(k,j)} = N, \forall k \in U; \forall j \in T \\ N \in \{N_R, N_y\} \\ X_{(k,j)} \in \{0, 1\} \end{cases} \quad (4)$$

where N_R and N_y , respectively, denote the number of UAVs-

USVs for single target within different sensors. $\sum_{k=1}^n \sum_{j=1}^m X_{(k,j)} = N$

denotes the number of UAVs-USVs deployed to execute the j -th target.

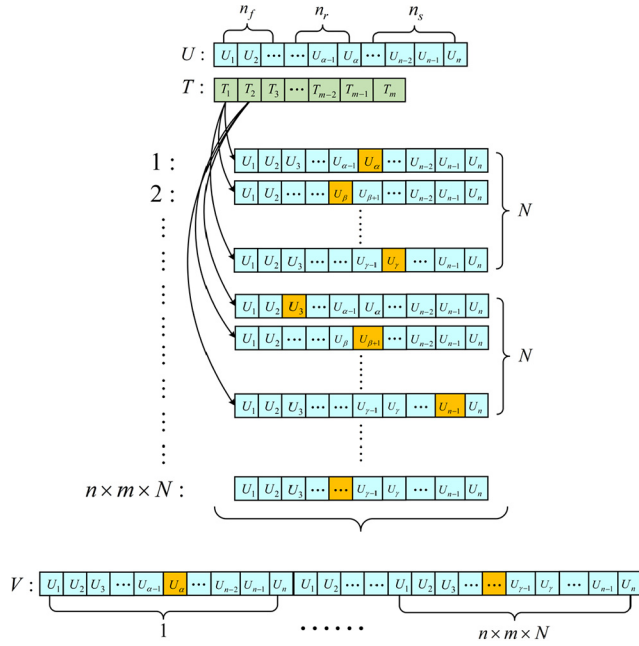
A minimum weight vertex cover model for mission planning problem

Considering an undirected weighted graph $G = \{V, E, W\}$, the set of edges $E = \{e_{ij} | i, j \in V\}$. A subset $S \in V$ is called a vertex cover (VC), whereas it satisfies coverage requirement, such as any edge $(i, j) \in E$ at least one vertex in S , the problem means that to find out an optimal S^* with the minimum sum of vertex weights among all VCs.

An adjacency matrix G is designed to address contradictions from equation (4) through the game converter (c). The vertex matrix V of the adjacency matrix in the GC is shown in Figure 2, it is calculated by $m \times n \times n$ as any UAVs-USVs with mission-related payloads can carry out any targets. We consider a time cost $C(i, j)$ of the i -th UAV-USV from current position to the preassigned position of the j -th target, given by:

$$C(i, j) = \frac{D_{ij}}{\|\mathbf{v}_i\|} \quad (5)$$

where \mathbf{v}_i denotes the i -th UAV-USV speed and the D_{ij} denotes the Euclidean distance between i -th UAV-USV and j -th target. The GC can be elaborated as two main procedures. At first, the executions of tasks are refined to schemes that single UAV-USV carries a payload. Based on this, a $m \times n \times N$ times loop formed, where payload types affect the carried number τ of

Figure 2 Mission planning schemes transform into vertex matrix

Source: Author's own creation

payloads. In Figure 2, the set U of UAVs-USVs, it is sorted by faster UAVs, normal UAVs and USVs from left to right.

Then, a null $m \times n \times n$ matrix G is constructed and restricted with three constraints in once mission planning, as follows:

- 1 Constraint (1) ensures each UAV-USV can only complete single task of a target. If the UAV-USV is required and identical for the j and k schemes, $G_{(j \times k) \& (k \times j)} \leftarrow 1$.
- 2 Constraint (2) ensures each task of target can only be done by single UAV-USV. If the target, the serial number of schemes and payload are identical for the j and k schemes, similarly, $G_{(j \times k) \& (k \times j)} \leftarrow 1$.
- 3 Constraint (3) ensures that there are no contradictions when different mission schemes for one target. If the targets of the j and k schemes are identical, correspondingly, $G_{(j \times k) \& (k \times j)} \leftarrow 1$.

When one of the vertices in the union is selected, the remaining vertices also are covered, then vertices with the same function cannot be selected repeatedly, therefore these conflicts can be addressed. In addition, the normalized time cost function is set as vertex weight value W , which is shown in equation (6):

$$W(j) \leftarrow C(\alpha(j), \beta(j)) / \sum_{j=1}^{m \times n \times N} C(\alpha(j), \beta(j)) \quad (6)$$

where the serial number of UAVs-USVs and targets can be recorded by $\alpha(j)$ and $\beta(j)$ in the j -th scheme, respectively. In other words, following the establishment of the game matrix, the vertex j signifies a strategy involving the deployment of UAVs-USVs toward the target. Here, the identification of the deployed UAV-USV is denoted by the symbol $\alpha(j)$, whereas the target is identified by the symbol $\beta(j)$.

For finding out the optimal set $S^* \in V$, the MWVC problem is treated as an unconstrained maximization problem as follows (Sun et al., 2019):

$$\begin{cases} \arg \max_{s \in S} f(s) \\ f(s) = - \left\{ \sum_{i=1}^n W(i) s_i + \lambda \sum_{i=1}^n \left[(1 - s_i) \sum_{j \in \Omega_i} (1 - s_j) \right] \right\} \\ S = \{s_1, s_2, \dots, s_n\} \\ s_i \in \{0, 1\} \end{cases} \quad (7)$$

where $\Omega_i = \{j | e_{ij} = 1; i, j \in V\}$ is the neighbor set of node i , λ is a positive penalty constant and the set S is the feasible vertex cover solution.

