Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# An Emergency Decision Support Framework for Flight-Deck Accidents: Integrating Dynamic Adjudication, Multi-Agent Cooperation, and Hierarchical Path Planning

**Jianwei Huang**
Naval University of Engineering
**Yue Hou**
houyue19820163.com

Naval University of Engineering

**Research Article**

**Keywords:** Flight Deck accident emergency response, Damage stability, Fire–casualty coupling, Multi-agent reinforcement learning, Hierarchical path planning, Emergency decision support

**Posted Date:** July 29th, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-10508347/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

**Additional Declarations:** The authors declare no competing interests.

# An Emergency Decision Support Framework for Flight-Deck Accidents: Integrating Dynamic Adjudication, Multi-Agent Cooperation, and Hierarchical Path Planning

Jianwei Huang [1,2], Yue Hou [1*], Kangbo Wang [1], Bo Wang[1], Jiahe Wang [1]
*1.Naval University of Engineering, Wuhan 430033, China*
*2. Maritime College, Fujian Chuanzheng Communications College, Fuzhou 350007, China;*
*Corresponding author: houyue1982@163.com

**ABSTRACT**

The flight deck of a naval vessel is a high-risk operational environment in which a single initiating event can precipitate coupled hazard cascades encompassing concurrent fire propagation, progressive casualty deterioration, and damage-induced stability degradation. Conventional pre-plan-driven response approaches are architecturally ill-suited to such multi-hazard, temporally evolving scenarios, as they cannot adapt to stochastic hazard interactions in real time. This paper presents the Intelligent Emergency Decision-support Framework for Flight-deck Safety (IEDFS), a closed-loop architecture comprising three tightly coupled modules: (1) a knowledge–data dual-driven dynamic reasoning engine that performs online fire propagation modelling, casualty condition state estimation, real-time confidence updating, and damage stability constraint evaluation; (2) an attention-enhanced multi-agent cooperative decision module in which the task-allocation problem is formalised as a Decentralised Partially Observable Markov Decision Process (Dec-POMDP) and solved via an Attention-QMIX network preserving heterogeneous agent role attributes; and (3) a hierarchical time-varying path planning module combining a temporally extended A* algorithm for global route generation with D* Lite for local dynamic replanning, enabling simultaneous hazard avoidance, stretcher-carry task scheduling, and inter-agent conflict resolution. At each decision cycle, the reasoning engine furnishes updated situational assessments to the decision module, whose cooperative policy outputs are consumed by the path planner—completing a feedback loop that sustains adaptive response under continuously evolving hazard states. Validation on the FDAS-Sim platform across five progressively escalating accident scenarios demonstrates that *IEDFS improves the critical-casualty golden-hour treatment rate by **26.3 percentage points**, reduces fire containment time by **40.1%**, decreases flight deck operational recovery time by **36.1%**, and lowers the damage stability exceedance rate by **53.2%** relative to static rule-based baselines.*Ablation studies confirm the independent contribution of each module and quantify the synergistic performance gain from system integration. These results position IEDFS as a viable methodological foundation for intelligent decision support and simulation-based proficiency training in flight deck emergency response.

Keywords: Flight Deck accident emergency response; Damage stability; Fire–casualty coupling; Multi-agent reinforcement learning; Hierarchical path planning; Emergency decision support

## Highlights

- First closed-loop intelligent emergency framework for Flight Decks integrating adjudication, cooperation, and planning.
- The dynamic adjudication engine achieves unified modeling of fire propagation, casualty condition evolution, and damage stability constraints.
- Heterogeneous team cooperative decision-making under a Dec-POMDP framework supports task coordination under partial observability.

- Hierarchical time-varying path planning simultaneously addresses hazard avoidance, stretcher transport constraints, and team conflict resolution.
- Validation across five progressively escalating scenarios demonstrates significant improvements in treatment rate, fire control efficiency, and platform safety retention.

## 1. Introduction

### 1.1. Background and Motivation

The Flight Deck supports aircraft launch and recovery, fuel replenishment, weapons loading, towing, repositioning, and continuous logistical operations. It is therefore one of the most densely populated and hazard-intensive operational areas on a naval vessel.Once a collision, hard landing, fuel leak ignition, ammunition anomaly, or structural damage occurs, the accident can rapidly evolve into a multi-physics cascading process encompassing "localized fire–fuel dispersion–secondary explosion threat–personnel injury and entrapment–passage obstruction–flooding and damage stability degradation." Research on naval vessel risk-informed design, post-fire platform recoverability, damage stability and flooding risk assessment, and real-time fire-evacuation analysis has established that fire, structural damage, personnel behavior, and residual platform safety margin exhibit significant spatiotemporal coupling [1–5].

Flight Deck accident response simultaneously confronts three categories of challenge: state uncertainty, strongly coupled tasks, and constrained resources. First, fire conditions, smoke-heat exposure, casualty status, hazard boundaries, and stability state evolve continuously in response to intervention actions, rendering static pre-plans incapable of covering all possible evolution trajectories. Second, heterogeneous forces—fire suppression and explosion containment, search-and-rescue medical, isolation and debris clearance, and stability assurance—must operate concurrently, and their equipment, personnel, and traversal paths may compete. Third, response objectives extend beyond rapid fire suppression to include critical-casualty golden-hour treatment, residual stability maintenance, Flight Deck functional recovery, and secondary accident suppression. Existing work on stability-oriented decision support, fire evacuation timing optimization, dynamic evacuation path planning, and sequential maritime disaster decision-making has individually addressed portions of this problem [6–9], yet a closed-loop method capable of unifying accident state evolution, cross-team task coordination, and dynamic path execution remains absent.

### 1.2. Related Work

#### 1.2.1 Fligh Deck Accident Emergency Response

Research on shipboard damage control and accident response has progressively advanced from experience-dependent manual handling toward knowledge-driven, model-based, and performance-based decision support. [10] constructed an early knowledge-based system for safety control of damaged ships; [11] employed knowledge models to support damage identification, action sequencing, and system reconfiguration; Choi et al. [12] proposed an onboard support method for typical damage scenarios; and Lee et al. [13] developed a performance-based onboard damage control system integrating sensing, assessment, communication, procedure recommendation, and training support. Woolley and others [2] have developed a platform recoverability assessment toolset for post-fire missions in the navy and stressed that the accident response evaluation should consider personnel survival, system damage, and mission continuity at the same time.

Damage stability analysis, progressive flooding model and fire-evacuation coupling provide the engineering basis for dynamic accident state modelling. Vassalos et al. [3] systematically reviewed ship damage stability and flooding risk assessment workflows; Rodrigues [4] surveyed approaches to modeling flooding propagation and consequences; Ruponen et al. [6] compared alternative stability analysis approaches in decision support for damaged passenger ships; and Lee and Nam [5] integrated performance-based fire

analysis with personnel evacuation for real-time shipboard response, providing a reference for online coupling of fire evolution and action decision-making.

