Unmanned Systems, Vol. 14, No. 4 (2026) 819–838
© World Scientific Publishing Company
DOI: 10.1142/S2301385026500214

UNMANNED SYSTEMS
https://www.worldscientific.com/worldscinet/us

Hierarchical Attack-Defense Strategy of Unmanned Surface Vehicles for Domain Protection and Invasion Task

Jinduo Wang, Huayan Pu, Zhaoxiang Shi, Zilong Song, Heng Fu, Jianjun Wu, Yangmin Xie*

School of Mechatronics Engineering and Automation, Shanghai University, Shanghai 200444, P. R. China

This paper introduces an innovative hierarchical framework tailored to meet the offensive and defensive needs of swarm-based unmanned surface vehicles (USVs) for domain protection and invasion tasks. The framework is structured into three layers: The top layer handles target allocation, the middle layer focuses on action strategy, and the bottom layer ensures safety. For domain protection tasks, the top layer leverages auction theory to effectively match attackers and defenders, enabling the distributed allocation of multiple defenders to respective invaders. The middle layer employs an adaptive guidance method, allowing each defender to follow a guidance law aligned with the optimal interception point. The bottom layer uses an enhanced potential field function to guarantee collision-free interception of invaders. For invasion tasks, the top layer similarly applies auction theory for the allocation of virtual waypoints. The middle layer introduces an Artificial Potential Field (APF)-based angle-adaptive attack strategy, enabling invaders to evade defender interception. The bottom layer remains consistent with the domain protection framework. This paper marks the first hybrid virtual-real experiment in multi-USV confrontation, demonstrating its reliability and effectiveness in real maritime conditions. Experiments conducted within a 5000-m radius, with vessels traveling at speeds up to 12 m/s, confirmed the framework's robustness and practical applicability in near-realistic combat scenarios.

Keywords: Hierarchical attack-defense strategy; unmanned surface vehicle (USV); multi-USV confrontation; obstacle avoidance; maritime experiments.

1. Introduction

In recent years, Unmanned Surface Vehicle (USV) technology has advanced rapidly. As an autonomous robotic system, USVs demonstrate excellent performance in cruising, path planning, and intelligent obstacle avoidance [1]. Coordinated USV clusters overcome the limitations of individual USVs, allowing them to perform more complex and large-scale task. Supported by real-time data communication and multi-vehicle formation techniques, USV clusters can engage in coordinated combat and execute various military missions, such as infiltration detection, deception, and swarm attacks [2]. Consequently, they have garnered increasing research attention [3].

Current research on USV swarms primarily focuses on multi-USV path tracking [4–8], multi-USV target tracking [9, 10], and the formation control [11–13]. These studies are mainly suited for cooperative task in nonconfrontational scenarios. Regarding works involving noncooperative confrontations, most studies focus on the simple Nv1 pursuit problem. These studies either utilize traditional control methods, such as optimal control with state estimators [14], traditional planning methods like the Apollonius circle [15], and artificial potential field (APF) [15], or the

Received 4 September 2024; Revised 1 March 2025; Accepted 2 March 2025; Published 17 June 2025. This paper was recommended for publication in its revised form by editorial board member, Haitao Zhang.
Email Addresses: *xieym@shu.edu.cn
*Corresponding author.

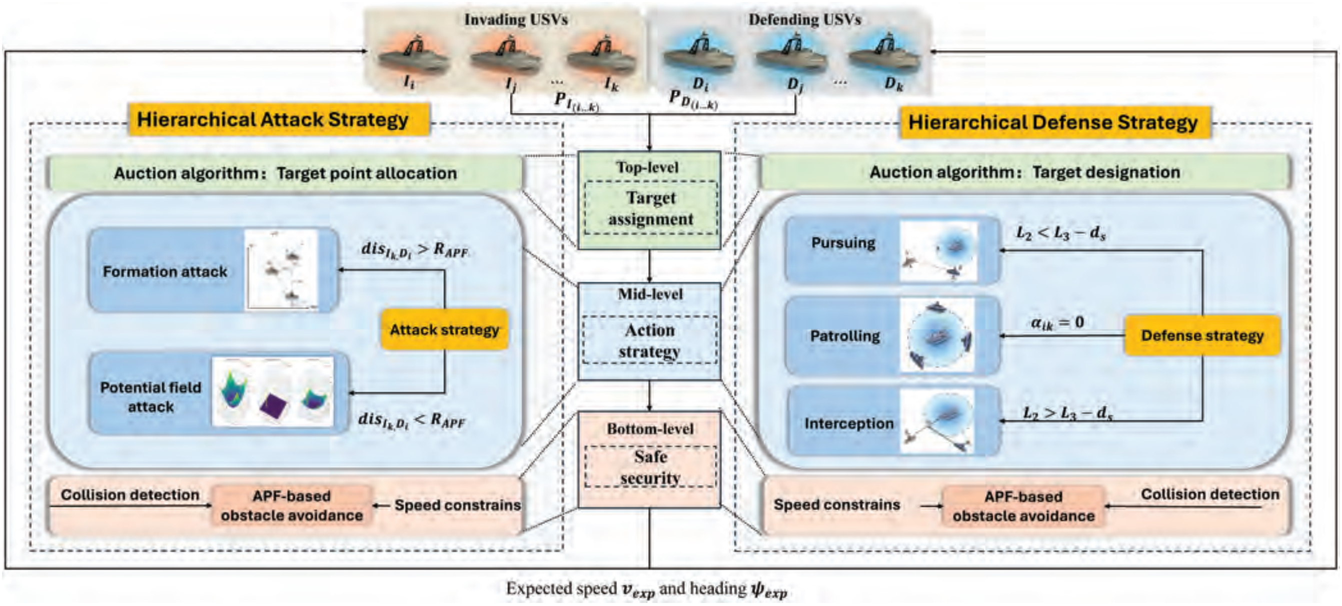


Fig. 1. Overview of the hierarchical attack-defense strategy.

recently popularized reinforcement learning methods [16]. For such deterministic and simple task, the above works generally only employ a single-layer action control layer.

In recent research addressing complex multi-agent tasks, the necessity of decomposing scenarios into hierarchical layers has been increasingly recognized. For instance, [17] proposed a hierarchical framework for 1V1 aerial combat, incorporating a higher-level decision-making layer for switching between offensive and defensive states and a lower-level collision avoidance layer for emergency control. Similarly, [18] introduced a grouping defensive strategy layer to enhance intra-group coordination in NVN engagements, while [19] addressed domain protection and invasion tasks using a three-layer strategy. However, these works are limited either to simulation-based evaluations or, as in [19], to small-scale real-world experiments within a 50-m radius water area.

While previous studies have shown that hierarchical frameworks provide greater flexibility, adaptability, and robustness in handling complex tasks and environments compared to end-to-end decision-making in unmanned system adversarial scenarios, significant gaps remain. Specifically, there is a lack of bilateral strategy frameworks tailored for both offensive and defensive tasks in multi-USV settings, as well as insufficient real-system validation under near-realistic USV actuation capabilities. To address these challenges, this paper proposes a hierarchical bilateral adversarial strategy framework, as illustrated in Fig. 1, with the key innovations and contributions summarized as follows:

(1) This paper presents a novel hierarchical framework specifically designed to address the complexities of multi-role confrontation tasks in multi-agent scenarios. Unlike previous approaches that focus on single-role (only for offenders or defenders) or symmetric confrontations (e.g. 1v1 scenarios), our framework enables dynamic coordination mechanisms for both offensive and defensive roles, allowing seamless adaptation to diverse task requirements and operational complexities. This supports robust multi-role cooperation and adversarial decision-making in complex multi-agent environments.

(2) The framework introduces a comprehensive set of modular strategies tailored for offensive and defensive tasks: On the offensive side, it includes a formation attack strategy and an angle-adaptive potential field-based attack method for precise target engagement and enhanced evasion. On the defensive side, it features an adaptive interception strategy and a patrol mechanism to ensure continuous domain protection and effective threat neutralization. The modular design enhances flexibility and scalability, allowing seamless upgrades or replacements of individual modules to adapt to dynamic scenarios and evolving mission requirements.

(3) To the best of our knowledge, this is the first work to conduct hybrid virtual-real experiments in the domain of multi-USV confrontations. By integrating simulated target vessels with real USVs in a shared maritime environment, we provide a realistic testbed for evaluating algorithm performance. Conducted under realistic conditions with USVs achieving speeds of up to 12 m/s within a 5000-m operational range, these experiments demonstrate the framework's reliability, practical applicability,

and effective transition from simulation to real-world deployment. Compared to state-of-the-art methods, our framework significantly improves collision avoidance efficiency, defense quality, and interception success in 3v3 confrontation tasks, highlighting its superior performance and adaptability.

The rest of this paper is organized as follows. Section 2 reviews related work. Section 3 presents the problem formulation. Section 4 details the hierarchical attack strategy for the domain protection task. Section 5 describes the hierarchical attack strategy for the invasion task. Section 6 provides the experimental results. Finally, Sec. 7 concludes this paper.

2. Related Work

In recent years, enhancing the cooperative capabilities of multi-agent systems in invasion and defense scenarios has become a major research focus, as it is crucial for achieving victory in adversarial environments. Previous studies have explored various approaches, which can be broadly categorized into rule-based methods, imitation-based learning methods, and reinforcement learning methods.

2.1. *Rule-based methods*

Rule-based approaches heavily rely on predefined rules and expert policies to guide agent actions. However, rule-based methods are susceptible to exploitation when adversaries target weaknesses identified through strategic analysis [20] and counter-attack tactics [21]. Therefore, several studies have integrated rule-based agents with other artificial intelligence methodologies. Masek *et al.* [22] used evolutionary FSM to discover emergent behavior [23] and employed flexible rules through the evolution of behavior trees. Developing a rule-based system that can adapt to all environments is impractical, and even well-designed rules can prove problematic. Consequently, the rule-based approach is now often employed as a supplementary module to enhance specific performance aspects in other methodologies. For instance, [24] used purchase strategy scripts, and [25] incorporated game-specific knowledge to filter out low-probability actions.

2.2. *Deep reinforcement learning*

Multi-agent reinforcement learning (MARL) methods enable multi-agent systems to make strategic decisions based on environmental conditions. The continuous decision-making process allows agents to learn cooperation and adapt adversarial strategies in dynamic and complex environments. In the context of adversarial air combat, deep RL decision-making methods based on DQN [26] and SAC [27] are proposed. To achieve better performance and stability in training, hierarchical reinforcement learning (HRL) methods have been introduced in adversarial environments, such as air-to-air combat [28], StarCraft II [29], and MOBA games [30]. Combined with self-play learning, [31] proposed a Multi-Agent Hierarchical Policy Gradient algorithm (MAHPG) to learn various strategies. Our framework advances beyond previous studies by supporting multi-role, multi-agent confrontation tasks, offering superior adaptability, scalability, and reduced dependence on expert data and extensive training.

Although Multi-agent Deep Reinforcement Learning (MADRL) has outperformed expert-level humans in target confrontation scenarios, the primary challenge remains how to narrow the gap between theoretical environments and real-world ones in adversarial decision-making. Our study uniquely demonstrates the algorithm's reliability and real-world applicability in multi-USV confrontations using a hybrid virtual-real system, highlighting its practicality under realistic maritime conditions.

To address the shortcomings of the aforementioned applications, this paper aims to establish a hierarchical framework suitable for multi-tasking in multi-UAV confrontation environments. Based on the above scenario, this paper proposes a hierarchical attack-defense strategy framework that solves the domain protection and domain invasion problems.

3. Problem Formulation

3.1. *Problem statement for USVs groups defense and attack scenarios*

In this paper, we consider two groups of USVs, which are defenders and invaders, as shown in Fig. 2, $D_i(i = 1, 2, \dots, n)$ represents the defenders, $I_i(i = 1, 2, \dots, n)$ represents the invaders. Define the domain protected by defenders as $\Omega_p = \{\boldsymbol{p} \in \mathbb{R}^2 | \|\boldsymbol{p} - \boldsymbol{p}_{\Omega_p}\| \leq r_{\Omega_p}\}$ with $\boldsymbol{p}_{\Omega_p} \in \mathbb{R}^2$ and $r_{\Omega_p} \in \mathbb{R}_+$. $\boldsymbol{p}$ represents the position vector of a point in space and $\boldsymbol{p}_{\Omega_p}$ represents the central position of the protected domain Ω_p. For the position of a USV i, define $\boldsymbol{p}_i = [x_i, y_i]^T$. Meanwhile, $\boldsymbol{p}_{\Omega_p} = [x_{\Omega_p}, y_{\Omega_p}]^T$. Define an interception domain $\Omega_I(\Omega_p \subset \Omega_I)$ of the defenders as $\Omega_I = \{\boldsymbol{p} \in \mathbb{R}^2 | \|\boldsymbol{p} - \boldsymbol{p}_{\Omega_I}\| \leq r_{\Omega_I}\}$ with $p_{\Omega_I} \in \mathbb{R}^2$ and $r_{\Omega_I} \in \mathbb{R}_+$. In this paper, we consider both defense and attack problems, which are described as follows. (1) For the domain protection task, the goal of the defender is to prevent the invaders from invading domain Ω_p; (2) For the invasion task, the goal of the invaders is to evade the defender's interception and enter Ω_p. If all invaders are intercepted before reaching Ω_p, the defenders win. Otherwise, the invaders win.