Each vertex in V is treated as rational player. Then, we consider a networked game $G = (V, \{S_i\}, \{\zeta_i\})$. For each joint action $s = (s_1, s_2, \dots, s_n)$, where $s_i = 1$ means $i \in S$, whereas $s_i = 0$ otherwise, \bar{s}_i are the actions different from s_i of the same player, s_{-i} are the actions of other players. ζ_i is the local utility function, $\phi(s)$ is a potential function. Combining these definitions with some formulations are denoted as:

$$\begin{cases} \xi_i(s) = \max_{s_i \in S_i} \zeta_i(\bar{s}_i, s_{-i}) \\ \zeta_i(s'_i, s_{-i}) - \zeta_i(s_i, s_{-i}) = \phi(s'_i, s_{-i}) - \phi(s_i, s_{-i}) \\ \phi(s) = f(s) \\ \zeta_i(s) = -W(i) s_i - 2\lambda(1 - s_i) \sum_{j \in \Omega_i} (1 - s_j) \end{cases} \quad (8)$$

The change of $\phi(s)$ when the player i unilaterally changes from s_i to s'_i , whereas the rest keep their actions s_{-i} is examined, then based on equations (7) and (8), we have:

$$\phi(s'_i, s_{-i}) - \phi(s_i, s_{-i}) = \xi_i(s'_i, s_{-i}) - \xi_i(s_i, s_{-i}) \quad (9)$$

From the definitions of game theory, it must be a potential game. And the MWVC problem not only is the optimal Nash equilibrium, but is the optimal $f(s)$ in equation (7).

Algorithm design and decision-making framework

Modified population-based game-theoretic optimizer for minimum weight vertex cover problem

Given n individuals of population P , the i -th individual X_i is defined by m nodes $\{x_1, x_2, \dots, x_m\}$, and $n = \{1, 2, \dots, n\}$. Before the population $P^{t=0}$ that is initialized randomly starts to iterate, it needs to learn in game (LIG). The LIG process generates a nonredundant VC solution from initial $P^{t=0}$, and it is required for each node x_j of individual $X_{i \in N}^{t=0}$ satisfy some formulations, expressed by:

$$\begin{cases} b_j = \arg \max_{x_j \in X_j} \zeta_j(\bar{x}_j^{t=0}, x_j^{t=0}) \\ \eta_j = \zeta_j(b_j, x_{-j}^{t=0}) - \zeta_j(x_j^{t=0}) \\ k = \arg \max_{j \in N} \zeta_j \\ x_k^{t=1} = b_k \end{cases} \quad (10)$$

where b_j and η_j denote the best response and the regret, respectively. Then the node k with the highest regret is selected and updated its action within its best response. The whole learning process breaks when system has reached Nash equilibrium.

To find out the global best X_s^t , the objective function is designed via the minimum weight value W and the penalty value of uncovering edges, which is shown in Algorithm 1.

Algorithm 1: Objective function

Input: G, P^t

Output: f, X_s^t

- 1 $X_i^t = \{x_1^t, x_2^t, \dots, x_m^t\};$
- 2 $f_1 \leftarrow \sum_{i=1}^m W(i) \cdot x_i^t;$
- 3 Find the node with $x_m^t = 0$ in the set X_i^t ;
- 4 **for** each node $x_j^t \in X_i^t, j = 1, 2, \dots, \eta$ **do**
- 5 **for** each node $x_k^t \in X_i^t, k = 1, 2, \dots, m$ **do**
- 6 Find out the number $\Delta\vartheta$ of neighbour nodes which equal to 0 of x_j^t ;
- 7 $\vartheta \leftarrow \vartheta + \Delta\vartheta;$
- 8 **end**
- 9 $f_2 \leftarrow \vartheta \cdot m;$
- 10 $f \leftarrow -(f_1 + f_2);$
- 11 Obtain X_s^t which with the top f in P^t .

Then the DMC-SE process is applied to obtain new individuals V_i^t from each $X_i^t \in P^t$ based on global best X_s^t and a probability ρ of adaptive mutation. The DMC-SE is designed in Algorithm 2, where τ is a $1 \times n$ random vector which entry is 0 or 1 with equal probabilities, and \circ denotes the hadamard product operator, is a mathematical operation by multiplying corresponding elements on two matrices or vectors of the same size.