Personnel evacuation, external search and rescue, and emergency plan selection research primarily addresses life safety and action organization under time- and resource-constrained conditions. Sarvari et al. [14] proposed an emergency evacuation decision support system for ferryboats; Xie et al. [7] analyzed the uncertainty in evacuation timing under safe-return-to-port conditions, and Xie et al. [15] optimized the issuance timing of passenger evacuation orders under ship fire scenarios; Liu et al. [8] applied an improved ant colony system to dynamic evacuation path planning on cruise ships. At the level of external rescue and integrated emergency response, Xiong et al. [16] constructed a decision support method for maritime search and rescue design and operationalization; Qiao et al. [17] combined functional resonance analysis with Bayesian networks to assess maritime liquid cargo emergency resilience; Zhao et al. [18] proposed a collaborative scheduling model for mass maritime rescue operations; Rempel [9] employed a sequential decision framework to model major maritime disaster response; and Zhang et al. [19] investigated emergency plan generation and selection from the perspectives of dynamic case-based reasoning and prospect theory. Taken together, these contributions form a reasonably complete foundation for damage control, evacuation, search and rescue, and plan selection, yet the internal accident adjudication, heterogeneous team coordination, and executable routing specific to flight decks remain typically treated in isolation.

### 1.2.2Dynamic Adjudication Rules and Accident State Modeling

Knowledge bases, rule-based inference, and scenario databases have long been employed in shipboard damage control decision-making owing to their interpretability, auditability, and compatibility with safety procedure embedding. The studies by Lee [10], Calabrese et al. [11], and Choi et al. [12] demonstrate that expert knowledge can effectively support damage state identification, action selection, and safety boundary judgment; Lee et al. [13] further coupled performance assessment results with damage control procedures so that action recommendations can be updated as sensor information evolves. Such approaches are well-suited to expressing mandatory safety constraints, but conventional rule bases are typically constructed offline and are difficult to revise continuously based on predictive performance and field feedback.

Existing accident state models can individually describe flooding, stability, fire, personnel evacuation, and emergency resilience, but two deficiencies persist for rapidly evolving flight deck accidents. First, different physical processes typically employ mutually independent time scales and state variables, precluding a unified online state transition. Second, model outputs are usually mapped unidirectionally to emergency actions, lacking a feedback mechanism through which actual intervention outcomes update adjudication rules and model parameters. A knowledge–data dual-driven dynamic adjudication engine is therefore needed to handle fire propagation, casualty deterioration, hazard boundary expansion, and damage stability constraints within a single timestep.

### 1.2.3Multi-Agent Reinforcement Learning and Cooperative Decision-Making

Multi-agent deep reinforcement learning provides an important technical foundation for distributed cooperation in complex, partially observable environments. Relevant surveys identify the core difficulties of this field as environmental non-stationarity, exponential growth of joint action spaces, credit assignment, communication constraints, and training stability [20-21]. Sunehag and his co-authors [22] put forward Value Decomposition Networks (VDN) to divide team value in cooperative tasks into a sum of individual values; Foerster et al. [23] improved multi-agent credit assignment using counterfactual baselines; and Rashid et al. [24] introduced QMIX for centralised training and decentralised execution via a monotonic mixing network.

Regarding value representation, role specialization, and relational modeling, Iqbal and Sha [25] introduced attention mechanisms into a centralized critic, enabling agents to selectively aggregate information from other agents; Yang et al. [26] proposed Qatten, capturing individual value contributions to team value via an attention structure; Wang et al. [27] proposed QPLEX to enhance joint value expressiveness under individual-global consistency; and Wang et al. [28] proposed ROMA to support agent specialization and

dynamic role assignment through emergent roles. In emergency applications, Yang et al. [35] validated the potential of reinforcement learning for multi-agent cooperative evacuation.

These approaches are primarily validated on standard benchmarks, traffic systems, or general cooperative environments, and lack joint modeling of heterogeneous team capability differences, communication constraints, stability safety boundaries, dynamic task priority changes, and path executability specific to shipboard accidents. A type-aware attention mechanism within the value decomposition framework is therefore required to explicitly encode capability parameters, local observations, and cooperative relationships of different teams, achieving task coordination under partial observability through centralized training–decentralized execution.

#### 1.2.4Dynamic Path Planning and Multi-Agent Navigation

Dynamic path planning determines whether higher-level task plans can be executed safely in rapidly changing accident environments. Sharon et al. [29] proposed Conflict-Based Search (CBS), obtaining optimal conflict-free paths through high-level constraint search and low-level single-agent planning; Stern et al. [30] unified the fundamental definitions, objective functions, and benchmarks of multi-agent path planning; and Ma et al. [31] proposed prioritized planning with consistent priority ordering to reduce solution failures while maintaining high solution quality. These works provide a systematic algorithmic foundation for multi-team conflict resolution.

Li et al. [32] have developed an anytime multi-agent path planning method for large-scale and more dynamic environments by means of a large neighborhood search that continuously improves solutions over time; Li et al. [33] proposed EECBS to significantly accelerate multi-agent path solving with bounded suboptimality guarantees; and Ren et al. [34] proposed multi-objective Path-Based D* Lite, which accelerates incremental replanning after environmental changes by reusing prior search results. Although these algorithms handle map changes, scale expansion, and team conflicts, flight deck accidents further require consideration of smoke-heat and fuel exposure, explosion risk, stretcher transport width and turning constraints, heavy equipment passage, temporary closures, and the influence of heel and trim on path executability.

#### 1.2.5Positioning of the Present Work

The preceding review shows that existing research has independently produced mature methods in shipboard damage control decision support, accident state modeling, multi-agent cooperation, and dynamic path planning. However, three critical gaps persist in cross-layer integration: first, the absence of an online adjudication mechanism that couples fire, casualties, hazard dispersion, and damage stability within a unified timestep; second, the lack of a partially observable cooperative decision model tailored to heterogeneous teams performing fire suppression and explosion containment, search-and-rescue medical, isolation and debris clearance, and stability assurance tasks; and third, the absence of a closed-loop system linking upper-level task allocation with lower-level time-varying path execution, conflict resolution, and feedback correction. The contribution of this paper is not to replace existing physical models or individual algorithms, but to organize them through modular interfaces into an online-executable, interpretable, and extensible emergency decision support framework.