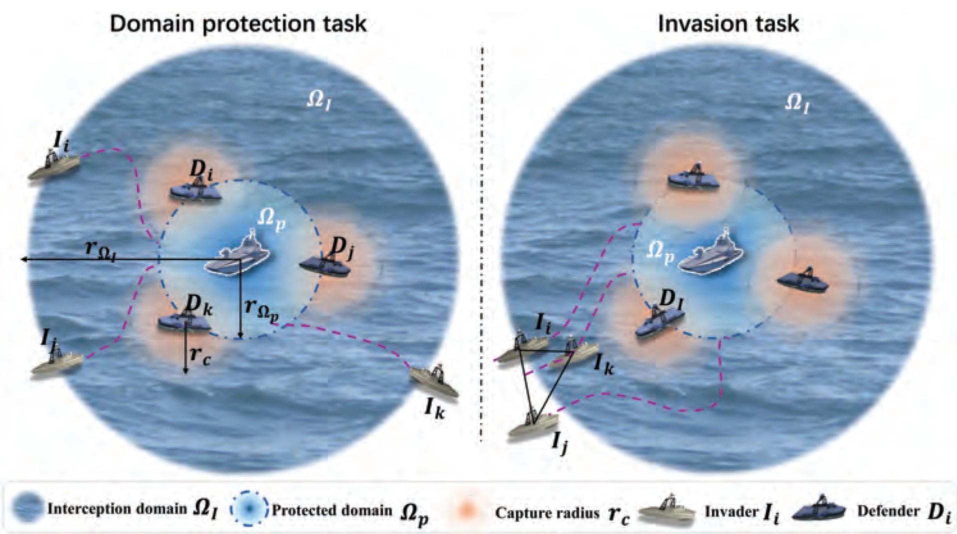


Fig. 2. USV's group defense and attack scenarios.

In this study, only the defender has an attack capability. To quantify the survival status of the invaders, the concept of Health Points (HPs) is introduced. When an invader remains within the attack zone, its HP denoted as ph_I for invaders, decreases by the following equation:

$$ph_I = \hat{ph}_f - \eta t_{in}, \tag{1}$$

where ph_I represents the remaining HP of the invaders, $\hat{ph}_f$ signifies the invader's full HP, η is the decreasing rate, and t_{in} is the total time that the vessel is inside the attackers' attack zones.

The results for the USV group's defense and attack scenarios are classified into two cases: win and lose. Denote the distance between the invader I_i and the Ω_p as $\text{dis}_{I_i,\Omega_p}(t)$ at time t. The criteria for the two results at time t are listed as follows. If any condition of these cases is triggered, the game is over.

- Win: $\forall\, i \in \{1, 2, \dots, n\}, ph_{I_i}(t) = 0 \wedge \text{dis}_{I_i,\Omega_p}(t) > 0$.
- Loss: $\exists\, i \in \{1, 2, \dots, n\}, \text{dis}_{I_i,\Omega_p}(t) \leq 0$.

Considering the problem formulation outlined above, the setup for the confrontational groups is based on the following assumptions:

Assumption 1. The defenders and invaders have identical movement capabilities, as specified in Sec. 3.2. Specifically, their linear velocities are constrained to the same range: $[v_{\min}, v_{\max}]$, such that:

$$\forall\, i \in \{1, 2, \dots, n\}, v_i^D, v_i^I \in [v_{\min}, v_{\max}]. \tag{2}$$

Assumption 2. The defenders and invaders can always access global information, including the kinematics of the vessels' real-time positions and velocities in both groups.

3.2. *USV kinematic model*

The USV's kinematic model describes its movement capability during fighting. Define $\boldsymbol{p}_i^D = [x_i, y_i]^T$ and ψ_i as the position and the heading angle, respectively, of the D_i. For both the defender and the invader, their kinematic model in a 2D space is formulated as shown as follows:

$$\begin{cases} \dot{v} = a_v, \\ \dot{p_x} = v \cos\psi, \\ \dot{p_y} = v \sin\psi, \\ \dot{\psi} = \omega, \end{cases} \tag{3}$$

where (p_x, p_y) represents the position of a USV; v, ψ, ω are its speed, heading angle, and angular velocity, respectively. $a_v \in [a_{v,\min}, a_{v,\max}]$ is the acceleration. $v_{\min}$ and $v_{\max}$ denote the minimum and maximum values for linear velocities. Usually, we do not consider a negative linear velocity, and thus $v_{\min} = 0$.

3.3. *Artificial potential field method*

The APF method, first proposed by Khatib in 1985 [32], utilizes target forces to navigate robots. The core idea is to place the USV within its environment and design its movement based on an artificial gravitational field. This field consists of two primary forces: an attractive force that attracts the USV toward the target waypoint and a repulsive force that pushes the USV away from obstacles in its path, preventing collisions. The strength of the repulsive force increases as the USV gets closer to an obstacle, thereby ensuring safe navigation. The combined effect of these forces generates a resultant force that determines the USVs movement. The resultant force is a vector sum of the attractive and repulsive forces, guiding the USV toward the target while avoiding obstacles.

The attractive potential U_{att} is defined to pull the USV toward the target. It can be expressed as

$$\boldsymbol{U}_{\mathrm{att}} = \frac{1}{2}\mu_{\mathrm{att}}\|\boldsymbol{p}_{\mathrm{USV}} - \boldsymbol{p}_{\mathrm{target}}\|^2, \tag{4}$$

where μ_{att} is a positive constant, p_{USV} is the current position of the USV, and p_{target} is the position of the target.

The attractive force F_{att} is the negative gradient of the attractive potential U_{att}:

$$F_{\mathrm{att}} = -\nabla U_{\mathrm{att}} = -\mu_{\mathrm{att}}(p_{\mathrm{USV}} - p_{\mathrm{target}}). \tag{5}$$

The repulsive potential U_{rep} is designed to push the USV away from obstacles. It can be expressed as

$$U_{\mathrm{rep}} = \begin{cases} \frac{1}{2}\mu_{\mathrm{rep}}\left(\frac{1}{d_a} - \frac{1}{d_0}\right)^2 & \text{if } d_a \leq d_0, \\ 0 & \text{if } d_a > d_0, \end{cases} \tag{6}$$

where $d_a = \|p_{\mathrm{USV}} - p_{\mathrm{obs}}\|$, which is the distance from the USV to the nearest obstacle. d_0 is the influence range of the repulsive potential field. When the distance between the USV and the obstacle exceeds this distance, the repulsive potential field is 0. μ_{rep} is the repulsive force coefficient, which controls the magnitude of the repulsive force on the USV.

The repulsive force F_{rep} is the negative gradient of the attractive potential U_{rep}:

$$F_{\mathrm{rep}} = \begin{cases} \mu_{\mathrm{rep}}\left(\frac{1}{d_a} - \frac{1}{d_0}\right)\frac{1}{d_a}^{2}\nabla d_a & \text{if } d_a \leq d_0, \\ 0 & \text{if } d_a > d_0, \end{cases} \tag{7}$$

where $\nabla d_{\mathrm{USV,obs}}$ is the direction of the distance gradient to the obstacle, usually pointing toward the obstacle.

The total force F_{tot} of the robot at any position is the resultant force of attraction and repulsion:

$$F_{\mathrm{tot}} = F_{\mathrm{att}} + F_{\mathrm{rep}}. \tag{8}$$

4. Hierarchical Defense Strategy for the Domain Protection Task

For the domain protection task, this study proposes the hierarchical defense strategy to intercept invaders and protect Ω_p. The overall architecture of this strategy is shown in Fig. 1, comprising three components: target designation, defense strategy, and APF-based obstacle avoidance.

4.1. *Auction-based target designation*

Decentralized methods distribute decision-making across multiple agents, offering improved scalability, robustness, and adaptability compared to centralized approaches. By eliminating reliance on a single point of control, they reduce the risk of failure and enable real-time responses to dynamic changes, making them ideal for complex, distributed systems. Among commonly used decentralized methods, such as auction algorithms and greedy methods, we choose the consensus-based auction algorithm for its superior global optimality over the latter [33]. The overall computational complexity is $O(N^2)$, where N is the number of USVs and targets. The detailed description is as follows:

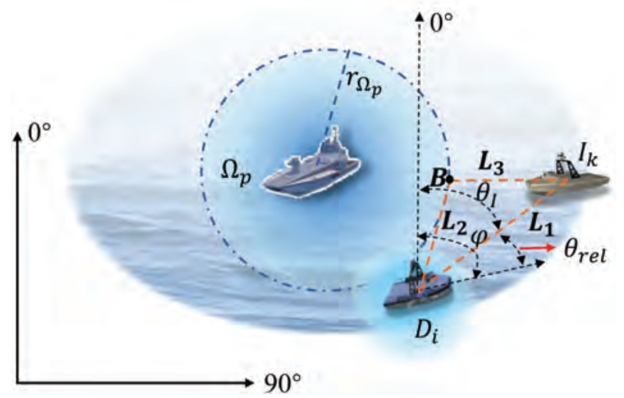


Fig. 3. Schematic diagram of the position of the USVs. (a) In the target designation process, L_1 and L_2 are used to determine the optimal pairing solution; (b) In the attack strategy, L_2 and L_3 are used to determine the optimal guidance solution.

$$\begin{aligned} \max \quad & \sum_{i=1}^{N}\sum_{k=1}^{N} a_{ik}\alpha_{ik} \\ \text{s.t.} \quad & \sum_{i=1}^{N}\alpha_{ik} = 1, \quad k = 1, \ldots, N, \\ & \sum_{k=1}^{N}\alpha_{ik} = 1, \quad i = 1, \ldots, N, \end{aligned} \tag{9}$$

where $\alpha_{ik} \in 0, 1$ denotes a decision variable, and $\alpha_{ik} = 1$ means that the I_i is assigned to the D_k for intercepting; As illustrated in Fig. 3, point B represents the ultimate interception point of D_i against I_k. The distance L_1 denotes the separation between D_i and I_k, while L_2 indicates the distance between D_i and the ultimate interception point. $a_{ik} \in \mathbb{R}^2$ denotes a cost variable, and $a_{ik} = \delta/(L_1 + L_2)$ with $\delta \in \mathbb{R}^2$. According to Eq. (9), if the sum of the distances between all matched defenders and invaders and the distance between the defender and the ultimate interception point is minimized, the objective function reaches its maximum value. For the specific target designation process, please refer to [34].

4.2. *Defense strategy*

In the previous section, we addressed the target designation problem using the auction algorithm. This section proposes an interception strategy to effectively intercept invaders. To successfully complete the interception mission, we introduce the concept of the extreme interception point, illustrated as point B in Fig. 3. Geometrically, point B represents the intersection of the line connecting the invader with the center of the Ω_p and the boundary of the Ω_p, indicating the point closest to the protected area from the invader's current position. Its definition is as follows:

$$\boldsymbol{B} = \arg\min_{\boldsymbol{B}\in\Omega_p} \|\boldsymbol{p}_i^I - \Omega_p\|. \tag{10}$$

Subsequently, we propose an adaptive interception strategy for the D_i to intercept the matched I_k:

$$\psi_{\text{esp}} = \begin{cases} \tan^{-1}(\boldsymbol{p}_i^D - \boldsymbol{p}_k^I) & \text{if } L_2 < L_3 - d_s \\ \tan^{-1}(\boldsymbol{p}_i^D - \boldsymbol{B}) & \text{if } L_2 > L_3 - d_s, \\ \tan^{-1}(\boldsymbol{p}_i^D - \boldsymbol{B}_{\text{patrol}}) & p_i^I \notin \Omega_I, \end{cases} \tag{11}$$

where ψ_{esp} is the desired heading angle, d_s is a hyperparameter, $\boldsymbol{B}_{\text{patrol}}$ is the patrol point on the Ω_p. According to Eq. (11), the defender's behavior depends on the distance between the matched invader and the defender relative to the critical interception point. If $L_2 < L_3 - d_s$, the invader is considered far from Ω_p, and the defender adopts a pure interception strategy. Conversely, if $L_2 > L_3 - d_s$, the invader is deemed to be close to Ω_p, and the defender executes an interception point strategy. If there is no matched invader, the defender performs patrol tasks until a new invader is matched. In this study, the USV executes a pure interception strategy and an interception point strategy at $v_{\max}$ and undertakes patrol missions at $\alpha v_{\max}$, where $\alpha \in (0, 1)$.