Algorithm 2: DMC-SE

Input: $P^t, X_s^t, R, \rho_{\max}, \rho_{\min},$ and τ

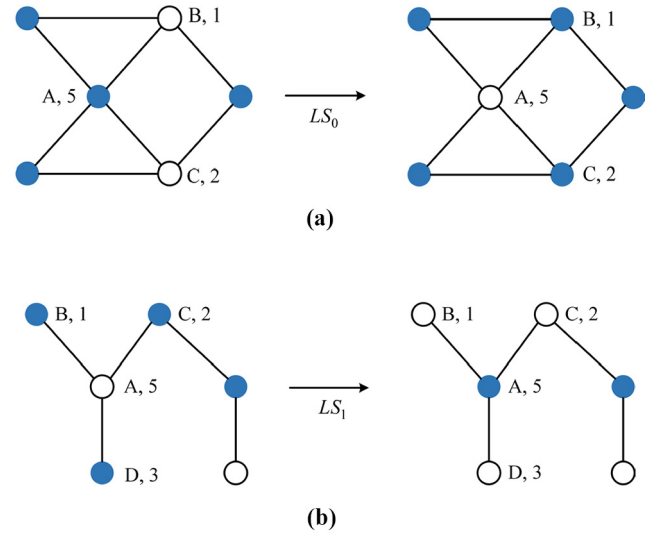
Output: V_i^t

- 1 **for** each individual $X_i^t, i = 1, 2, \dots, n$
- 2 **do** $v_1 \leftarrow \tau \circ (X_s^t - X_i^t);$ v_1 is a $1 \times m$ vector denotes an operator from X_s^t to X_i^t
- 3 **for** each $v_1(k), k = 1, 2, \dots, m$ **do**
- 4 **if** $\mu > e^{-Rt}$ **then** \cdot the random number $\mu \in (0, 1)$
- 5 $v_2(k) \leftarrow I - v_1(k);$
- 6 **else** $v_2(k) \leftarrow v_1(k);$ \cdot update velocity
- 7 **end**
- 8 **end**
- 9 $V_i^t \leftarrow X_i^t + v_2;$ \cdot update position
- 10 $\rho \leftarrow \rho_{\max} - (t - 1) \cdot (\rho_{\max} - \rho_{\min}) / (t_{\max} - 1);$ $\cdot \rho$ denotes mutation probability
- 11 **if** $\varsigma < \rho$ **then** \cdot the random number $\varsigma \in (0, 1)$
- 12 $V_i^t \leftarrow I - V_i^t;$
- 13 **else**
- 14 $V_i^t \leftarrow V_i^t;$ \cdot mutation process
- 15 **end**
- 16 **end**

In addition, to obtain a nonredundant VC solution, these individuals V_i^t that have passed DMC-SE need to through LIG again in equation (10), then new individuals are generated $[\text{Errorhx2202}]_i^t \leftarrow LIG(V_i^t)$. $[\text{Errorhx2202}]_i^t$ should experience LS procedure for acquiring a more superior VC solution, $\kappa_i^t \leftarrow LS([\text{Errorhx2202}]_i^t)$. By exchanging the actions of certain vertices, VC solution can obtain the minimum weight of covered nodes, the selected vertices are marked with blue.

In Figure 3, through a local exchange, the covered vertex A is replaced with the uncovered vertices B and C to obtain an improved VCs, as the weight of all vertices satisfies $W(A) - [W$

Figure 3 Local search of node A in LS_0 and LS_1



Notes: (a) LS_0 ; (b) LS_1

Source: Figure courtesy of Qiu *et al.* (2022)

$(B) + W(C)] = 2$. Similarly, the covered vertices B, C and D are replaced with the uncovered vertex A to obtain an improved VCs, as the weight of all vertices satisfies $[W(B) + W(C) + W(D)] - W(A) = 1$. Finally, all original individuals X_i^t are transformed into the optimal solutions κ_i^t , then an elite retention strategy is used that the n individuals that possess the top fitness value in the population P_i^t are inherited to the next iteration population P_i^{t+1} .

Decision-making framework

When UAVs-USVs perform reconnaissance missions, the fact that the position and status information of known targets is always changing, the unforeseen addition of unknown targets and the position and status information of UAVs-USVs heterogeneous system, which is also changing. As a result, a cooperative decision-making framework is devised to solve the unexpected combinatorial optimization problem, which is showed in Figure 4.

Information entropy evaluation

Moreover, information entropy may be used to analyze information transmission and reception, and it has steadily evolved into a tool for quantifying the information content of a system. As a result, in addition to examining the effectiveness of optimization algorithm in solving the reconnaissance issue, the ability to grasp the veracity of all target information also should be assessed. The absolute information entropy approach is used. Given a set $W = \{\omega_1, \omega_2, \dots, \omega_n\}$ of group events and the prior probabilities of these events and entropy function $H(t)$ are expressed as:

$$\begin{cases} H(t) = -\sum_{i=1}^n p_i(t) \log(p_i(t)) \\ p_i(t) \geq 0, \sum_{i=1}^n p_i(t) = 1 \end{cases} \quad (11)$$

where $H(t)$ is the information content of the set W , and the more chaotic the system, the higher the $H(t)$ value. We use the

traditional approach based on the binary assumption, which assumes that there are two states (authentic or deceptive) of every target with equivalent probability.

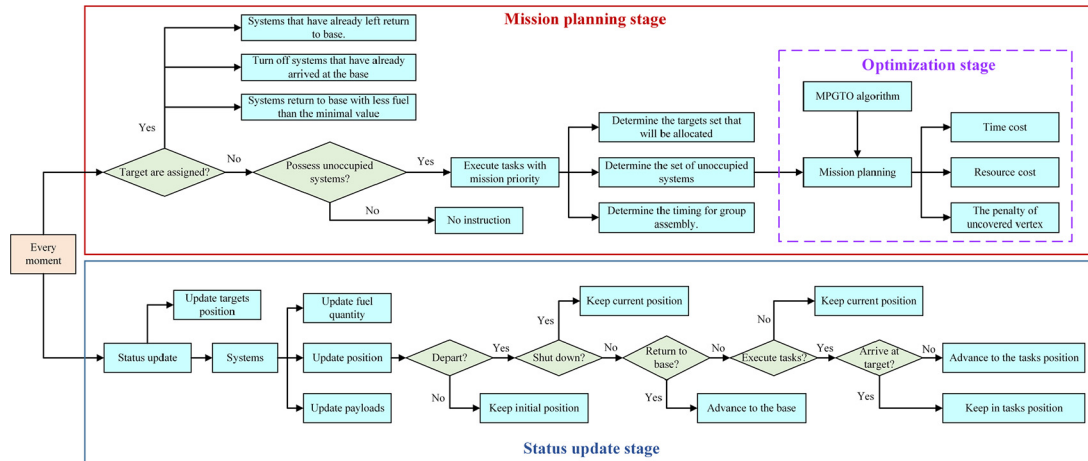
Simulation

Mission execution analysis

In this section, the whole scenario is set to a square of $1,650 \times 1,650 \text{ m}^2$, the initial positions of UAVs–USVs are arranged

randomly in a square of $200 \times 200 \text{ m}^2$; similarly, the initial positions of dynamic vessel targets are arranged randomly in a square of $1,000 \times 1,000 \text{ m}^2$. And the scenarios and constraints of UAVs–USVs heterogeneous system and targets are presented in Table 1. Aiming at the types of targets, the unknown targets contain the green authentic targets and the known targets contain the white deceptive targets and red authentic targets, are shown in Figure 5(a).