### 1.3. Research Gaps and Contributions

The foregoing literature demonstrates that knowledge-based damage control systems and performance-based analysis can provide auditable accident state judgment and action recommendations [11] [13]; value decomposition, role learning, and attention mechanisms can support multi-agent cooperation under partial observability [24-28]; and conflict search, bounded suboptimal planning, and incremental search can improve path solving efficiency in dynamic environments [29][32-34]. When directly applied to flight decks, however, these methods must additionally address unified representation of multi-physics states, heterogeneous team

capability constraints, life–platform–mission multi-objective trade-offs, and bidirectional feedback between decision-making and execution.

Based on the above analysis, this paper proposes an Intelligent Emergency Decision-support Framework for Flight-deck Safety (IEDFS) for flight deck accident response. The framework integrates a dynamic adjudication engine driven by both expert knowledge and data for dynamic adjudication rule engine, attention-enhanced multi-agent reinforcement learning, and hierarchical time-varying path planning to form a closed-loop decision chain encompassing "accident state evolution–task cooperative allocation–dynamic path execution–outcome feedback correction."

The principal contributions are as follows:

Knowledge–data dual-driven dynamic adjudication engine: An online-updated rule framework is constructed for flight deck accident scenarios, unifying the modeling of fire propagation, casualty condition evolution, hazard source dispersion, and damage stability constraints, enabling real-time accident state adjudication, intervention effectiveness assessment, and rule confidence updating.

Heterogeneous multi-agent cooperative decision model: flight deck accident emergency response is formulated as a Dec-POMDP, and an Attention-enhanced value decomposition network (Attention-QMIX) is designed to support cooperative optimization among heterogeneous teams—fire suppression and explosion containment, search-and-rescue medical, isolation and debris clearance, and stability assurance—under partially observable conditions.

Hierarchical Time-varying Path Planning Method: Differentiated Time-varying Edge Cost Functions They are for different kinds of problems, combine time-extended A* global planning with D* Lite local replanning, and support joint handling of hazard-avoidance stretcher transport constraints, team conflict resolution, and rapid replanning after environmental changes.

Closed-loop integrated verification platform: Five successively increasing scenarios are designed Within a flight deck accident simulation environment, systematically quantify the overall system performance, the contribution of individual modules, and the synergistic gains achieved by module integration.

## 2. Materials and Methods

### 2.1. System Architecture and Problem Formulation

#### 2.1.1System Overview

The five modules of the Intelligent Emergency Decision-support Framework (IEDFS) in this proposal operate in a discrete-time closed loop, as shown in Figure 1.

The scenario management module initialises the flight deck topological graph, carrier-based aircraft and hazard source distributions, team compositions, sea state parameters, and communication constraints. Dynamically update fire spread, casualty changes, hazard expansion and damage stability over time in the adjudication engine. The module of multi-agent cooperative decision-making generates a task allocation strategy based on local observations and estimated global state. The top-down Path planning module converts abstract task objectives into concrete, executable Paths. The visualisation and Analysis Module Support accident simulation playback and evaluation metric calculation.

The modules are connected in a closed loop as follows: at each timestep t, the adjudication engine computes the updated state s(t+1) according to the current environment state s(t) and the previous timestep's executed actions a(t−1); the cooperative decision-making module outputs joint actions a(t) based on local observations $o_i(t)$; the path planning module converts a(t) into executable paths $p_i(t)$ and feeds execution feedback back to the adjudication engine.

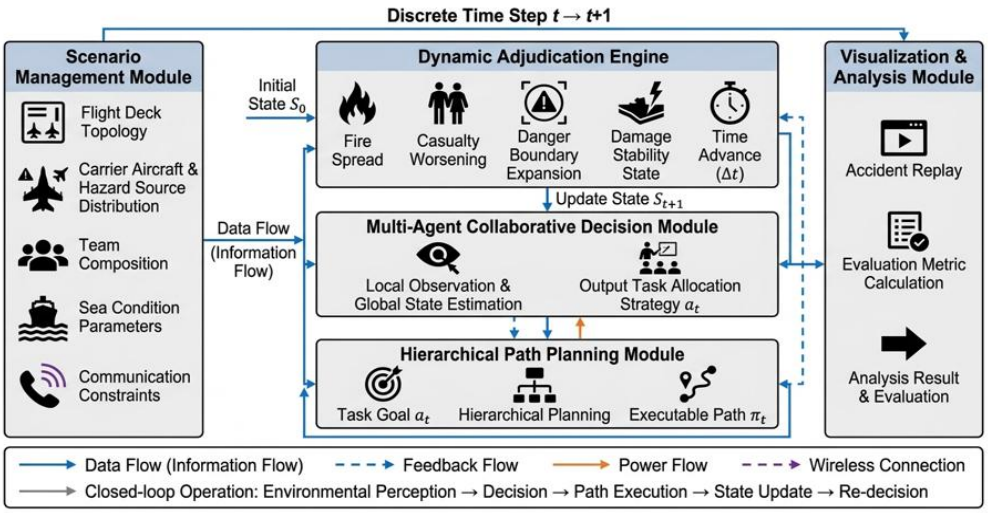



***Figure1.*** *Functional module diagram of the Intelligent Emergency Decision-support Framework (IEDFS).*

2.1.2 Flight Deck Graph Model

The flight deck and its adjacent critical areas are modeled as a time-varying weighted graph:

$$G(t) = \big(V, E, W(t)\big) \tag{1}$$

where: $V$ denotes the set of nodes, $E$ denotes the set of edges, and $W(t)$ denotes the time-varying edge cost matrix.

The node set $V$ encompasses all functional zones of the flight deck (parking areas, launch/recovery zones, ammunition areas, fuel zones, fire stations, deck-edge safety zones), vertical passage entries (elevator openings, hatches, stairwells), hangar key nodes, and medical evacuation points, totaling $|V| = N_v$ nodes.

The state vector of each node v at time t is:

$$d_v(t) = \left[d_v^{\text{fire}}(t),\ d_v^{\text{inj}}(t),\ d_v^{\text{haz}}(t),\ d_v^{\text{deb}}(t),\ d_v^{\text{acc}}(t)\right]^{\mathsf{T}} \tag{2}$$

where: $d_v^{\text{fire}}$ represents fire intensity index, $d_v^{\text{inj}}$ represents a composite indicator of casualty count and injury severity, $d_v^{\text{haz}}$ represents hazard source exposure level, $d_v^{\text{deb}}$ represents debris obstruction degree, and $d_v^{\text{acc}}$ represents area trafficability, respectively.