4.3. *APF-based obstacle avoidance*

In the previous section, an interception strategy was proposed. In this section, an APF-based safety avoidance mechanism is proposed for each defender to ensure collision-free behavior during the domain protection task. The overall computational complexity of APF is $O(N \times n)$, where n represents the number of interactions each defender has with targets or obstacles.

The repulsion function is expressed as follows:

$$F_{\text{rep}}^{D_i} = \begin{cases} 0 & \text{if } D_{D_i,I_i} \ge D_{\text{avo}}, \\ \mu_{\text{rep}}^D \cdot \left(\dfrac{D_{\text{avo}} - D_{D_i,I_i}}{D_{\text{avo}} - D_{\min}}\right) & \text{if } D_{\min} \le D_{D_i,I_i} < D_{\text{avo}}, \\ \mu_{\text{rep}}^D \cdot \left(\beta - \dfrac{D_{D_i,I_i}}{D_{\text{avo}}}\right) & \text{if } D_{D_i,I_i} < D_{\min}, \end{cases} \tag{12}$$

where $D_{D_i,I_i} = \|p_i^D - p_i^I\|$, $U_{\text{rep},t}^{D_i}$ refers to the repulsive field in which D_i is located at time t, while μ_{rep}^D represents the respective repulsive field coefficients. D_{avo} is a scalar that controls the influence of the repulsion field, $D_{\min}$ represents the emergency avoidance distance, β is a hyperparameter used to regulate the strength of the repulsive field in emergency scenarios. Dividing the repulsive force into multiple segments helps prevent the defender from over-maneuvering or under-reacting during obstacle avoidance, ensuring smooth transitions and making the defender's movement more natural and fluid. The direction of the repulsive force, θ_{rep}^D, is adjusted based on whether the obstacle is on the left or right side of the defender, ensuring safe obstacle avoidance. The adjustment is designed as follows:

$$\theta_{\text{rep}}^D = \begin{cases} \theta_I + \pi - \Delta\alpha & \text{if } \theta_{\text{rel}} < 0, \\ \theta_I + \pi + \Delta\alpha & \text{if } \theta_{\text{rel}} \ge 0, \end{cases} \tag{13}$$

where $\theta_{\text{rel}} = \theta_I - \psi$, as shown in Fig. 3. θ_{rel} represents the relative bearing of the target with respect to the USV, θ_I is the bearing of the target in the northeast coordinate system, and ψ is the heading angle of the USV in the same coordinate system. When $\theta_{\text{rel}} < 0$, the target is to the left of the USV; when $\theta_{\text{rel}} \ge 0$, the target is to the right. $\Delta\alpha$ is an empirically determined value obtained from real-world experiments, designed to ensure that the USV can safely and stably avoid the target. After fine-tuning, this value optimizes the obstacle avoidance strategy, enabling the USV to effectively steer clear of the target while maintaining smooth navigation and reliable task execution. The attraction function is expressed as follows:

$$F_{\text{att}}^{D_i} = \begin{cases} F_{\text{rep}}^{D_i} \cdot \dfrac{D_{D_i,I_i}}{D_{\min}} & \text{if } F_{\text{rep}}^{D_i} > 0, \\ 0 & \text{others.} \end{cases} \tag{14}$$

Here, we define the distance ratio $\frac{D_{D_i,I_i}}{D_{\min}}$, where the attraction is proportional to the distance to the target I_i. By dynamically adjusting $\frac{D_{D_i,I_i}}{D_{\min}}$, the coupling between attraction and repulsion ensures that the unmanned boat can effectively avoid obstacles while tracking targets in complex scenarios. Additionally, the presence of repulsion prevents the boat from becoming trapped in local minima, while the dynamic adjustment of attraction further enhances the convergence efficiency and overall path optimization. Therefore, the net force on defender D_i during the mission is

$$F_{\text{tot}}^{D_i} = F_{\text{att}}^{D_i} + \sum_{k=0}^{n} F_{\text{rep}_k}^{D_i}, \tag{15}$$

The direction of the resultant force can be expressed as follows:

$$\theta_{\text{tot}}^{D_i} = \arctan\left(\frac{F_{\text{tot},y}^{D_i}}{F_{\text{tot},x}^{D_i}}\right), \tag{16}$$

where $F_{\text{tot},y}^{D_i}$ and $F_{\text{tot},x}^{D_i}$ represent the components of the resultant force in the y-direction and x-direction, respectively. $\theta_{\text{tot}}^{D_i}$ will be the desired heading angle of the defender D_i, controlling the defender D_i to avoid obstacles.

5. Hierarchical Attack Strategy for the Invasion Task

This section focuses on designing an attack strategy for an invasion task. It is divided into three main aspects: (1) target point allocation based on the auction algorithm; (2) attack strategy: formation attack and angle-adaptive potential field attack; and (3) APF-based obstacle avoidance.

5.1. *Auction-based target point allocation*

In an invasion task, a stable formation can significantly improve the success rate of the invasion. To achieve this, we designed a formation attack using a leader-follower model, as shown in Fig. 4. This section utilizes the auction algorithm to assign the target position to the following USVs. First, the leader is determined based on the distance between each invader and p_{Ω_p}:

$$L = \arg\min_i \|\boldsymbol{p}_i^I - \boldsymbol{p}_{\Omega_p}\|^2, \tag{17}$$

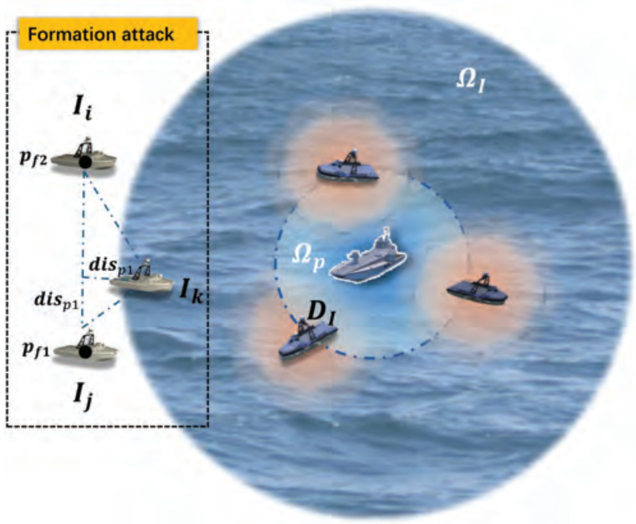


Fig. 4. Formation attack strategy.

where $\boldsymbol{p}_i^I$ refers to the position of the ith invader, $\boldsymbol{p}_{\Omega_p}$ represents the center position of the Ω_p, and finally, I_L is the selected leader. Points $\boldsymbol{p}_{f1}$ and $\boldsymbol{p}_{f2}$ are two target-following points. The role of $\boldsymbol{p}_{f1}$ is to provide follow-up targets for the invader boats other than the leader, enabling the rapid formation of the offensive formation at the start of the task. The following USV will use the auction algorithm to determine the matching target point by its relative position and relative orientation with the target point, as shown in the following equation:

$$\begin{aligned}
&\min \quad \sum_{i=1}^{N}\sum_{j=1}^{M} c_{ij}\delta_{ij} \\
&\text{such that} \quad \sum_{i=1}^{N} \delta_{ij} \le 1 \\
&\qquad \times \forall\, j \in \{1, 2, \dots, M\} \\
&\delta_{ij} \in \{0, 1\} \quad \forall\, (i, j) \in \{1, 2, \dots, N\} \\
&\qquad \times \{1, 2, \dots, M\} \\
&\delta_{ij} = 0 \quad \text{if } j \notin O_i,
\end{aligned} \tag{18}$$

where N is the number of followers, M is the number of target points, the cost function $c_{ij} = d_{ij} + \lambda\theta_{ij}$, d_{ij} represents the distance from I_i to the target point p_{fj}, δ_{ij} represents the difference between the heading angle of I_i and the azimuth angle of p_{fj}, δ_{ij} is decision variable. By calculating virtual target points through the auction algorithm, follower boats can select their optimal following positions based on the distance and heading angle difference with the target points. This ensures formation coordination and stability. Minimizing the distance and heading difference in the cost function reduces unnecessary movements and angle adjustments, thereby enhancing overall task efficiency and success rate. O_i is the set of target-following points to which I_i can be assigned and λ is the weight factor.

5.2. *Attack strategy*

The leader I_L's desired heading is directed toward the center of the protected area, and its speed is set to $v_{\max}$, as shown in (22).

$$\begin{aligned}
\psi_{\exp_L} &= \tan^{-1}(\boldsymbol{p}_L - \boldsymbol{p}_{\Omega_p}), \\
v_{\exp_L} &= v_{\max},
\end{aligned} \tag{19}$$

where $\boldsymbol{p}_L$ is the current position of the leader I_L.

The expected headings of the followers point to their corresponding target following points, and their expected

Table 1. Parameters in each chapter.

Section	Parameter
Problem formulation	$r_{\Omega_p} = 2000$ m, $r_{\Omega_I} = 5000$ m
APF-based obstacle avoidance	$D_{\text{avo}} = 500$ m, $D_{\min} = 50$, $\mu^D_{\text{rep}} = 40\sqrt{2}$, $\Delta\alpha = 10$
Auction-based target point allocation	$\lambda = 0.5$
Attack strategy	$\beta = 3, \gamma = 2, k = 0.8, \rho = 50, \epsilon = 50$

speeds are as follows:

$$v_{\text{rel}_f}\begin{cases} v_{\max} & \text{if } D_{I_j,I_L} \geq D_{I_L,p_{f_i}} + \beta, \\ v_{\max}\cdot a & \text{if } D_{I_L,p_{f_i}} \leq D_{I_j,I_L} \leq D_{I_L,p_{f_i}} + \beta, \\ -v_{\max}\cdot b & \text{if } D_{I_L,p_{f_i}} - \gamma \leq D_{I_j,I_L} \leq D_{I_L,p_{f_i}}, \\ -v_{\max} & \text{otherwise}, \end{cases} \tag{20}$$

where $a = (\frac{\text{dis}_{I_j,I_L}-\text{dis}_{I_L,p_{f_i}}}{\beta})^2$, $b = (\frac{\text{dis}_{I_j,I_L}-\text{dis}_{I_L,p_{f_i}}}{\gamma})^2$, parameters a and b are nonlinear coefficients designed to dynamically adjust speed based on the relative distance between the follower boats and the leader boat. Their purpose is to achieve smooth transitions in speed across different distance ranges. β and γ are distance parameters, which are used to adjust v_{exp} according to the distance.

To avoid interception by the defenders, this paper proposes an angle-based adaptive potential field method. This method adaptively adjusts the influence range of the potential field, R_{APF}, based on the angle between the invader's velocity vector and the line connecting the invader and the defender. The angle between the velocity vector $\boldsymbol{v}_{Ik}$ of I_k and the connecting line vector $\boldsymbol{r}_{Ik,D_i}$ is defined as ϕ:

$$\phi = \cos^{-1}\left(\frac{\boldsymbol{v}_{Ik}\cdot\boldsymbol{r}_{Ik,D_i}}{\|\boldsymbol{v}_{Ik}\|\|\boldsymbol{r}_{Ik,D_i}\|}\right), \tag{21}$$

where $\boldsymbol{r}_{Ik,D_i} = \boldsymbol{p}_{I_k} - \boldsymbol{p}_{D_i}$.

$$R_{\text{APF}} = D_{\text{avo}}(1 + k\cdot\cos\phi), \tag{22}$$

where k is the adjustment coefficient, which controls the variation of R_{APF} with ϕ. By dynamically adjusting the range of R_{APF}, the potential field method can flexibly adapt based on the relative position and directional relationship between the invader and the defender, thereby improving avoidance efficiency. The linear correction based on $\cos\phi$ is simple to compute, easy to implement, and effectively captures the relationship between velocity direction and line direction.

In this case, the repulsive force on I_i is as follows:

$$F^{I_i}_{\text{rep}} = \begin{cases} 0 & \text{if } D_{D_i,I_k} > R_{\text{APF}}, \\ \mu^I_{\text{rep}}\cdot\left(\frac{1}{D_{D_i,I_k}} - \frac{1}{R_{\text{APF}}}\right)^2 & \text{if } D_{D_i,I_k} \leq R_{\text{APF}}, \\ \rho & \text{if } D_{D_i,I_k} < D_{\min}, \end{cases} \tag{23}$$

where μ^I_{rep} represents the repulsion coefficient. When $D_{D_i,I_k} < D_{\min}$, the invader is subjected to a significantly large and constant repulsive force ρ, ensuring its ability to successfully evade interception by the defender. This mechanism provides a robust safeguard against close-range encounters, enhancing the invader's maneuverability and overall system performance in dynamic scenarios. The attractive force on I_i is as follows:

$$F^{I_i}_{\text{att}} = \begin{cases} F^{I_i}_{\text{rep}}\cdot\frac{D_{I_i,D_i}}{\epsilon} & \text{if } F^{I_i}_{\text{rep}} > 0, \\ 0 & \text{others}. \end{cases} \tag{24}$$

This design follows the same principles as (14), where ϵ is a constant representing the distance. In attack scenarios, the large spatial scale of the environment often causes the invader to experience a repulsive force from the defender that is significantly weaker than the attractive force from the central target. This force imbalance can result in the invader failing to evade interception, ultimately compromising mission effectiveness. To mitigate this issue, this study introduces the distance ratio $\frac{D_{I_i,D_i}}{\epsilon}$, facilitating the dynamic adjustment of the attractive force. This adjustment balances the interaction between target attraction and repulsive forces, thereby enhancing the invader's obstacle-avoidance capabilities and improving mission success rates.