Figure 4 Decision-making framework of reconnaissance missions planning



Source: Author's own creation

Table 1 Scenarios and constraints of UAVs–USVs and targets

Name	Variable	Faster UAVs	Normal UAVs	USVs	Payload types
Case 1	n_1	40	40	20	Optical
Case 2	n_2	16	16	16	Optical
Case 3	n_3	16	16	16	Radar
Case 4	n_4	8	8	8	Radar
Velocity(m/s)	v	10	6	3	–
Detection range(m)	R_d	100	80	60	–
Fuel	F	500	500	500	–
Minimum fuel	F_{\min}	100	100	100	–
Base	(x, y)	(0,0)	(0,0)	(0,0)	–
Radar reconnaissance time	t_R	10	10	10	–
Number of radar payloads	N_R	2	2	2	–
Optical reconnaissance time	t_J	30	30	30	–
Number of optical payloads	N_J	4	4	4	–
Targets		Variable			Value
Known targets		T^k			20
Unknown targets		T^u			10
Deceptive targets		T^t			9
Authentic targets		T^a			21
x -axis unit velocity(m/s)		u			1
y -axis unit velocity(m/s)		v			1
x -axis change rate		r_1			(–1,1)
y -axis change rate		r_2			(–1,1)
Threat radius(m)		R_L			50
Authentic probability		p_1			0.5
Deceptive probability		p_2			0.5