The global platform state vector is:

$$g(t) = \left[\phi(t),\ \theta(t),\ \text{GM}(t),\ \Delta_{\text{free}}(t),\ R_{\text{ops}}(t)\right]^{\mathsf{T}} \tag{3}$$

where: $\phi$ represents heel angle, $\theta$ represents trim angle, $GM$ represents residual metacentric height, $\Delta_{free}$ represents free surface correction, and $R_{ops}$ represents flight deck mission recovery capability index, respectively.

The time-varying edge cost function is defined as:

$$w(e,t) = w_{\text{len}}(e) + \alpha \cdot w_{\text{haz}}(e,t) + \beta \cdot w_{\text{trans}}(e,t) + \gamma \cdot w_{\text{cong}}(e,t) + \eta \cdot w_{\text{stab}}(t) \tag{4}$$

where: $w_{\text{len}}$ is the geometric transit time; $w_{\text{haz}}$ is the composite hazard cost incorporating fire zone, smoke-heat, fuel exposure, and explosion risk; $w_{\text{trans}}$ is the special constraint cost for stretcher transport and heavy equipment passage; $w_{\text{cong}}$ is the congestion conflict cost arising from concurrent multi-team operations; $w_{\text{stab}}$ is the influence term of platform attitude and stability margin changes on path executability; and $\alpha, \beta, \gamma, \eta \geq 0$ are weighting coefficients initialized through expert calibration and subject to adaptive adjustment through validation.

2.1.3Accident State Evolution Model

The complete accident state of the flight deck consists of node-level local states and the global platform state:

$$S(t) = \{d_v(t)\}_{v \in V} \cup g(t) \tag{5}$$

Fire sub-model: The fire intensity update at node v follows:

$$d_v^{\text{fire}}(t+1) = d_v^{\text{fire}}(t) + \sum_{u \in \mathcal{N}(v)} \lambda_{uv}(t) \cdot d_u^{\text{fire}}(t) - \mu_v^{\text{fire}}(t) \cdot a_v^{\text{sup}}(t) + \epsilon_v^{\text{fire}}(t) \tag{6}$$

where $\lambda_{uv}(t)$ is the heat propagation coefficient from node u to v (influenced by fuel distribution, ventilation state, and barrier effectiveness), $\mu_v^{\text{fire}}(t)$ is the fire suppression intervention efficiency coefficient, $a_v^{\text{sup}}(t)$ is the fire suppression resource input at the current time step, and $\epsilon_v^{\text{fire}}$ is the stochastic disturbance term.

Casualty condition sub-model,The composite injury index update for zone v is:

$$d_v^{\text{inj}}(t+1) = d_v^{\text{inj}}(t) + \delta_v^{\text{env}}(t) \cdot \left[d_v^{\text{fire}}(t) + d_v^{\text{haz}}(t)\right] - \mu_v^{\text{med}}(t) \cdot a_v^{\text{med}}(t) \tag{7}$$

where $\delta_v^{\text{env}}$ is the acceleration coefficient of environmental hazards on injury deterioration, $\mu_v^{\text{med}}(t)$ is the medical treatment efficiency, and $a_v^{\text{med}}$ is the medical resource input.

Stability loss sub-model: The evolution of residual metacentric height is influenced by fire-fighting water accumulation Δw(t), structural damage ΔK(t), and sea state disturbance ξ(t):

$$\text{GM}(t+1) = \text{GM}(t) - k_w \cdot \Delta w(t) - k_d \cdot \Delta K(t) + \xi(t) \tag{8}$$

The stability safety constraints are defined as $GM(t) \geq GM_{\min}$ and $|\varphi(t)| \leq \varphi_{\max}$; violation triggers a stability assurance priority dispatch mechanism.

2.1.4 Heterogeneous Emergency Teams and Action Space

Considering the task characteristics of flight deck accident response, five types of heterogeneous agents are defined:

$$\mathcal{A} = \left\{\mathcal{A}^{\text{cmd}}, \mathcal{A}^{\text{fire}}, \mathcal{A}^{\text{med}}, \mathcal{A}^{\text{iso}}, \mathcal{A}^{\text{stab}}\right\} \tag{9}$$

corresponding to the incident command team, fire suppression and explosion containment team, search-and-rescue medical team, isolation and debris clearance team, and stability assurance team, respectively, as shown in Table 1.

Each agent's action space comprises target node selection $v^* \in V$, response mode selection $m \in M_i$, path following $p_i$, and task switching. Heterogeneous team attributes are explicitly preserved in both policy learning and path planning.

| Team Type | Primary Function | Core Capability Parameters | Primary Action Set |
|---|---|---|---|
| Incident Command $\mathcal{A}^{\text{cmd}}$ | Priority adjustment, resource coordination | Global situational awareness radius | Task dispatch, priority switching, information sharing |
| Fire Suppression$\mathcal{A}^{\text{fire}}$ | Foam coverage, boundary cooling | Suppressant capacity, heat resistance rating | Zone selection, coverage/cooling/containment |
| Search-and-Rescue Medical $\mathcal{A}^{\text{med}}$ | Casualty search, on-site treatment, evacuation | Carrying capacity, medical equipment level | Search, triage, primary treatment, stretcher evacuation |
| Isolation and Clearance $\mathcal{A}^{\text{iso}}$ | Hazardous material isolation, debris removal | CBRN protection level, clearance equipment | Isolation, clearance, passage restoration |
| Stability Assurance $\mathcal{A}^{\text{stab}}$ | Liquid accumulation control, attitude monitoring | Drainage capacity, monitoring precision | Drainage organization, ballast adjustment, attitude warning |

Table 1. Attributes of five heterogeneous emergency agent types.

2.1.5Overall Optimization Objective

The overall objective of flight deck accident response is formulated as the following multi-objective optimization problem:

$$\max_{\pi} J(\pi) = w_1 J_{\text{surv}} + w_2 J_{\text{fire}} + w_3 J_{\text{ops}} + w_4 J_{\text{stab}} - w_5 C_{\text{time}} - w_6 C_{\text{risk}} \tag{10}$$

where: $J_{\text{surv}}$ is the critical-casualty golden-hour treatment rate; $J_{\text{fire}}$ is the fire containment benefit; $J_{\text{ops}}$is the flight deck operational recovery benefit; $J_{\text{stab}}$is the stability safety maintenance benefit; $C_{\text{time}}$ is the response time cost; $C_{\text{risk}}$ is the penalty for secondary accidents and safety boundary violations; and $w_1, \ldots, w_6$ are objective weight coefficients initialized through expert assignment.

2.2. Dynamic Adjudication Rule Engine

2.2.1Rule Tuple Structure

Each adjudication rule is represented as a unified tuple:

$$r = (\text{cond},\ \text{act},\ \text{eff},\ c_r,\ \text{meta}) \tag{11}$$

where *cond* denotes the trigger condition (which may include both precise and fuzzy conditions); *act* the corresponding emergency action; *eff* the expected effect distribution of the action; $\rho \in [0,1]$ the current confidence level; and *meta* records the source annotation, activation count, and historical effectiveness.