Therefore, the net force on defender I_i during the mission is

$$F^{I_i}_{\text{tot}} = F^{I_i}_{\text{att}} + \sum_{k=0}^{n} F^{Ii}_{\text{rep}_k}. \tag{25}$$

The direction of the resultant force can be expressed as

$$\theta^{I_i}_{\text{tot}} = \arctan\left(\frac{F^{I_i}_{\text{tot},y}}{F^{I_i}_{\text{tot},x}}\right), \tag{26}$$

where $F^{I_i}_{\text{tot},y}$ and $F^{I_i}_{\text{tot},x}$ represent the components of the resultant force in the y-direction and x-direction, respectively. $\theta^{I_i}_{\text{tot}}$ will be the desired heading angle of the invader I_i, controlling the invader I_i to avoid defender.

6. Experiments

To validate the effectiveness of our framework, we conducted evaluations through simulation tests (Sec. 6.1) and real-vessel experiments (Sec. 6.2). The parameter settings for both the simulation and real-vessel environments are provided in Tables 1 and 2. To ensure generality, the initial headings of both sides were randomized within the range of $[-\pi, \pi]$, and the initial positions of the defenders and invaders were randomly placed within the circular rings defined by the radii listed in Table 2. The initial velocities of

Table 2. Parameters in the experiment.

Parameters	Defender	invader
$v_{\max}/(\text{m/s})$	12.5	12.5
$r_c/(\text{m})$	300	0
$a_{\omega,\max}/(\text{deg/s}^2)$	6	6
$a_{v,\max}/(\text{m/s}^2)$	0.6	0.6
$\hat{ph}_f$	—	100
η	—	10

all USVs were set to 0 m/s. The USVs speed is dynamically adjusted based on its distance from the target, with different velocity values assigned for varying ranges. For the domain protection task, when $D_{D_i,I_i} \geq D_{\text{avo}}$, $v_{\text{exp}} = 12\,\text{m/s}$; For $D_{\min} \leq D_{D_i,I_i} < D_{\text{avo}}$, $v_{\text{exp}} = 8\,\text{m/s}$; When $D_{D_i,I_i} < D_{\min}$, $v_{\text{exp}} = 2\,\text{m/s}$. For the domain invasion task, when $D_{D_i,I_i} \geq D_{\text{APF}}$, $v_{\text{exp}} = 12\,\text{m/s}$; For $D_{D_i,I_i} < D_{\text{APF}}$, $v_{\text{exp}} = 8\,\text{m/s}$; When $D_{D_i,I_i} < D_{\min}$, $v_{\text{exp}} = 2\,\text{m/s}$.

6.1. ***Simulation tests***

6.1.1. *Auction versus greedy*: *Task allocation efficiency*

To demonstrate the advantages of the auction algorithm, this study compares it with the greedy algorithm in terms of computation time and task allocation cost for the target designation task. As shown in Fig. 5, while the computation time of the auction algorithm increases with the number of agents, it remains within an acceptable range compared to the greedy algorithm. More importantly, in the comparison of task allocation costs, we observe that as the number of agents increases, the allocation cost of the greedy algorithm is significantly higher than that of the auction algorithm. This indicates that the allocation results of the greedy algorithm are not globally optimal, as its early local decisions may negatively impact subsequent matching processes, leading to higher overall costs. In contrast, the auction algorithm dynamically optimizes the allocation scheme among USVs through price adjustments and competitive mechanisms, effectively avoiding the local optimality issues often encountered by the greedy algorithm. Our experimental results provide clear evidence supporting the selection of the auction algorithm. In a target designation task with varying numbers of agents, the auction algorithm reduced task allocation costs by up to 36% compared to the greedy method, as shown in Fig. 5, while maintaining computation times suitable for real-time applications. These findings underline its effectiveness and justify its use over other methods in the studied context.

6.1.2. *Comparison with learning-based method*

To further evaluate the effectiveness of the proposed framework in addressing the multi-UAV defense problem, we conducted a comparison between the hierarchical defense strategy and a learning-based algorithm, MADDPG [35], in 1V1, 2V2, and 3V3 scenarios. For invader strategies, we adopted the two approaches: (1) Rule-based invaders: Invaders move toward the center of the protected area at the maximum speed $v_{\max}$. (2) Manually controlled invaders: The direction and speed of the invaders are controlled in real-time using an X-box controller. To reduce the impact of human factors, the operator is required not to escape in the direction away from the protection area. In our simulation, three evaluation metrics are used to assess the performance of the proposed method:

- Relative Winning Rate (RWR): Based on 10 rounds of confrontation tests, RWR measures the quality of defense. It represents the percentage of successful defenses across the 10 confrontations, with the calculation method

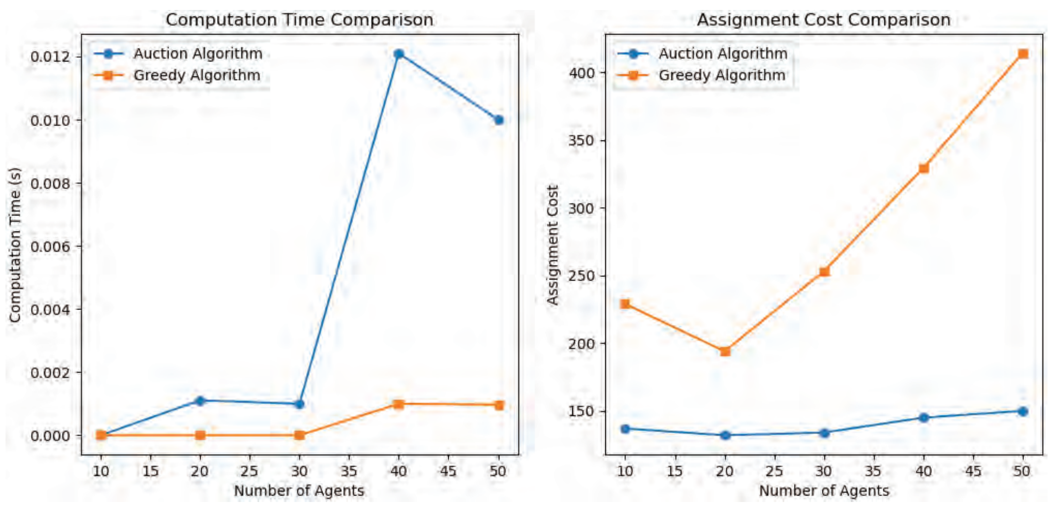


Fig. 5. Comparison of auction and greedy algorithm in computation time and task allocation cost.

detailed as follows: RWR $= W/(W + L)$ where W represents the number of wins, L represents the number of losses, $W + L = 10$.

- Average Task Completion Time (ATCT): This metric evaluates the efficiency of defense, representing the average time required to complete the defense tasks over the 10 confrontations.
- Average Minimum Distance to Invaders during Interception (AMD): This metric measures collision avoidance efficiency, representing the AMD to invaders during the 10 confrontations.

We conducted 10 rounds of testing for different invader strategies across three scenarios, with the AMD comparison results shown in Table 3. The data reveal that, regardless of the invader's strategy, our method consistently achieves a larger AMD across all scenarios. This demonstrates that our approach provides enhanced safety in defense tasks, effectively avoiding close encounters with invaders and significantly reducing collision risks.

As the scenarios scale from 1V1 to 3V3, our method continues to exhibit remarkable performance advantages. For example, in the 3V3 scenario, our method achieves a 36% improvement in AMD compared to the learning-based approach under the Rule-based invader strategy, and a 5.0% improvement under the Manually controlled invader strategy. These results fully validate the applicability and superiority of our method in multi-target and multi-agent task scenarios.

Additionally, Fig. 6 presents the comparison of RWR and ATCT across different scenarios. The data clearly show that our method demonstrates significant advantages in both key metrics. Across all scenarios, our method consistently outperforms the learning-based approach in terms of RWR, with the most significant improvement observed in complex manually controlled invader scenarios (e.g. an 80% increase in RWR in the 3V3 scenario). This highlights the exceptional dynamic adaptability and policy generalization capabilities of our approach. Furthermore, our method significantly reduces the ATCT, especially in high-dimensional scenarios

Table 3. AMD comparison across scenarios and strategies.

Scenario	Invader policy	Ours (AMD, m)	Learning-based (AMD, m)	Improvement (%)
1V1	Rule-based	**168.9**	128.3	+31.6%
	Manually controlled	**276.1**	273.8	+0.8%
2V2	Rule-based	**172.5**	132.7	+29.9%
	Manually controlled	**276.6**	268.3	+3.1%
3V3	Rule-based	**172.7**	126.7	+36.3%
	Manually controlled	**247.1**	235.3	+5.0%

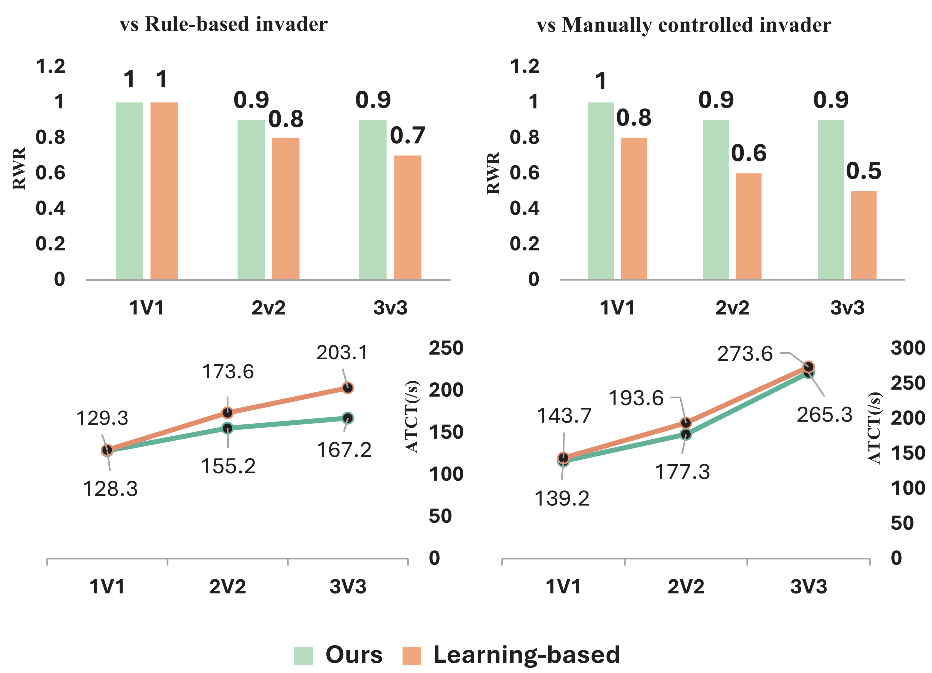


Fig. 6. Comparisons of RWR and ATCT across different scenarios.