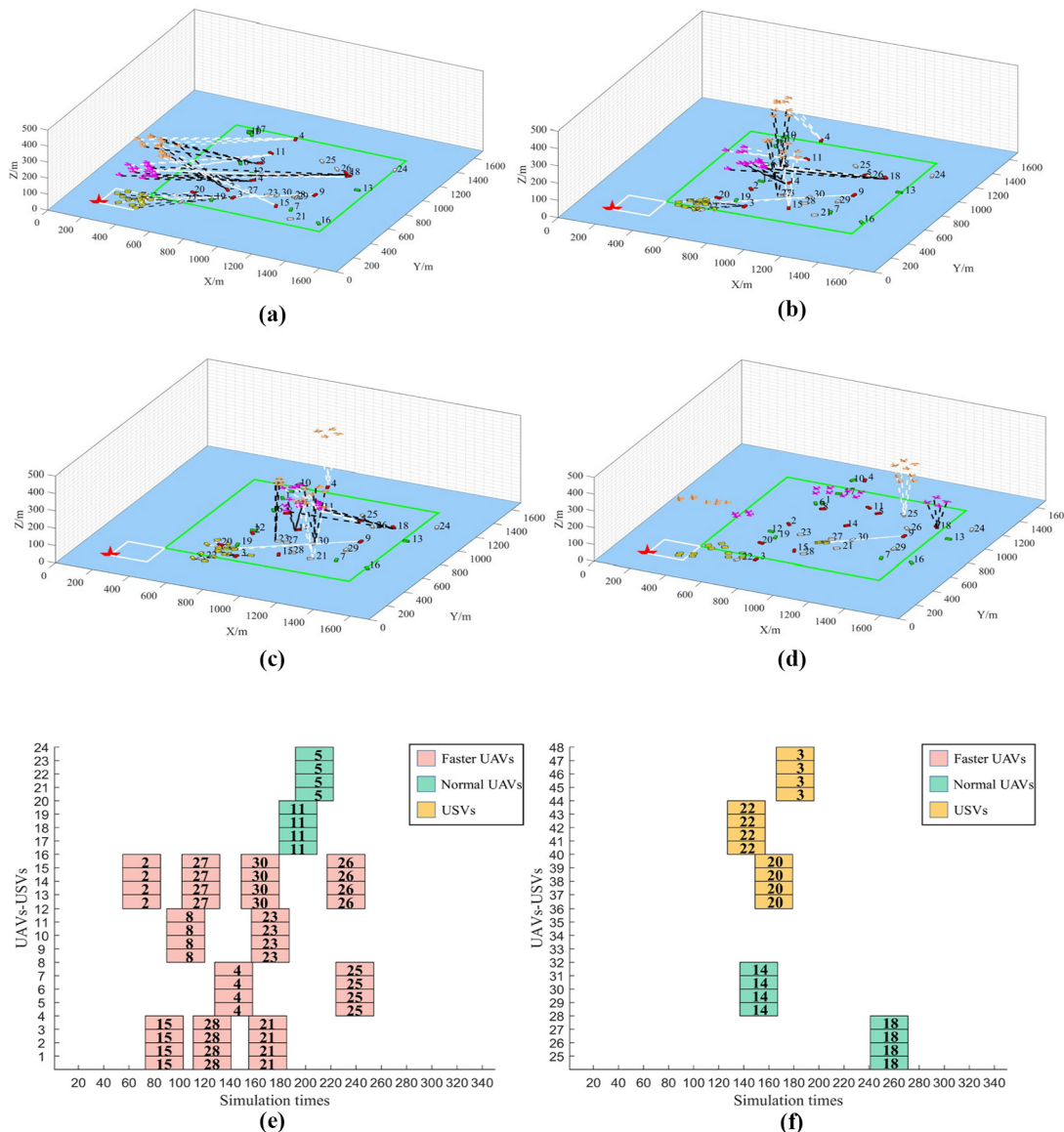
Source: Authors' own creation

From Figure 5(e), at $T_s = 20$, targets are assigned the UAVs–USVs initially and move randomly. Approximately, $T_s = 60$, the first group of UAVs–USVs reach the boundary of target No. 2, then turn on payloads and start the $\Delta T_s = 30$ reconnaissance mission around the boundary. From Figure 5(b) and 5(c), target No. 15 and target No. 8 are executed, and the UAVs–USVs are preparing to execute the target No. 27, meanwhile the other UAVs–USVs are advancing to the assigned targets. In Figure 5(c), the UAVs–USVs are reassigning and executing reconnaissance mission constantly. And in Figure 5(d), the UAVs–USVs that have completed their missions and have not been assigned missions is returning to the base.

When aiming at the radar reconnaissance missions. From Figure 6, UAVs–USVs turn on the payloads and start the $\Delta T_s = 10$ reconnaissance missions when approaching the boundary of targets, then consider to execute the remaining tasks or return to the base. Compared to using optical payloads, the network complexity of missions is predigested in the mission via the number and boot time of payloads.

Cases 2 and 3 show that when the detecting capability of the payloads is strong, the required reconnaissance time and number of payloads are reduced and just part of the UAVs need to be assigned tasks to complete the reconnaissance more effectively. When the detecting

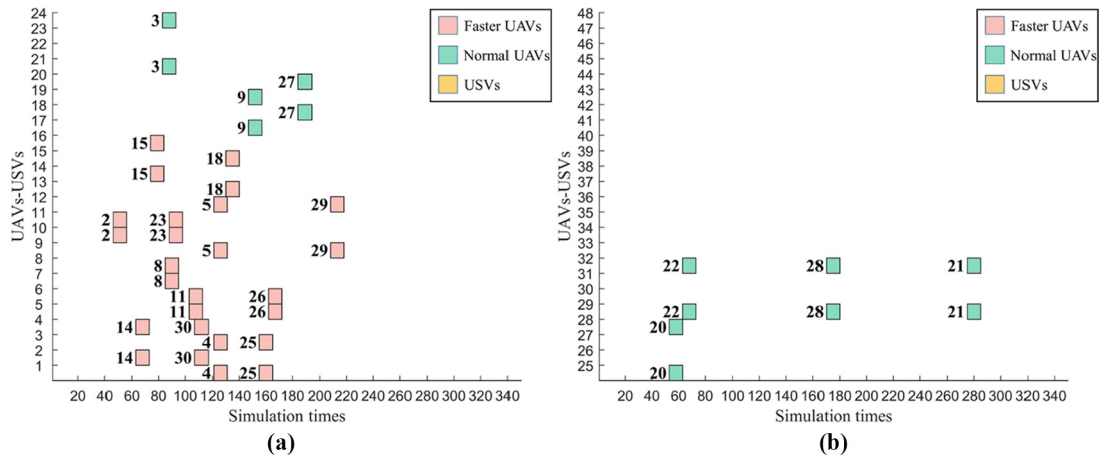
Figure 5 Process reconnaissance with 48 UAVs–USVs on 30 dynamic targets



Notes: (a) $T_s = 20$; (b) $T_s = 100$; (c) $T_s = 150$; (d) $T_s = 250$; (e) 1–24 UAVs–USVs; (f) 25–48 UAVs–USVs
(e) and (f) denote the gantt chart for reconnaissance mission planning using optical sensors

Source: Author's own creation

Figure 6 Gantt chart for reconnaissance mission planning



Notes: (a) 1 - 24 UAVs-USVs; (b) 25 - 48 UAVs-USVs
Source: Author's own creation

Table 2 Parameters of each optimization algorithms

Algorithm	Parameter	Variable	Value
MPGTO	Maximum iteration	T_{\max}	10
	DMC-SE operator	R	0.8
	Maximum mutation probability	ρ_{\max}	0.15
	Maximum mutation probability	ρ_{\min}	0
	Population size	n	5
PGTO	Maximum iteration	T_{\max}	10
	Mutation probability	ρ	0.15
	Population size	n	5
PBIG	Maximum iteration	T_{\max}	10
	Determination rate	α	0.5
	Degree of destruction	ρ_{de}	0.1
	Population size	n	5
MS-ITS	Maximum iteration	T_{\max}	10

Source: Authors' own creation

capability of the payloads is weak, the required reconnaissance time and number of payloads increased, both UAVs-USVs need to be deployed and faster UAVs need to accomplish more tasks.

Algorithm comparison

We created Cases 1, 3 and 4 with the same target status. Then, it was compared to various existing algorithms, such as basic PGTO, population-based iterated greedy (PBIG), multi-start iterated tabu search (MS-ITS), which can deliver remarkable optimization results for the MWVC problem (Bouamama *et al.*, 2012; Zhou *et al.*, 2016). Table 2 shows the parameters of each optimization method. After the amount of experiments, the timeliness of all algorithms is first considered, which includes the mean time $\bar{\phi}$, the minimum time ϕ_{\min} , the maximum time ϕ_{\max} , the variance δ and the total time T , as given in Table 3.