2.2.2Rule Initialization: Knowledge–Data Dual-Drive

Rule initialization comes from two places:

Knowledge-driven component: The initial rules are extracted from flight deck accident response pre-plans, shipboard firefighting and medical emergency regulations, expert interviews and historical exercise experience. The above sources constitute the initial rule base $R_0$, which includes basic scenarios such as fire detection and first response, hazardous material isolation procedures, casualty triage and evacuation workflows, and stability exceedance warnings and countermeasures.

Data-driven component: $A_{priori}$ association rule mining and sequential pattern analysis are applied to historical simulation exercise samples and training records to automatically discover accident evolution regularities and intervention effect patterns not explicitly covered by pre-plans, forming the extended rule set $R_{\text{ext.}}$ The complete initial rule base is:

$$\mathcal{R} = \mathcal{R}_0 \cup \mathcal{R}_{\text{ext}} \tag{12}$$

2.2.3Online Confidence Update

During system operation, the decision engine predicts the deviation between the predicted outcome based on the rules and the actual observation, and continuously updates the confidence level $c_r$ (t) ∈ [0,1] for each rule. The outcome matching signal is defined as

$$g_r(t) = \exp\left(-\frac{D\left(\widehat{\text{eff}}_r(t), \text{eff}_r^{\text{obs}}(t)\right)}{\delta}\right) \tag{13}$$

where $\widehat{\text{eff}}_r(t)$ is the predicted distribution of intervention effects following the triggering of rule r, and $\text{eff}_r^{\text{obs}}$ is the empirical distribution of observed effects within the window; $D\left(\widehat{\text{eff}}_r(t), \text{eff}_r^{\text{obs}}(t)\right)$ is the Wasserstein distance between the two, and δ > 0 is the tolerance threshold. Confidence updates employ first-order exponential contraction towards$g_r(t)$:

$$c_r(t+1) = c_r(t) + \eta_r\big(g_r(t) - c_r(t)\big) \tag{14}$$

where $\eta_r \in (0,1]$ is the rule-level learning rate. Since $c_r(t)$, $g_r(t)$and $\eta_r$ all lie within [0, 1], the update forms a convex combination; $c_r(t)$ is thus always confined to [0, 1], requiring no additional clipping. $g_r(t)$depends solely on the magnitude of the deviation and is independent of its direction, consistent with the semantic interpretation of confidence as a measure of monotonic reliability. When the confidence of rule *r* remains below the threshold $T_{retire}$ consecutive time steps, the rule is placed in a dormant state and removed from the current set of adjudication candidates, whilst remaining in the rule library; Dormant rules are re-evaluated every $T_{check}$ steps based on their performance in the most recent window; if $g_r(t)$ rises above $c_{min}$, the rule is reactivated to prevent permanent deactivation caused by short-term abnormal operating conditions. When the data-driven module identifies a new pattern, it automatically adds the corresponding rule, with its initial confidence set to the median confidence $c_{init}$ of rules in the library for the same scenario, after which it is incorporated into the update process described above.

2.2.4Coupled Adjudication Process

At each simulation timestep, the adjudication engine updates the global accident state according to the procedure illustrated in Fig. 2.

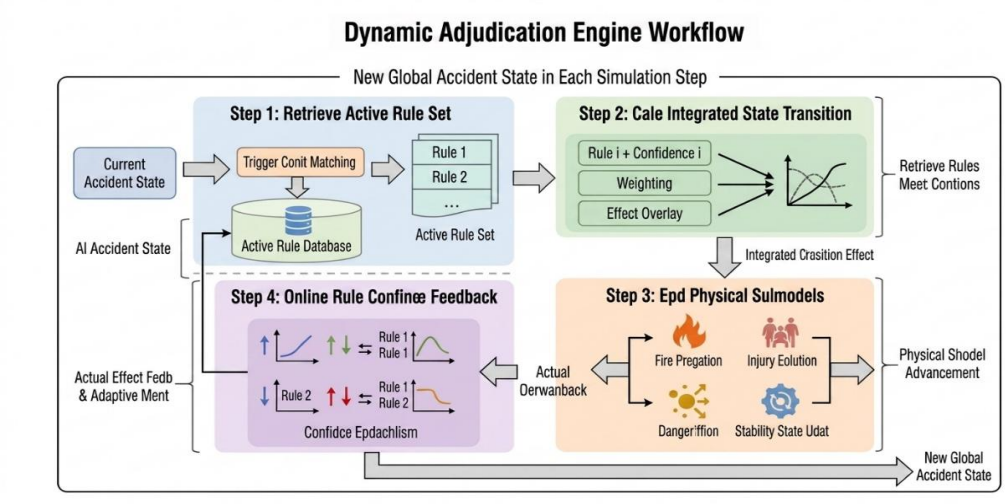



***Figure 2.*** *Dynamic adjudication engine workflow diagram.*

Step 1: Retrieve the active rule set $R_active(t) \subseteq R$ satisfying the trigger conditions of the current accident state $S(t)$.

Step 2: Compute the composite state transition by confidence-weighted superposition of intervention effects from all active rules.

Step 3: Execute physical sub-model advancement, updating fire, casualty, hazard dispersion, and stability states.

Step 4: Feed actual observations back to the confidence update mechanism.

This mechanism ensures that the adjudication engine maintains adherence to engineering knowledge and safety constraints while adaptively adjusting based on actual effects in complex, atypical accidents—preventing static rule base failure under extreme conditions.

## 2.3. Attention-Enhanced Multi-Agent Reinforcement Learning for Cooperative Decision-Making

### 2.3.1 Dec-POMDP Formulation

Flight Deck accident emergency response is modeled as an $n$-agent Decentralized Partially Observable Markov Decision Process (Dec-POMDP), defined as the seven-tuple:

$$\mathcal{M} = \langle \mathcal{S}, \mathcal{A}_1, \ldots, \mathcal{A}_n, P, R, \mathcal{O}_1, \ldots, \mathcal{O}_n \rangle \tag{15}$$

where $S$ is the global state space, $A_i$ is the action space of agent $i$, P: $S \times A \times S \rightarrow [0,1]$ is the state transition probability, R is the shared reward function, and $O_i$ is the local observation space of agent $i$. The local observation $o_i(t) \in O_i$ of agent $i$ covers only the neighborhood state within perception radius $r_i$, including: fire intensity of visible nodes, casualty positions and injury severity, hazard boundaries, passage trafficability, and positions of other teams (if within perception range). The global state $S(t)$ is accessible only during training for joint value function estimation; during execution, each agent relies solely on local observations and communication messages for decision-making.