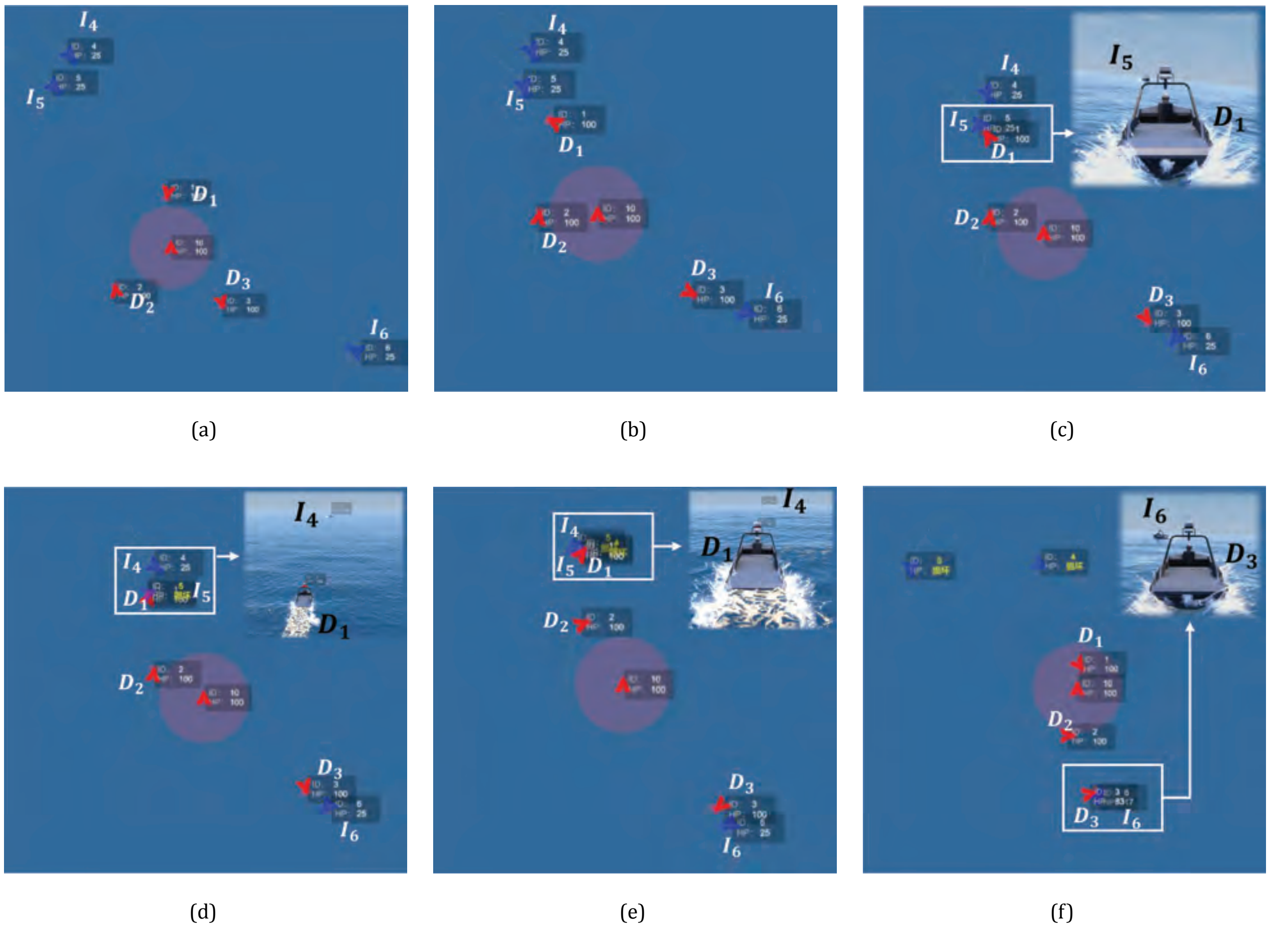


Fig. 7. The confrontation process involving three manually controlled invaders and three defenders employing the hierarchical defense strategy.

such as the 3V3 Rule-based invader strategy, where ATCT is optimized by 17.7%, showcasing improved efficiency in defense task execution. It is worth noting that the uncertainty inherent in manually controlled invader strategies poses a significant challenge to the robustness and dynamic adjustment capability of defense methods. However, leveraging its modular design and efficient dynamic task allocation mechanism, our approach performs exceptionally well even under such high-uncertainty conditions, demonstrating its reliability and practicality in complex scenarios. As the scenario complexity scales from 1V1 to 3V3, our method consistently maintains stable performance, further validating its scalability and potential for application in multi-target and multi-agent defense tasks.

In conclusion, the experimental results demonstrate that our method not only provides enhanced safety and defense efficiency but also exhibits outstanding dynamic adaptability and scalability, making it a reliable solution for addressing complex multi-target defense tasks.

To provide a more intuitive demonstration of the confrontation process, we present an example involving three manually controlled attackers and three defenders based on the Hierarchical Defense Strategy, as shown in Fig. 7. The entire confrontation is divided into six frames. Figure 7(a) illustrates the initial positions of both the attackers and defenders, which are randomly generated. The attackers move directly toward the defense area, while the defenders take actions based on the Hierarchical Defense Strategy. Figure 7(b) shows the task allocation results. Defender D_1 first intercepts Attacker I5, and the current situation is

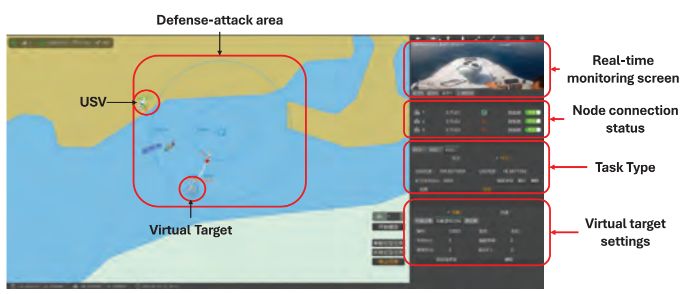


Fig. 8. Shore-based monitoring screen.

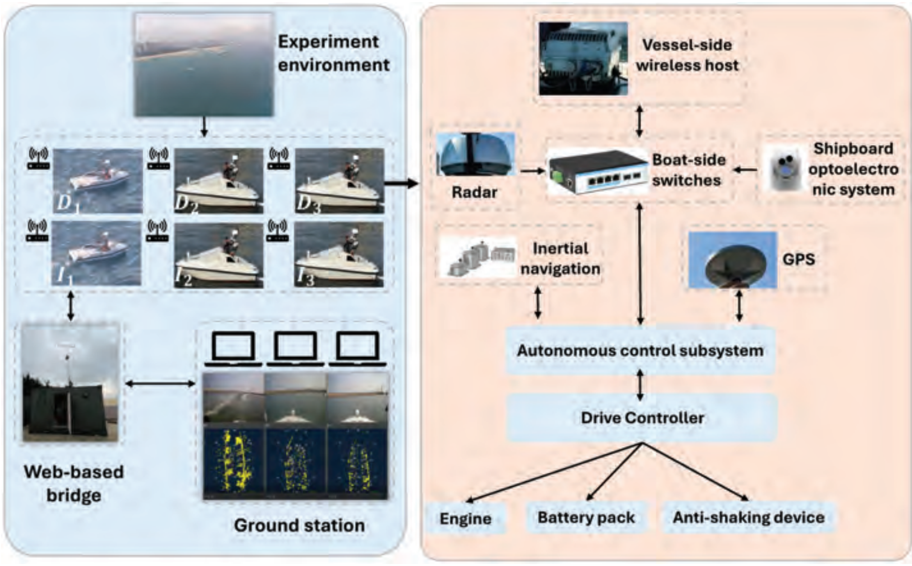


Fig. 9. Hardware structure of the attack-defense experiment platform.

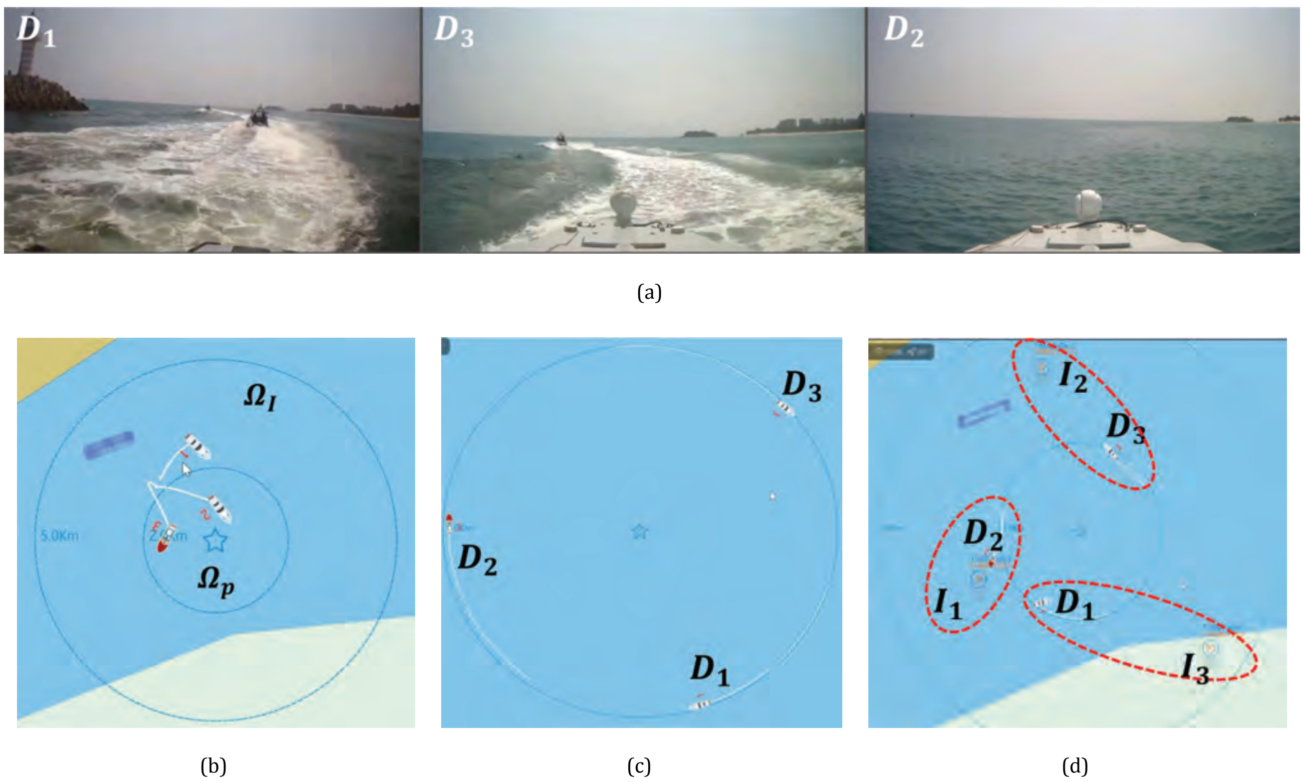
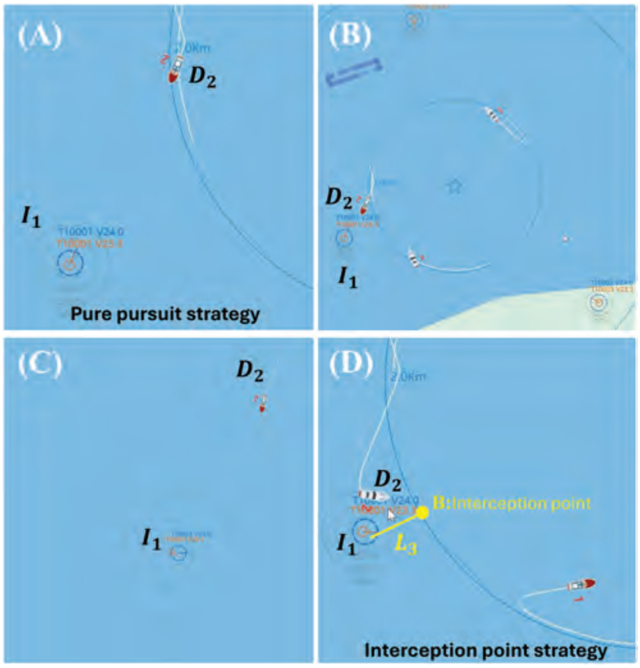
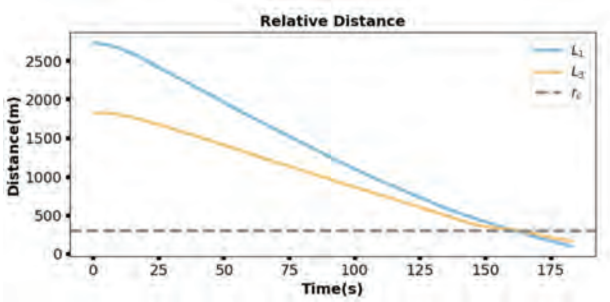


Fig. 10. The initial stage of the domain protection task. (a) Shore-based monitoring interface: three USVs were setting sail. (b) The defenders begin their patrol mission. (c) The defenders have reached the patrol point and are performing their patrol mission. (d) Target designation: The defenders are matched with the corresponding invaders.

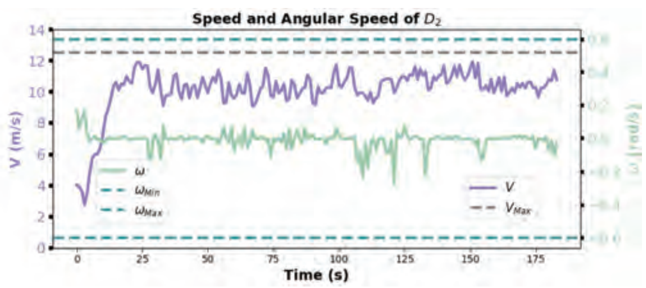
clearly visible from D_1's main view in the top-right corner of Fig. 7(c). Figure 7(d) shows that after D_1 successfully intercepts I_5, I_5 stops moving and is labeled as "damaged". Since Defender D_2 is relatively far from I_4, the interception task is reassigned to D_1. Figure 7(e) demonstrates D_1 successfully intercepting I_4. Throughout the task, Defender D_3 is consistently responsible for intercepting I_6, and Fig. 7(f) shows that D3 successfully completes the interception.