In Case 4, when the mission network is tiny, the PBIG is more timely, and the MPGTO in second place is roughly 21.74

Table 3 Comparison results for cases with various algorithms

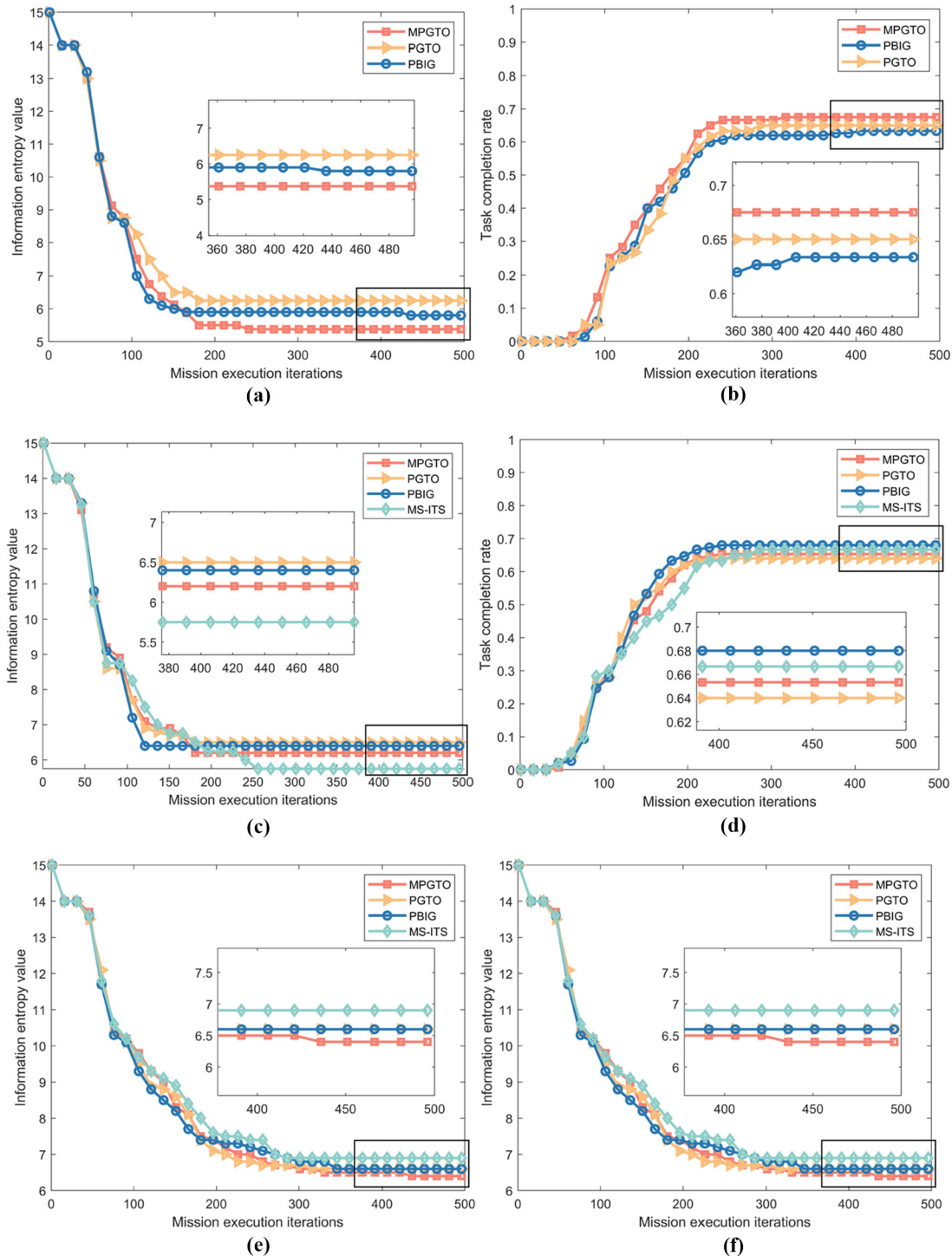
Scenario	Algorithm	$\bar{\phi}(s)$	$\phi_{\min}(s)$	$\phi_{\max}(s)$	δ	$T(s)$	$Q(\%)$	H
Case 1	MPGTO	2.23E-01	7.41E-04	9.84E + 01	1.94E + 01	1.11E + 02	67.50	5.38
	PGTO	2.83E-01	6.53E-04	1.26E + 02	3.19E + 01	1.42E + 02	66.67	6.10
	PBIG	1.04E + 00	7.47E-04	5.16E + 02	5.32E + 02	5.20E + 02	63.33	5.80
	MS-ITS	—	—	—	—	>7.20E + 03	—	—
Opt		2.23E-01	6.53E-04	9.84E + 01	1.94E + 01	1.11E + 02	67.50	5.38
Case 3	MPGTO	9.30E-03	4.99E-04	3.16E + 00	2.01E-02	4.65E + 00	65.33	6.20
	PGTO	1.16E-02	4.93E-04	4.09E + 00	3.38E-02	5.78E + 00	64.17	6.50
	PBIG	1.09E-02	5.08E-04	4.63E + 00	4.28E-02	5.45E + 00	68.00	6.40
	MS-ITS	7.53E + 00	4.96E-04	3.61E + 03	2.60E + 04	3.76E + 03	66.67	5.75
Opt		9.30E-03	4.93E-04	3.16E + 00	2.01E-02	4.65E + 00	68.00	5.75
Case 4	MPGTO	2.80E-03	4.63E-04	7.31E-01	1.10E-03	1.38E + 00	58.14	6.40
	PGTO	3.10E-03	4.65E-04	9.16E-01	1.70E-03	1.56E + 00	59.33	6.60
	PBIG	2.30E-03	4.62E-04	6.25E-01	7.80E-04	1.13E + 00	59.97	6.60
	MS-ITS	1.03E + 00	4.57E-04	4.95E + 02	4.89E + 02	5.15E + 02	61.33	6.90
Opt		2.30E-03	4.57E-04	6.25E-01	7.80E-04	1.13E + 00	61.33	6.40

Source: Authors' own creation

longer than the PBIG in the $\bar{\phi}$. However, when the amount of UAVs–USVs is increased till Case 3, compared to the PBIG, MPGTO is more timely and $\bar{\phi}$ is improved by 14.7%, ϕ_{\max} is modified by 22.7% with PGTO, δ and T both have an exceptional improvement in comparison to other algorithms.