### 2.3.2 Reward Function Design

To illustrate the reward structure guiding multi-agent policy optimization, Fig. 3 presents the average per-episode contribution of each reward component after training for the IEDFS. Incentives for Casualty Rescue ($R_1$) and stability safety incentive ($R_4$) account for the majority of the total reward due to higher weighting coefficients and the operational concept of "life first, platform integrity first". The per-step time penalty ($P_1$) continuously imposes efficiency pressure, and the safety violation penalty ($P_2$) strongly restricts crossing the stability boundary.

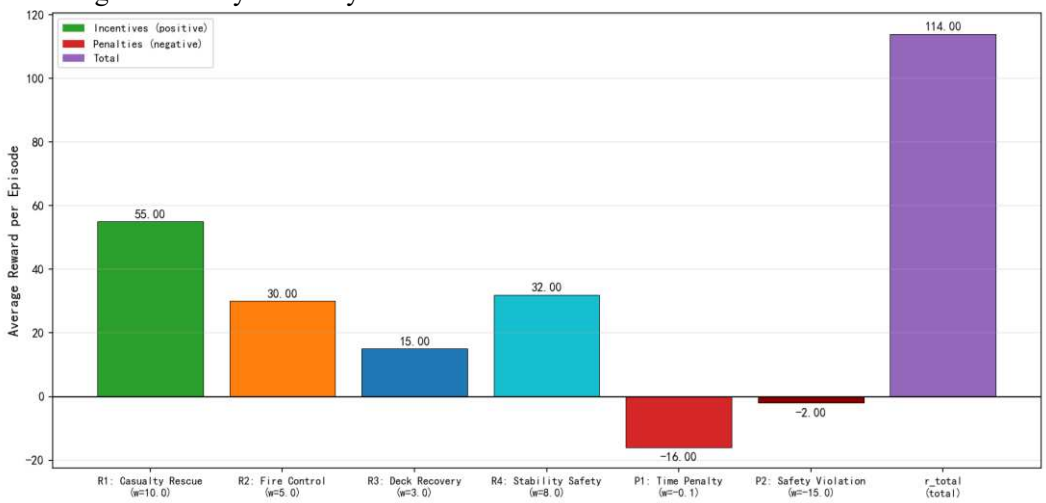



**Figure 3.** Average reward component contributions per episode for IEDFS.

Figure 4 shows the breakdown of the per-step reward in a typical episode for scenario S3. The decomposition shows that the learned policy naturally develops phased cooperative patterns: early fire suppression actions yield higher rewards; after controlling the fire, the contributions of casualty evacuation and stability maintenance gradually increase. This phase alignment emerged without explicit temporal reward shaping; thus, the composite reward design successfully encoded the desired multi-objective priority order.

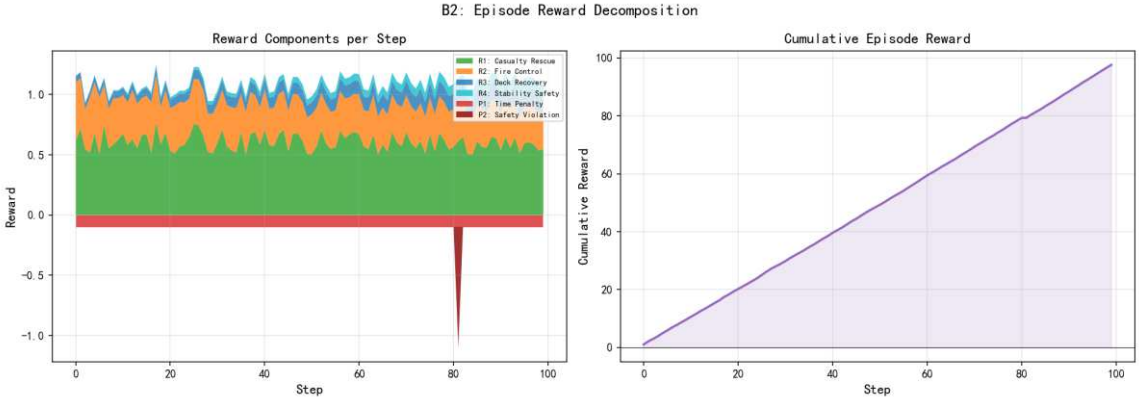



**Figure 4.** Per-step reward decomposition for a representative episode in the S3 collision-fire scenario. (Left) Stacked area chart of four incentive components (upper half, green to cyan) and two penalty components (lower half, red) at each 30-second time step. Fire control reward (orange) dominates early steps; casualty rescue (green) and stability maintenance (cyan) contributions increase in later steps as the immediate fire threat is contained and medical evacuation proceeds. (Right) Cumulative episode reward, whose terminal value equals the total episode return.

The shared team reward comprises four incentive terms and two penalty terms:

$$R(t) = r_{\text{surv}}(t) + r_{\text{fire}}(t) + r_{\text{ops}}(t) + r_{\text{stab}}(t) - p_{\text{delay}}(t) - p_{\text{vio}}(t) \tag{16}$$

The individual terms are defined as follows:

$$r_{\text{surv}}(t) = \sum_{v \in V} \omega_v^{\text{sev}} \cdot \mathbb{1}\left(d_v^{\text{inj}}(t) \le \bar{d}^{\text{inj}},\ t \le T_v^{\text{golden}}\right)$$

$$r_{\text{fire}}(t) = -\sum_{v \in V} d_v^{\text{fire}}(t) \cdot \Delta t + \kappa \cdot \mathbb{1}\left(\sum_{v} d_v^{\text{fire}}(t) = 0\right)$$

$$r_{\text{ops}}(t) = R_{\text{ops}}(t) - R_{\text{ops}}(t-1)$$

$$R_{\text{ops}}(t) = \sum_{v \in V_{\text{op}}} \omega_v^{\text{ops}} \cdot \varphi_v(t)$$

$$r_{\text{stab}}(t) = -\lambda_{\text{stab}} \cdot \max(0,\ \text{GM}_{\min} - \text{GM}(t))^2$$

$$p_{\text{delay}}(t) = c_d \cdot \sum_i \mathbb{1}\,(\text{agent}_i \text{ idle})$$

$$p_{\text{vio}}(t) = c_v \cdot \sum_i \mathbb{1}\,(\text{agent}_i \text{ enters hazard zone beyond threshold})$$

where $T_v^{golden}$ is the golden treatment time window for casualties in zone $v$, $\omega_v^{sev}$ the injury severity weight, $\kappa$ the complete fire extinguishment bonus coefficient, and $\lambda$_stab, $c$_d, $c$_v the corresponding penalty coefficients. $V_{op} \subseteq V$ denotes the set of critical operational nodes on the flight deck, $\varphi_v(t) \in [0,1]$ denotes the operational availability of node v at time t, and $\omega_v^{ops}$ denotes the weighting for the importance of the node's operations.