(a)

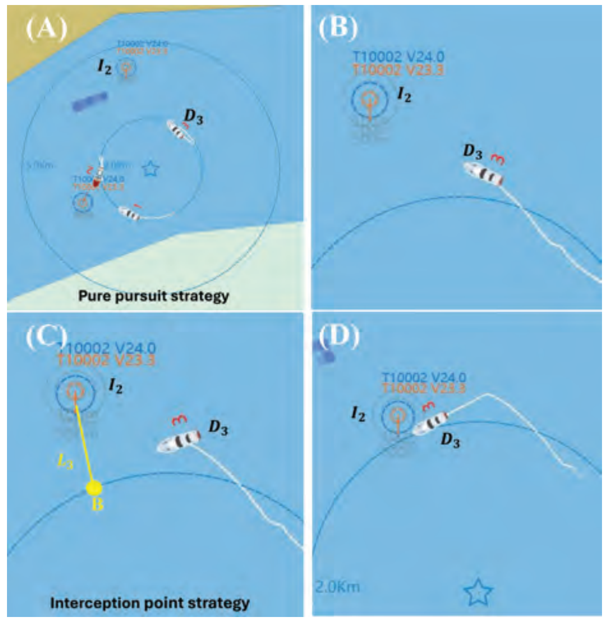


(b)

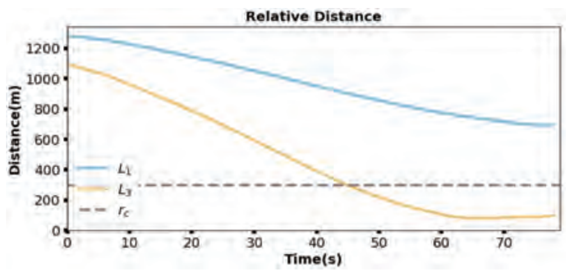


(c)

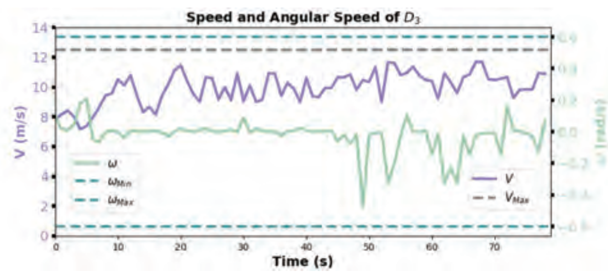
Fig. 11. The process of D_2 intercepting I_1.



(a)



(b)



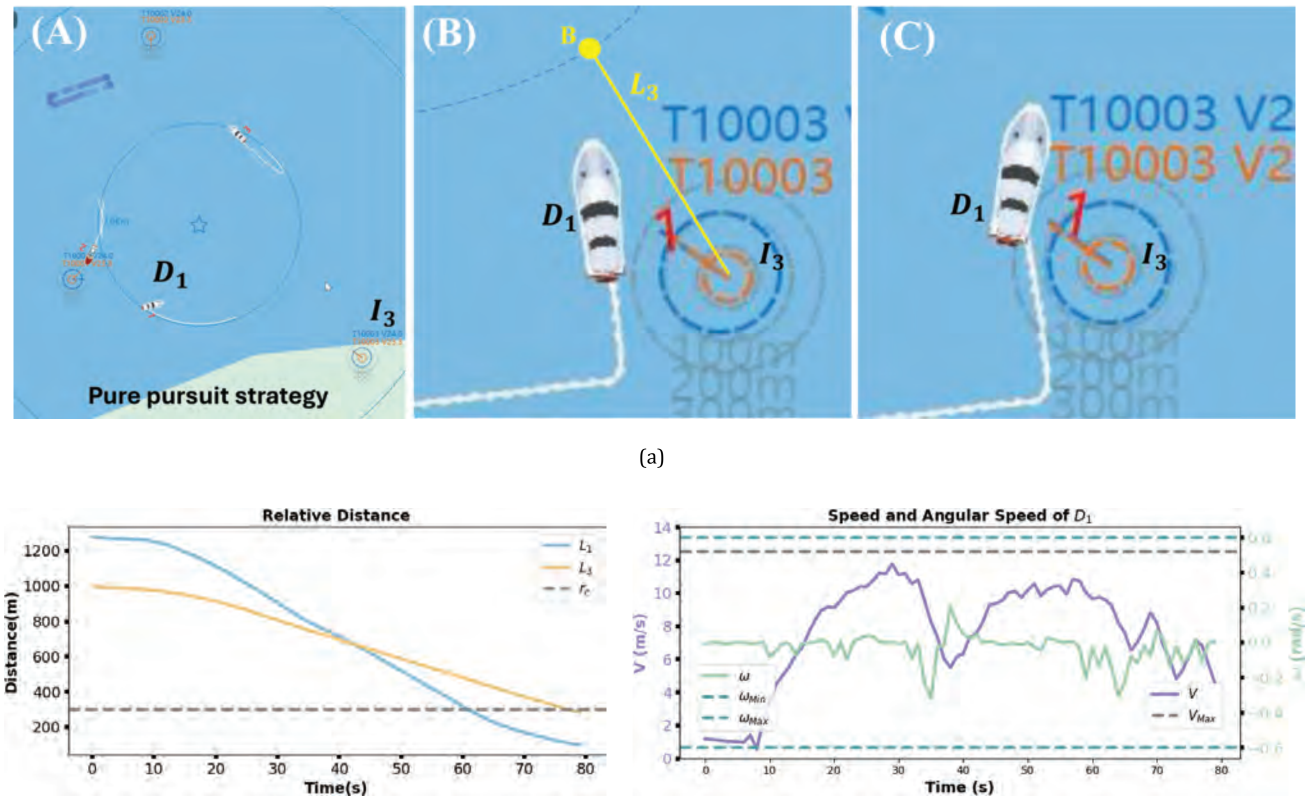
(c)

Fig. 12. The process of D_3 intercepting I_2.

This process clearly demonstrates the effectiveness of the Hierarchical Defense Strategy in dynamic task allocation and multi-target interception. Particularly when facing multiple invaders, the defenders are able to adapt flexibly and allocate tasks effectively, ensuring the efficient completion of the defense task.

6.2. *Real-vessel experiments*

This section presents experimental results to demonstrate the effectiveness of the proposed hierarchical defense



(a)

(b)

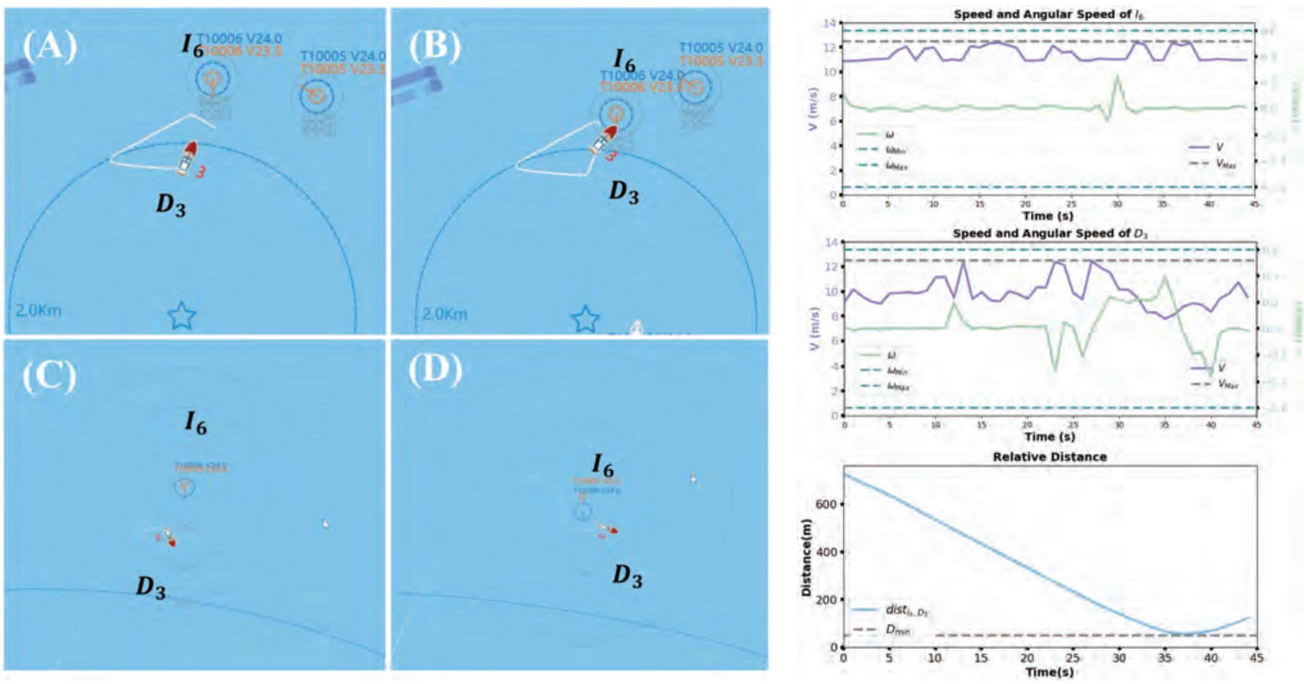
(c)

Fig. 13. The process of D_1 intercepting I_3.

strategy and hierarchical attack strategy in domain protection tasks and domain invasion tasks, respectively. The experiments were conducted in the waters of Lingshui County, Sanya City. On the day of testing, the MK III wave buoy measured an average significant wave height of 1.470 m, and the AWS2700 automatic weather station recorded an average wind speed of 8.500 m/s. The test environment was characterized by a sea state of level 4. Additionally, the vessel was equipped with a wireless ad-hoc network communication system, with a maximum communication range of 10 km. This experiment is carried out in a hybrid virtual-real approach :

(1) Domain protection task: In this task, the defenders comprised three USVs, while the attackers are virtual targets within the system as shown in Fig. 8. The defenders are controlled by a hierarchical defense strategy. Two of the invaders are controlled by a rule-based approach, while the third invader is manually controlled by shore-based personnel to breach the protected area.
(2) Domain invasion task: In this task, the defender, manually controlled, performed the interception mission. The attackers consisted of three USVs controlled by a hierarchical attack strategy to execute the invasion mission.

The experimental platform consisted of three USVs and a ground control station. Each USV is 7.5 m long, 3 m wide, and has a displacement of 3 tons. The hardware structure of the experimental platform is depicted in Fig. 9. The USVs are powered by dual engines with dual jets, enabling safe operation undersea state four conditions. They are equipped with an ad hoc communication system that supports the dynamic entry and exit of multiple USVs to form a cluster. This system also features advantages such as same-frequency node communication, network load balancing, and beyond-line-of-sight transmission, ensuring the safety of network-controlled USVs. Figure 9 illustrates the communication architecture of the experimental network system, where all communication ports use the UDP protocol. A network-based bridge was employed to establish communication between the USVs and the ground control station. Due to the large range of the experimental area, the real-time status of the USVs is displayed on a shore-based monitoring screen, as shown in Fig. 8. In the attack-defense experiment, the protection domain Ω_s and the invasion domain Ω_I are set as closed circular areas with radii of 2 km and 5 km, respectively, as shown in Fig. 8(b). The star marks the center of the protection domain, whose location is determined by the



(a) (b)

Fig. 14. APF-based obstacle avoidance verification.

shore-based personnel. To simulate realistic attack-defense scenarios and fully utilize the maneuverability of the USVs, the USVs are allowed to operate at their maximum speed $v_{\max}$.

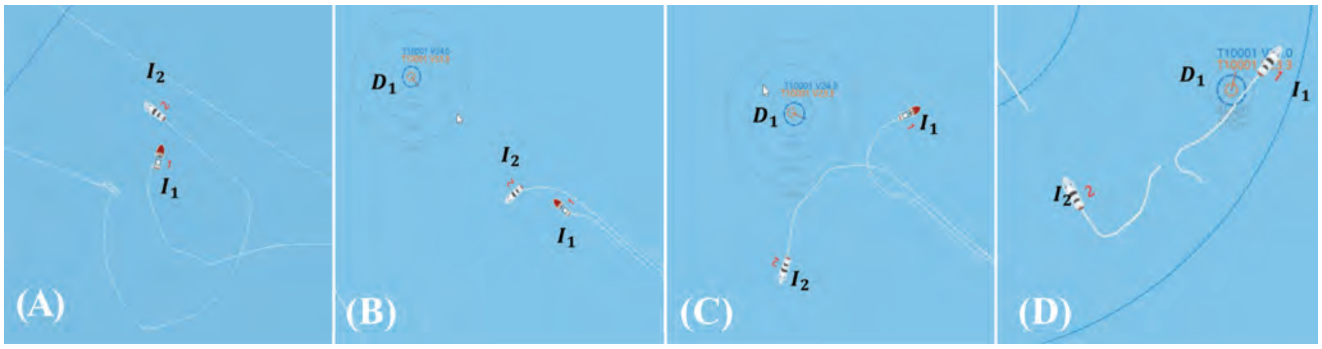
6.2.1. *The domain protection task experiment*

In this study, to ensure complete surveillance coverage of Ω_s, the patrol task is executed at three equidistant points on Ω_s, each separated by 120, as depicted in Fig. 10(c). The patrol algorithm enables defenders to effectively move along a predetermined patrol route by continuously calculating and updating the position of target points. The position of these target points is determined based on the center point, r_{Ω_p}, and the speed of USVs. The position of the target points is updated according to changes in speed and time. The proportional navigation algorithm is used to calculate the desired speed and heading. By adjusting the speed and heading of the defenders in real time, the algorithm ensures that the defenders can patrol along the planned path.