Because the system has already converged and the optimal solution has generated in the later simulation time, the ϕ_{\min} is irregular number that denotes the time of running once main program without executing optimal program. Comparing Cases 1 and 3, military from 48 UAVs–USVs for radar

Figure 7 Results of algorithm comparison



Notes: (a) Information entropy value; (b) task completion rate; (c) information entropy value; (d) task completion rate; (e) information entropy value; (f) task completion rate

Source: Author's own creation

reconnaissance missions to 100 UAVs-USVs for optical reconnaissance missions, MS-ITS has been exceeded two hours in once optimization process, which can hardly be considered in real-time mission planning. For the most complicated Case 1, the advantages that MPGTO is applied to vast network are reflected. Compared with PGTO that is the second optimal algorithm, the ϕ is improved by 21.2%, ϕ_{\max} is modified by 21.9%, other indicators also promote significantly.

Next, the quality of mission planning solution is considered. In Case 1, as MS-ITS algorithm processes the mission planning problem inefficiently, thus it is not appropriate for real-time comparison. Then comparing Cases 1, 3 and 4, the four algorithms can always reach convergence and approximate optimal solution with the given required computation time, which is shown in Figure 7. For the task completion rate, as only authentic targets can be executed by UAVs-USVs, therefore it will approach 0.7 from the data in Table 1. For the information entropy value, at the beginning of mission, UAVs-USVs cannot distinguish the status of all targets, thus the entropy stay at a large value. When the missions are executed continuously, the entropy value will continue to approach 4.5 from the data in Table 1 and definition in equation (11). In addition, through analysing the tendency of information entropy value H and task completion rate Q with the simulation time, when mission planning problem with small-size network and low complexity of tasks, four algorithms can all reach a superior convergence value with the given required computation time. However, with the size of network and the complexity of tasks are increased, MS-ITS will be unsuited in the scenario, the performance of PBIG and PGTO also will fall behind MPGTO.

Finally, the proposed algorithm solves the reconnaissance mission planning problem in less time and has certain benefits over some current algorithms in terms of solution quality. Furthermore, with the increased complexity of tasks and multiple expansion of UAVs-USVs, the optimization ability is even more prominent due to the mechanism of internal competition and collaboration.

Conclusions and future plans

The resolution of the dynamic task allocation problem in traditional, large-scale heterogeneous USs poses challenges in achieving both low computational complexity and excellent optimization results. This paper designed a transformation of the task allocation problem into the MWVC problem within a game-theoretic framework. This approach introduces competitive and cooperative autonomy in the allocation process, thereby mitigating the difficulties associated with centralized allocation methods, such as high computational complexity, poor real-time performance and inability to obtain solutions. To optimize the MWVC problem, a novel MPGTO algorithm is proposed, capable of obtaining approximate optimal solutions. A series of simulation experiments is conducted to validate the feasibility and effectiveness of the MPGTO algorithm after incorporating the evolutionary strategy. With the increasing complexity of US reconnaissance missions and expanding networks, the algorithm overcomes the challenge of requiring more time and resources for large-scale tasking problems in terms of mission planning, resource scheduling for USs in the military. These

findings demonstrate the significance of the improvement in PGTO and the design of the MWVC approach. Future work should consider more complex models involving UAVs-USVs with payload and resource constraints. In addition, further analysis from a game-theoretic perspective is needed to enhance the autonomy of UAVs-USVs and authentic scenarios should be considered as well.

References

- Alazzam, A., Sharieh, A. and Sabri, K.E. (2020), "A feature selection algorithm for intrusion detection system based on pigeon inspired optimizer", *Expert Systems with Applications*, Vol. 148, p. 113249.
- Berger, J., Barkaoui, M. and Boukhtouta, A. (2007), "A hybrid genetic approach for airborne sensor vehicle routing in real-time reconnaissance missions", *Aerospace Science and Technology*, Vol. 11 No. 4, pp. 317–326.
- Bouamama, S., Blum, C. and Boukerram, A. (2012), "A population-based iterated greedy algorithm for the minimum weight vertex cover problem", *Applied Soft Computing*, Vol. 12 No. 6, pp. 1632–1639.
- Chen, J. and Li, X. (2023), "Toward the minimum vertex cover of complex networks using distributed potential games", *Science China Information Sciences*, Vol. 66 No. 1, p. 112205.
- Duan, H. and Qiao, P. (2014), "Pigeon-inspired optimization: a new swarm intelligence optimizer for air robot path planning", *International Journal of Intelligent Computing and Cybernetics*, Vol. 7 No. 1, pp. 24–37.
- Duan, H., Zhao, J., Deng, Y., Shi, Y. and Ding, X. (2021), "Dynamic discrete pigeon-inspired optimization for multi-UAV cooperative search-attack mission planning", *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 57 No. 1, pp. 706–720.
- Gao, T. and Bhattacharya, S. (2019), "Multirobot charging strategies: a game-theoretic approach", *IEEE Robotics and Automation Letters*, Vol. 4 No. 3, pp. 2823–2830.
- Ghassemi, P. and Chowdhury, S. (2022), "Multi-robot task allocation in disaster response: addressing dynamic tasks with deadlines and robots with range and payload constraints", *Robotics and Autonomous Systems*, Vol. 147, p. 103905.
- Hassija, V., Saxena, V. and Chamola, V. (2020), "Scheduling drone charging for multi-drone network based on consensus time-stamp and game theory", *Computer Communications*, Vol. 149, pp. 51–61.
- Kusyk, J., Uyar, M.U., Ma, K., Samoylov, E., Valdez, R., Plishka, J., Hoque, S.E., Bertoli, G. and Boksiner, J. (2021), "Artificial intelligence and game theory controlled autonomous UAV swarms", *Evolutionary Intelligence*, Vol. 14 No. 4, pp. 1775–1792.
- Mebrek, A. and Yassine, A. (2021), "Intelligent resource allocation and task offloading model for IoT applications in fog networks: a game-theoretic approach", *IEEE Transactions on Emerging Topics in Computational Intelligence*, pp. 1–15.
- Nemer, I.A., Sheltami, T.R. and Mahmoud, A.S. (2020), "A game theoretic approach of deployment a multiple UAVs for optimal coverage", *Transportation Research Part A: Policy and Practice*, Vol. 140, pp. 215–230.