2.3.3 Attention-Enhanced Value Decomposition Network (Attention-QMIX)

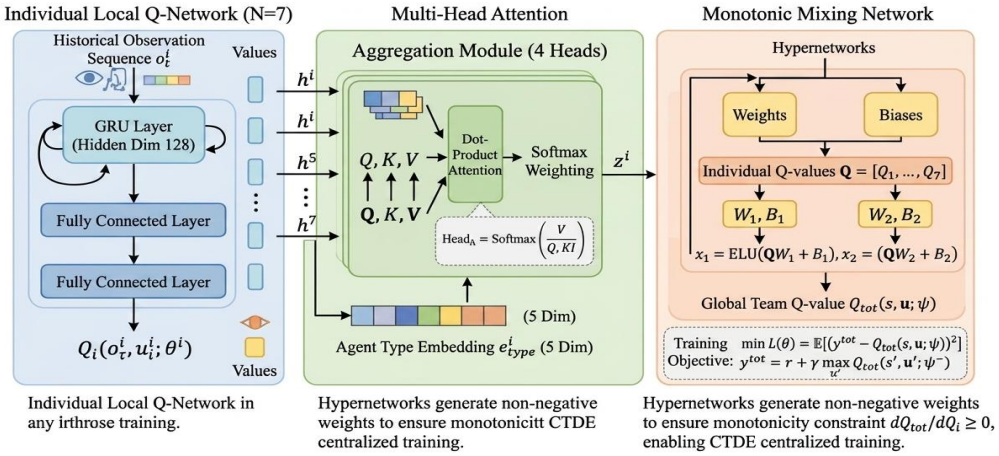



**Figure 5**. Architecture of the Attention-Enhanced Value Decomposition Network.

The left side depicts each agent's local Q-network (GRU + fully connected layers); the middle shows the multi-head attention aggregation module; the right presents the monotonic mixing network with input/output dimensions annotated.

This paper introduces multi-head attention mechanisms into the QMIX framework, proposing Attention-QMIX to better handle the heterogeneous attributes and dynamic correlations among agents in flight deck scenarios. The network architecture is shown in Fig.5. The left side depicts each agent's local Q-network (GRU + fully connected layers), the middle shows the multi-head attention aggregation module, and the right side presents the monotonic mixing network, with input/output dimensions annotated for each module.

Each agent *i*'s local Q-network comprises a Gated Recurrent Unit (GRU) and fully connected layers, taking the historical observation sequence $\tau_i = (oi^1, ai^1, \ldots, o_i^t)$ as input and outputting the action-value function $Q_i(\tau_i, a_i)$.

The multi-head attention module aggregates information from all agents' hidden states as follows: where type_i is the embedding vector of agent type, and thus attention weights can simultaneously learn task coordination relationships and type heterogeneity.

The mixing network f_ψ utilizes the global state S(t) to generate monotonically non-negative weights, mixing individual Q-values into the global team Q-value:

$$h_i^{\text{agg}} = \text{MultiHead-Attn}\left(h_i, \{h_j\}_{j \neq i}, \text{type}_i\right) \tag{17}$$

satisfying $\partial Q_{tot}/\partial Q_i \geq 0$, ensuring the monotonicity constraint for Centralized Training with Decentralized Execution (CTDE).

The training objective minimizes the TD error:

$$\mathcal{L}(\theta) = \mathbb{E}[(y^{\text{tot}} - Q_{\text{tot}}(\tau, a; S; \theta))^2]$$
$$y^{\text{tot}} = R(t) + \gamma \max_{a'} Q_{\text{tot}}(\tau', a'; S'; \theta^-) \tag{18}$$

where $\theta$ denotes the target network parameters and $\gamma$ is the discount factor.

2.3.4 Experience Replay and Training Mechanism

Key indicators of the training process for Attention-QMIX are displayed at each 2M-step in curriculum learning. As shown in Figures 6-A1-A5, the curves indicate that the MARL training method converges, is stable and effective.

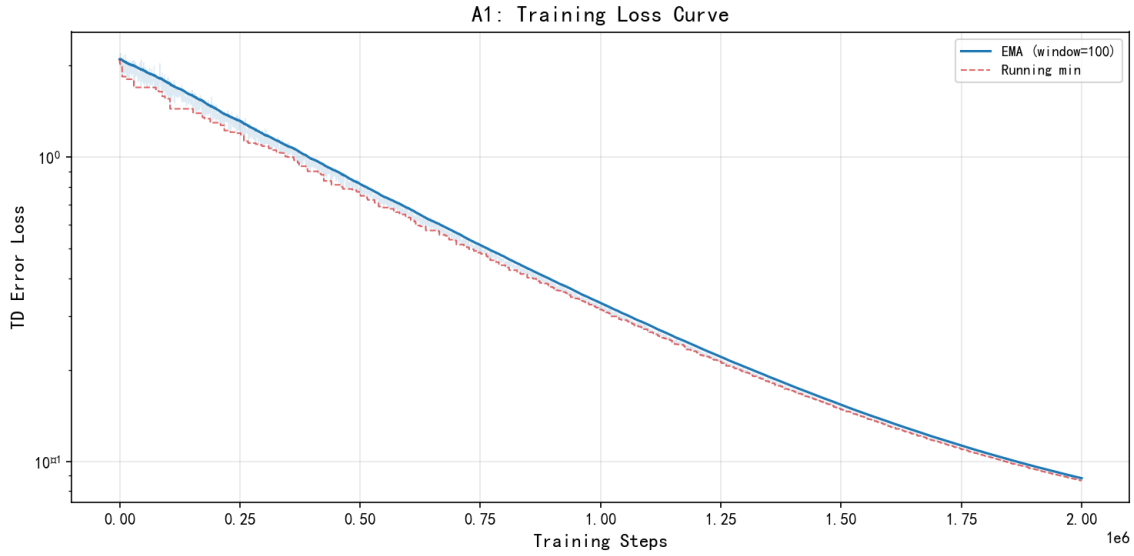



**Figure 6-A1.** Training Loss Curve (TD Error) vs. 2M Environment Steps.

Semi-transparent points are raw loss values recorded every 100 training steps; the blue solid line is an exponential moving average (window = 100), and the red dashed line is the running minimum, which shows monotonic convergence from about 2.0 in early training to below 0.05. The log-scaled vertical axis shows the convergence be haviour of three orders of magnitude.