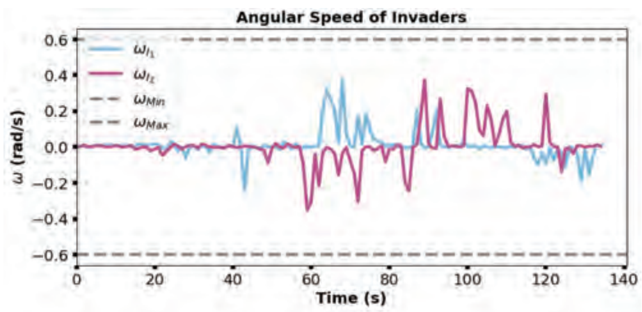
Once the patrolling state is stabilized, three invaders are randomly generated at different locations within Ω_I. To evaluate the performance of the algorithm, two kinds of invaders are adopted in this domain protection task experiment. Specifically, I_1 is a manually controlled invader, while I_2 and I_3 are rule-based invaders. The defenders are controlled based on the hierarchical defense strategy.

Figure 8(d) shows the results of the target designation after the mission begins, with each defender matched to an appropriate invader. Figure 11 illustrates the process of D_2 intercepting I_1. At the moment A shown in Fig. 11(a), $L1 < L2 + L3$, D_2 executes a pursuing strategy. Upon noticing that D_2 is approaching, I_1 changes its attack direction to move away from D_2 and closer to Ω_p, at the moment C shown in Fig. 11(a). When $L1 > L2 + L3$, at the moment D shown in Fig. 11(a), D_2 switches to the interception strategy, bringing I_1 into the capture range and ultimately completing the interception task. Figure 11(b) shows the variations of L_1 and L_3. The gray dashed line represents the capture radius r_c. When L_1 falls below this dashed line, at time = 165, the invader is within the capture range and remains within r_c thereafter. Upon successful capture of the invader, $L_3 > 0$, indicating that D_2 successfully completed the interception task before I_1 entered Ω_p. Figure 11(c) illustrates the actual speed and angular velocity of D_2 during the mission. The gray dashed line represents $v_{\max}$, and the green dashed lines represent $\omega_{\min}$ and $\omega_{\max}$. Both the speed and angular velocity remain within the constraint limits, demonstrating that the proposed guidance method is feasible under input constraints.

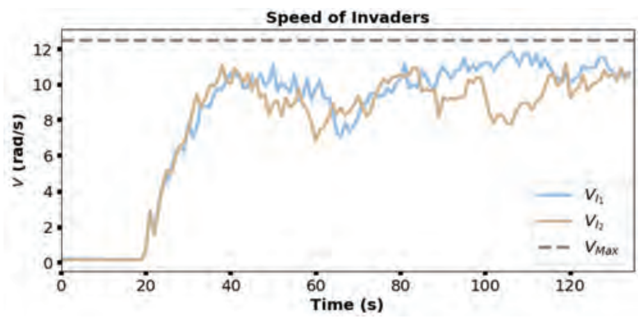
Figures 12 and 13 illustrate the interception processes of D_3 intercepting I_2 and D_1 intercepting I_3, respectively. Both D_3 and D_1 successfully completed their interception task, demonstrating the effectiveness of the proposed hierarchical defense strategy in solving the domain protection task.



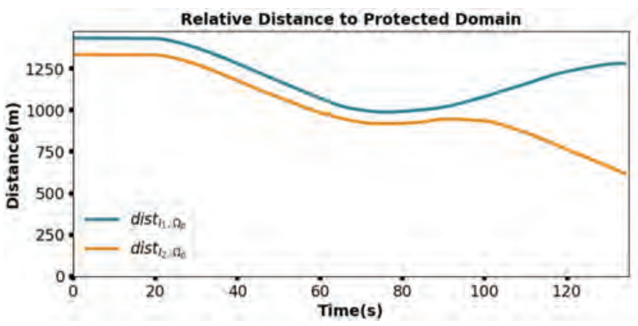
(a)



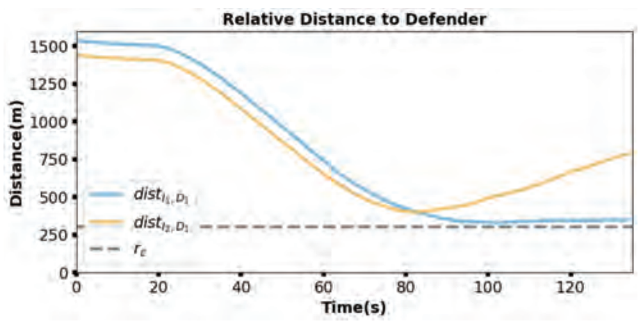
(b)



(c)



(d)



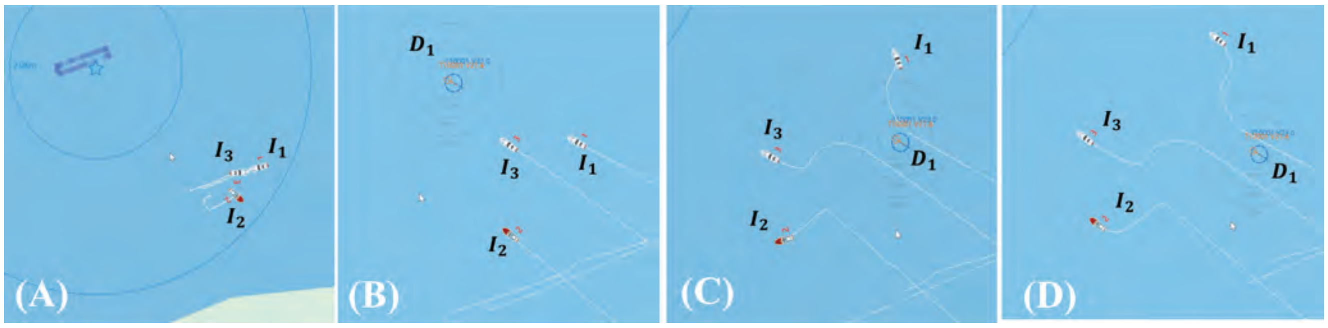
(e)

Fig. 15. Domain invasion experiment: 2 invaders invasion task.

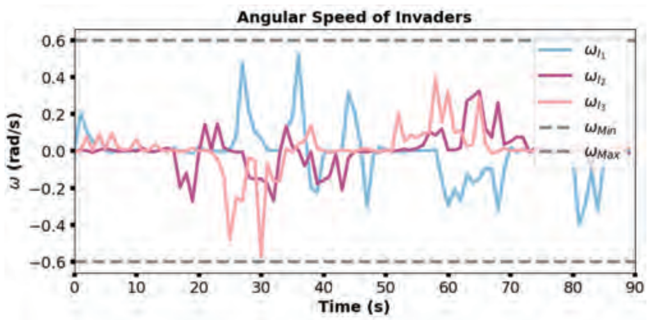
Additionally, the USV's ability to stably perform patrol and interception task at high speeds further validates the effectiveness and reliability of the navigation and control systems. This also indicates that the USVs can accurately execute predetermined task in dynamic and potentially high-risk environments, highlighting the feasibility and adaptability of the algorithm design in practical applications.

6.2.2. *APF-based obstacle avoidance verification*

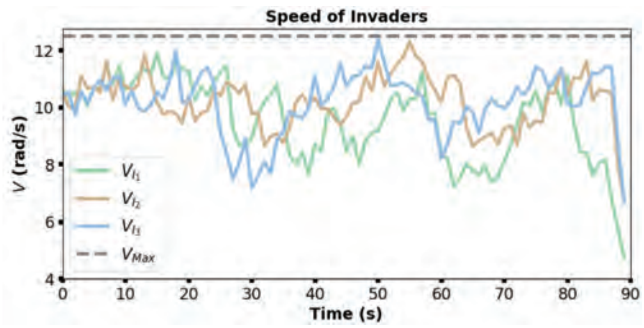
During the execution of the above-mentioned domain protection task, all defenders successfully completed their task, demonstrating the effectiveness of both the pursuing strategy and the interception strategy. Once invaders remain within the interception range for a certain period, the task is considered complete, and defenders will no longer receive information about these invaders. To simulate a real scenario where captured invaders still exist on the water surface, defenders in this experiment continued to receive information from captured invaders to validate the effectiveness of the APF-based obstacle avoidance. Figure 14 illustrates the APF-based obstacle avoidance verification process. In this scenario, after D_3 successfully intercepted I_2, a target invader, I_6, was randomly generated near D_3. The invader moved toward Ω_s at a speed of 12 m/s. Upon being



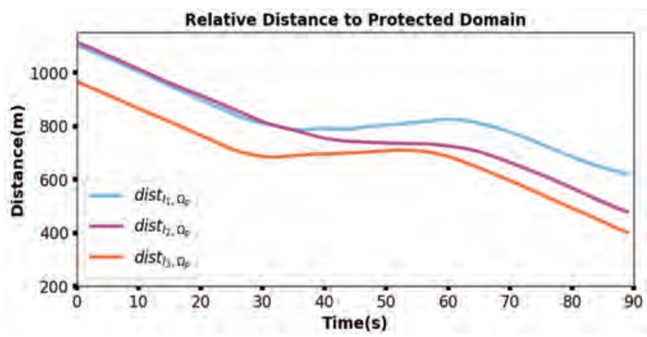
(a)



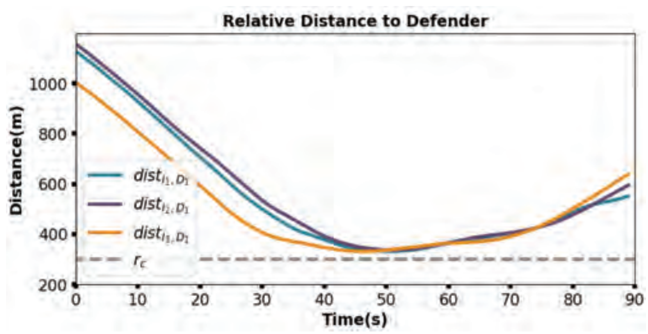
(b)



(c)



(d)



(e)

Fig. 16. Domain invasion experiment: 3 invaders invasion task.

assigned to D_3, D_3 first executed the pursuing strategy, as shown at moment A in Fig. 14(a). After successfully intercepting I_6, D_3 continued to receive I_6's state information. When $\text{dis}_{D_3,I_6} < D_{\text{avo}}$, the APF-based obstacle avoidance was triggered, as shown at moment C in Fig. 14(a). At this moment, I_6 still maintains v_{max}, D_3's speed decreased from 12.5 m/s to 8 m/s with a large turn angle, as depicted in Fig. 14(b). At the moment D in Fig. 14(a), D_3 successfully evaded I_6, ensuring $\text{dis}_{D_3,I_6} < D_{\text{min}}$, thus verifying the effectiveness of the security evasion mechanism. Throughout the process, I_6 traveled at nearly v_{max}, posing a significant challenge to the USV's maneuverability and obstacle avoidance algorithm.

6.2.3. *The domain invasion task experiment*

To validate the effectiveness of the hierarchical attack strategy, maritime tests were conducted for two scenarios: two invaders invasion task (Fig. 15(a)) and three invaders invasion task (Fig. 16(a)). In Fig. 15(a), D_1 was controlled by shore-based personnel, while I_1 and I_2 were controlled by the hierarchical attack strategy to execute the invasion task. The initial positions of the invaders at time A in Fig. 15(a) show I_2 as the leader and I_1 as the follower. When the invaders encountered interception by the defender at time B in Fig. 15(a), $\phi = 3$, and R_{APF} was relatively large, indicating a significant influence range of U^I_{rep}. At a distance of 900 m from D_1, I_2 initiated the potential field attack strategy, turning left to avoid D_1's interception. As shown in Figs. 15(b) and 15(c) at time = 58, ω_{I_2} was −0.38 rad/s, and v_{I_2} decreased from 11.5 m/s to 7 m/s. At time C in Fig. 15(a), I_1 turned right to avoid D_1's interception. By time D in Fig. 15(a), both I_1 and I_2 successfully evaded D_1's interception, with I_1 maintaining a safe distance during D_1's pursuit, as illustrated in Fig. 15(e).