- Qiu, H., Sun, C., Wang, X., Sun, W. and Zhou, Q. (2022), “A population-based game-theoretic optimizer for the minimum weighted vertex cover”, *Applied Soft Computing*, Vol. 116, p. 108272.
- Safar, M., Taha, M. and Habib, S. (2007), “Modeling the communication problem in wireless sensor networks as a vertex cover”, *IEEE/ACS International Conference on Computer Systems and Applications*, pp. 592–598.
- Semasinghe, P., Maghsudi, S. and Hossain, E. (2017), “Game theoretic mechanisms for resource management in massive wireless IoT systems”, *IEEE Communications Magazine*, Vol. 55 No. 2, pp. 121–127.
- Sheng, M., Zhao, C., Liu, J., Teng, W., Dai, Y. and Li, J. (2022), “Energy-efficient trajectory planning and resource allocation in UAV communication networks under imperfect channel prediction”, *Science China Information Sciences*, Vol. 65 No. 12, p. 222301.
- Stolfi, D.H., Brust, M.R., Danoy, G. and Bouvry, P. (2021), “UAV-UGV-UMV multi-swarms for cooperative surveillance”, *Frontiers in Robotics and AI*, Vol. 8, p. 616950.
- Sun, C., Sun, W., Wang, X. and Zhou, Q. (2019), “Potential game theoretic learning for the minimal weighted vertex cover in distributed networking systems”, *IEEE Transactions on Cybernetics*, Vol. 49 No. 5, pp. 1968–1977.
- Sun, C., Wang, X., Qiu, H. and Zhou, Q. (2021), “Game theoretic self-organization in multi-satellite distributed task allocation”, *Aerospace Science and Technology*, Vol. 112, p. 106650.
- Suresh, M. and Ghose, D. (2012), “UAV grouping and coordination tactics for ground attack missions”, *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 48 No. 1, pp. 673–692.
- Tong, B., Wei, C. and Shi, Y. (2022), “Fractional order Darwinian pigeon-inspired optimization for multi-UAV swarm controller”, *Guidance, Navigation and Control*, Vol. 2 No. 2, p. 2250010.
- Wu, H. and Shang, H. (2020), “Potential game for dynamic task allocation in multi-agent system”, *ISA Transactions*, Vol. 102, pp. 208–220.
- Wu, W., Cui, N., Shang, W. and Wang, X. (2018), “Distributed task allocation for multiple heterogeneous UAVs based on consensus algorithm and online cooperative strategy”, *Aircraft Engineering and Aerospace Technology*, Vol. 90 No. 9, pp. 1464–1473.
- Wu, H., Ghadami, A., Bayrak, A.E., Smereka, J.M. and Epureanu, B.I. (2021), “Impact of heterogeneity and risk aversion on task allocation in multi-agent teams”, *IEEE Robotics and Automation Letters*, Vol. 6 No. 4, pp. 7065–7072.
- Yoon, S., Do, H. and Kim, J. (2020), “Collaborative mission and route planning of multi-vehicle systems for autonomous search in marine environment”, *International Journal of Control, Automation and Systems*, Vol. 18 No. 3, pp. 546–555.
- Yuan, Y. and Duan, H. (2022), “Active disturbance rejection attitude control of unmanned quadrotor via paired coevolution pigeon-inspired optimization”, *Aircraft Engineering and Aerospace Technology*, Vol. 94 No. 2, pp. 302–314.
- Yuan, Y., Xu, X., Duan, H., Zeng, Z., Xu, D., Chen, R. and Wu, T. (2022), “Eagle vision-based coordinate landing control framework of unmanned aerial vehicles on an unmanned surface vehicle”, *Guidance, Navigation and Control*, Vol. 2 No. 4, p. 2250023.
- Zhang, D. and Duan, H. (2018), “Social-class pigeon-inspired optimization and time stamp segmentation for multi-UAV cooperative path planning”, *Neurocomputing*, Vol. 313, pp. 229–246.
- Zhou, T., Lü, Z., Wang, Y., Ding, J. and Peng, B. (2016), “Multi-start iterated Tabu search for the minimum weight vertex cover problem”, *Journal of Combinatorial Optimization*, Vol. 32 No. 2, pp. 368–384.

Corresponding author

Haibin Duan can be contacted at: hbduan@buaa.edu.cn