Figure 16(a) illustrates the three invaders invasion task. At time A shown in Fig. 16(a), the initial positions of the three invaders are depicted. After the mission begins, the invaders execute the formation attack strategy at time B shown in Fig. 16(a). Upon encountering D_1, I_3 turns left to evade the interception area due to the resultant force, followed by I_1 and I_2, which also maneuver to avoid D_1 as indicated at time C in Fig. 16(a). The speed graph in Fig. 16(c) shows a decrease in speed during evasive maneuvers, ensuring that the invaders can avoid D_1 by turning sharply at lower speeds. Throughout the invasion mission, the invaders maintained safe distances from each other, staying above the capture radius r_c. Ultimately, the invaders successfully evaded D_1 and completed the invasion mission.

7. Conclusion

This paper addresses both domain protection and invasion task by proposing a hierarchical attack-defense strategy framework. The following conclusions were drawn from extensive experiments, including real-world maritime tests:

- The proposed hierarchical strategies proved effective in both domain protection and invasion tasks. For defense, the strategy enabled dynamic switching between patrol, pursuit, and interception, allowing USVs to maintain the integrity of the protected area. For invasion, it guided USVs to navigate complex environments and adapt to dynamic threats, successfully overcoming defenders. Additionally, obstacle avoidance was validated through simulations, showcasing the system's ability to evade re-captured invaders and maintain mission integrity with real-time speed and heading adjustments to prevent collisions.
- Advantages of the hierarchical attack-defense strategy framework: (1) Convenient deployment: It can be directly implemented in USV systems; (2) Enhanced logicality and scalability: It is applicable to multi-task scenarios, allowing for customized decision-making plans based on task requirements; (3) High efficiency: It does not require extensive data training and can adjust strategies according to environmental changes.
- The APF method, while efficient, is challenged by local minima in dynamic and complex environments. To mitigate this, we implemented segmented repulsive forces for smoother obstacle avoidance and a coupled attraction-repulsion mechanism to escape local traps. An angle-adaptive adjustment further enhances adaptability to environmental changes. These improvements strengthen robustness without compromising simplicity.
- The proposed framework demonstrates strong scalability with larger USV groups and broader operational areas. While the upper layer's auction algorithm experiences a manageable increase in computational complexity and communication overhead due to its scalable and distributed design, the middle and lower layers operate independently for each USV. This independence ensures that their computational load and communication requirements remain unaffected by the number of agents, allowing the framework to maintain efficiency and robustness.

Acknowledgment

This research was funded by the National Natural Science Foundation of China, Grant Nos. 62325302 and 62173220.

ORCID

Yangmin Xie
https://orcid.org/0000-0002-6517-2917

References

[1] Z. Peng, J. Wang, D. Wang and Q.-L. Han, An overview of recent advances in coordinated control of multiple autonomous surface vehicles, *IEEE Trans. Ind. Inf.* **17**(2) (2020) 732–745.

[2] Y. Cheng and W. Zhang, Concise deep reinforcement learning obstacle avoidance for underactuated unmanned marine vessels, *Neurocomputing* **272** (2018) 63–73.

[3] G. Wu, T. Xu, Y. Sun and J. Zhang, Review of multiple unmanned surface vessels collaborative search and hunting based on swarm intelligence, *Int. J. Adv. Robot. Syst.* **19**(2) (2022), doi:10.1177/17298806221091885.

[4] H. Wei and Y. Shi, Mpc-based motion planning and control enables smarter and safer autonomous marine vehicles: Perspectives and a tutorial survey, *IEEE/CAA J. Autom. Sin.* **10**(1) (2022) 8–24.

[5] C. Shen, Y. Shi and B. Buckham, Path-following control of an AUV: A multiobjective model predictive control approach, *IEEE Trans. Control Syst. Technol.* **27**(3) (2018) 1334–1342.

[6] R. Wang, H. Li, B. Liang, Y. Shi and D. Xu, Policy learning for nonlinear model predictive control with application to USVS, *IEEE Trans. Ind. Electron.* **71**(4) (2023) 4089–4097.

[7] G. Wen, X. Fang, J. Zhou and J. Zhou, Robust formation tracking of multiple autonomous surface vessels with individual objectives: A noncooperative game-based approach, *Control Eng. Pract.* **119** (2022) 104975.

[8] N. Gu, H. Wang, A. Wang and L. Liu, Safety-critical game-based formation control of underactuated autonomous surface vehicles, *IEEE/CAA J. Autom. Sin.* **10**(4) (2023) 1102–1104.

[9] Y. Jiang, Z. Peng, D. Wang, Y. Yin and Q.-L. Han, Cooperative target enclosing of ring-networked underactuated autonomous surface vehicles based on data-driven fuzzy predictors and extended state observers, *IEEE Trans. Fuzzy Syst.* **30**(7) (2021) 2515–2528.

[10] L. Ma, Y.-L. Wang and Q.-L. Han, Cooperative target tracking of multiple autonomous surface vehicles under switching interaction topologies, *IEEE/CAA J. Autom. Sin.* **10**(3) (2022) 673–684.

[11] T. Li, R. Zhao, C. P. Chen, L. Fang and C. Liu, Finite-time formation control of under-actuated ships using nonlinear sliding mode control, *IEEE Trans. Cybern.* **48**(11) (2018) 3243–3253.
[12] Z. Peng, N. Gu, Y. Zhang, Y. Liu, D. Wang and L. Liu, Path-guided time-varying formation control with collision avoidance and connectivity preservation of under-actuated autonomous surface vehicles subject to unknown input gains, *Ocean Eng.* **191** (2019) 106501.
[13] X. Fang, J. Zhou and G. Wen, Location game of multiple unmanned surface vessels with quantized communications, *IEEE Trans. Circuits Syst. II*: *Exp. Briefs* **69**(3) (2021) 1322–1326.
[14] B.-B. Hu, H.-T. Zhang, B. Liu, H. Meng and G. Chen, Distributed surrounding control of multiple unmanned surface vessels with varying interconnection topologies, *IEEE Trans. Control Syst. Technol.* **30**(1) (2021) 400–407.
[15] Z. Sun, H. Sun, P. Li and J. Zou, Cooperative strategy for pursuit-evasion problem in the presence of static and dynamic obstacles, *Ocean Eng.* **279** (2023) 114476.
[16] X. Qu, W. Gan, D. Song and L. Zhou, Pursuit-evasion game strategy of USV based on deep reinforcement learning in complex multi-obstacle environment, *Ocean Eng.* **273** (2023) 114016.
[17] C. Qian, X. Zhang, L. Li, M. Zhao and Y. Fang, H3e: Learning air combat with a three-level hierarchical framework embedding expert knowledge, *Expert Syst. Appl.* **245** (2024) 123084.
[18] X. Nian, M. Li, H. Wang, Y. Gong and H. Xiong, Large-scale UAV swarm confrontation based on hierarchical attention actor-critic algorithm, *Appl. Intell.* **54**(4) (2024) 3279–3294.
[19] T. Kang, N. Gu, D. Wang, L. Liu, Q. Hu and Z. Peng, Neurodynamics-based attack-defense guidance of autonomous surface vehicles against multiple attackers for domain protection, *IEEE Trans. Ind. Electron.* (2024).
[20] G. H. Burgin and A. Owens, An adaptive maneuvering logic computer program for the simulation of one-to-one air-to-air combat, Vol. 2: Program description, Technical report, NASA (1975).
[21] J. Mcmanus and K. Goodrich, Application of artificial intelligence (AI) programming techniques to tactical guidance for fighter aircraft, in *Guidance*, *Navigation and Control Conf.* (1989), p. 3525.
[22] M. Masek, C. P. Lam, L. Benke, L. Kelly and M. Papasimeon, Discovering emergent agent behaviour with evolutionary finite state machines, in *Int. Conf. Principles and Practice of Multi-agent Systems* (Springer, 2018), pp. 19–34.
[23] M. Masek, C. P. Lam, L. Kelly and M. Wong, Discovering optimal strategy in tactical combat scenarios through the evolution of behaviour trees, *Ann. Oper. Res.* **320**(2) (2023) 901–936.
[24] C. Berner *et al.*, Dota 2 with large scale deep reinforcement learning, arXiv:1912.06680.
[25] J. Perolat *et al.*, Mastering the game of stratego with model-free multiagent reinforcement learning, *Science* **378**(6623) (2022) 990–996.
[26] Q. Yang, J. Zhang, G. Shi, J. Hu and Y. Wu, Maneuver decision of UAV in short-range air combat based on deep reinforcement learning, *IEEE Access* **8** (2019) 363–378.
[27] B. Li, J. Huang, S. Bai, Z. Gan, S. Liang, N. Evgeny and S. Yao, Autonomous air combat decision-making of UAV based on parallel self-play reinforcement learning, *CAAI Trans. Intell. Technol.* **8**(1) (2023) 64–81.
[28] A. P. Pope, J. S. Ide, D. Mićović, H. Diaz, D. Rosenbluth, L. Ritholtz, J. C. Twedt, T. T. Walker, K. Alcedo and D. Javorsek, Hierarchical reinforcement learning for air-to-air combat, in *2021 Int. Conf. Unmanned Aircraft Systems* (*ICUAS*) (IEEE, 2021), pp. 275–284.
[29] P. Sun *et al.*, Tstarbots: Defeating the cheating level builtin AI in starcraft II in the full game, arXiv:1809.07193.
[30] Z. Zhang, H. Li, L. Zhang, T. Zheng, T. Zhang, X. Hao, X. Chen, M. Chen, F. Xiao and W. Zhou, Hierarchical reinforcement learning for multiagent MOBA game, arXiv:1901.08004.
[31] Z. Sun *et al.*, Multi-agent hierarchical policy gradient for air combat tactics emergence via self-play, *Eng. Appl. Artif. Intell.* **98** (2021) 104112.
[32] O. Khatib, Real-time obstacle avoidance for manipulators and mobile robots, *Int. J. Robot. Res.* **5**(1) (1986) 90–98.
[33] S. Velhal, S. Sundaram and N. Sundararajan, A decentralized multirobot spatiotemporal multitask assignment approach for perimeter defense, *IEEE Trans. Robot.* **38**(5) (2022) 3085–3096.
[34] D. Morgan, G. P. Subramanian, S.-J. Chung and F. Y. Hadaegh, Swarm assignment and trajectory optimization using variable-swarm, distributed auction assignment and sequential convex programming, *Int. J. Robot. Res.* **35**(10) (2016) 1261–1285.
[35] L. Huang, M. Fu, H. Qu, S. Wang and S. Hu, A deep reinforcement learning-based method applied for solving multi-agent defense and attack problems, *Expert Syst. Appl.* **176** (2021) 114896.

Jinduo Wang received his M.E. degrees from the University of Jimei, in 2021. He is currently working toward a Ph.D. degree in Mechanical Engineering at the Shanghai University, Shanghai, China. His current research interests include USV swarm confrontation and multi-agent systems.

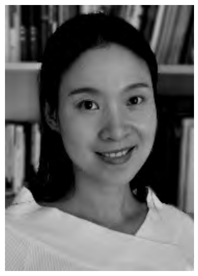
Huayan Pu received her M.S. and Ph.D. degrees in Mechatronics Engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2007 and 2011, respectively. She is currently Professor with the School of Mechatronic Engineering and Automation, at Shanghai University, Shanghai, China. Her current research interests include energy harvesting, vibration controlling, and robotics. Dr. Pu was the recipient of the Best Paper Award in Biomimetics at the 2013 IEEE International Conference on Robotics and Biomimetics. She was also nominated as the best conference paper finalist winner at the 2012 IEEE International Conference on Robotics and Biomimetics.

Zhaoxiang Shi received his B.E. degree in Mechanical Engineering from the Beijing Jiaotong University, Beijing, China in 2021. He is currently working toward an M.E. degree in Mechanical Engineering at Shanghai University, Shanghai, China. His research primarily focuses on the areas of USV swarm and imitation learning.

Zilong Song received his B.E. degree in automation from the Nanjing University of Posts and Telecommunications, Nanjing, China, in 2021. He is currently working toward an M.E. degree in Control Science and Engineering with Shanghai University, Shanghai, China. His current research interests include area coverage problems for USVs.

Heng Fu obtained a Bachelor's degree in Electrical Engineering and Automation from the Nanjing University of Technology in China in 2022. He is currently pursuing a Master's Degree at Shanghai University. His current research interests include unmanned boat collaborative search and swarm control.

Jianjun Wu received his B.E. degree in Mechanical Design, Manufacturing, and Automation from the Shanghai Maritime University, Shanghai, China, in 2017. He is currently working toward an M.E. degree in Electronic Information with Shanghai University, Shanghai, China. His current research interests include USV games and Decision control.



Yangmin Xie received her B.S. and M.S. degrees from the School of Mechanical Engineering, Beihang University, Beijing, China, and Ph.D. in Mechanical Engineering from the University of Illinois at Urbana-Champaign, Illinois, US. She is currently Full Professor with the School of Mechatronics Engineering and Automation, Shanghai University, Shanghai, China. Her current research interests include intelligent sensing and control of mobile robotics. She mainly focuses on environmental modeling, machine vision, and auto-navigation techniques for unmanned systems. Her recent work involves applications on planetary rovers, autonomous surveillance, and undersea mining.