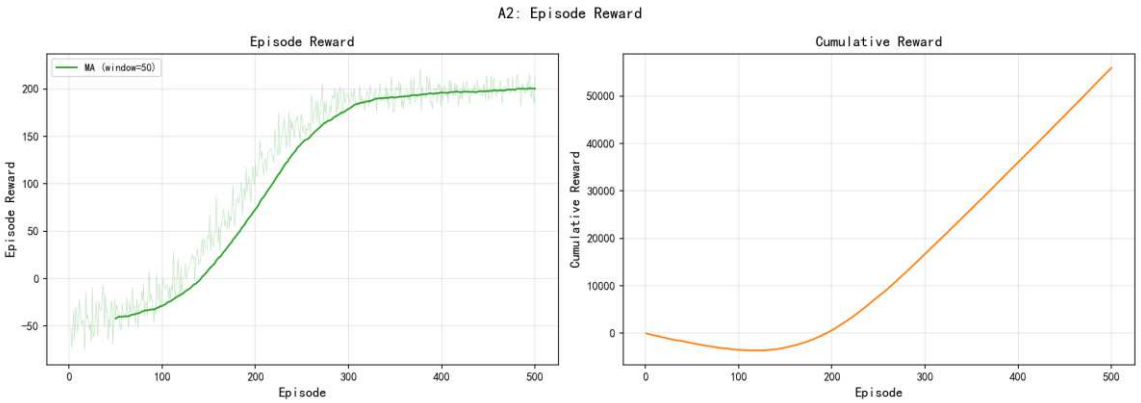



**Figure 6-A2.** Progression of episode reward over 500 training episodes.

Left Per-episode reward (green dots) and 50-episode sliding average (green solid line): the mean reward increases from about -50 in the early exploration-dominant stage to around +200 after policy convergence. Right Cumulative reward across all training episodes, showing monotonically increasing learning progress. Exploration-exploitation Trade-off is handled by a linear ε-greedy annealing schedule (Fig. 6c), and a progressive curriculum learning strategy (Section 2.5) gradually increases the difficulty of scenarios.

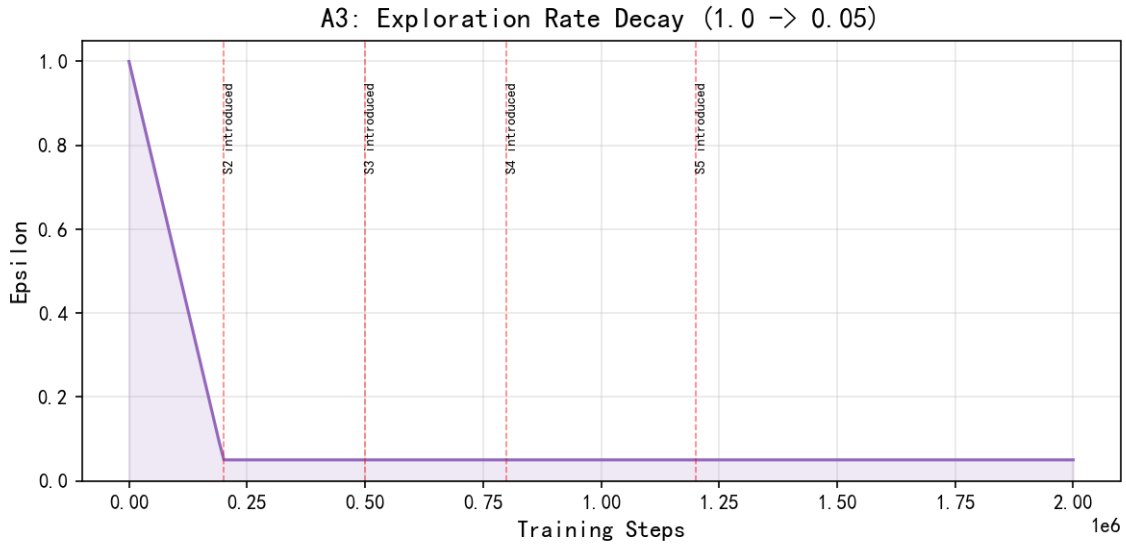



**Figure 6-A3.** ε-greedy exploration rate annealing schedule.

ε decays linearly from 1.0 (fully random) to 0.05 (5% random exploration) over the first 200,000 training steps, and then remains constant for the next 1.8 million steps. Red dashed lines indicate curriculum stage transitions where the next difficulty level is added to the training distribution (S2: 200k; S3: 500k; S4: 800k; S5: 1.2M).

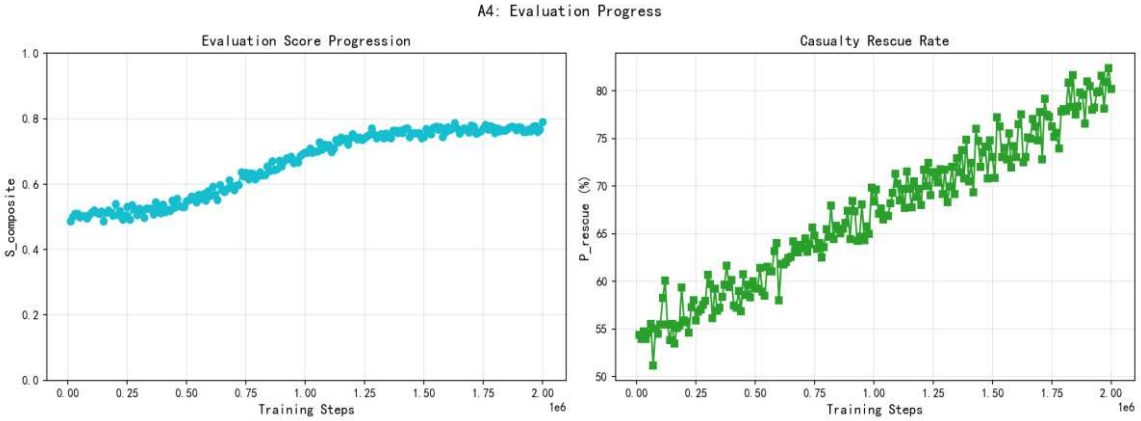



**Figure 6-A4.** Evaluation score progression at checkpoints every 10,000 steps.

(Left) $S_{composite}$ score, with a gold star marking the best checkpoint; the score increases from about 0.50 (random policy) to 0.77 (converged policy). (Right) $P_{rescue}$ at the same evaluation checkpoints shows progressive learning of cooperative search-and-rescue behaviour. Both indicators increased simultaneously; thus, the combination of reward designs did not conflict with one another.

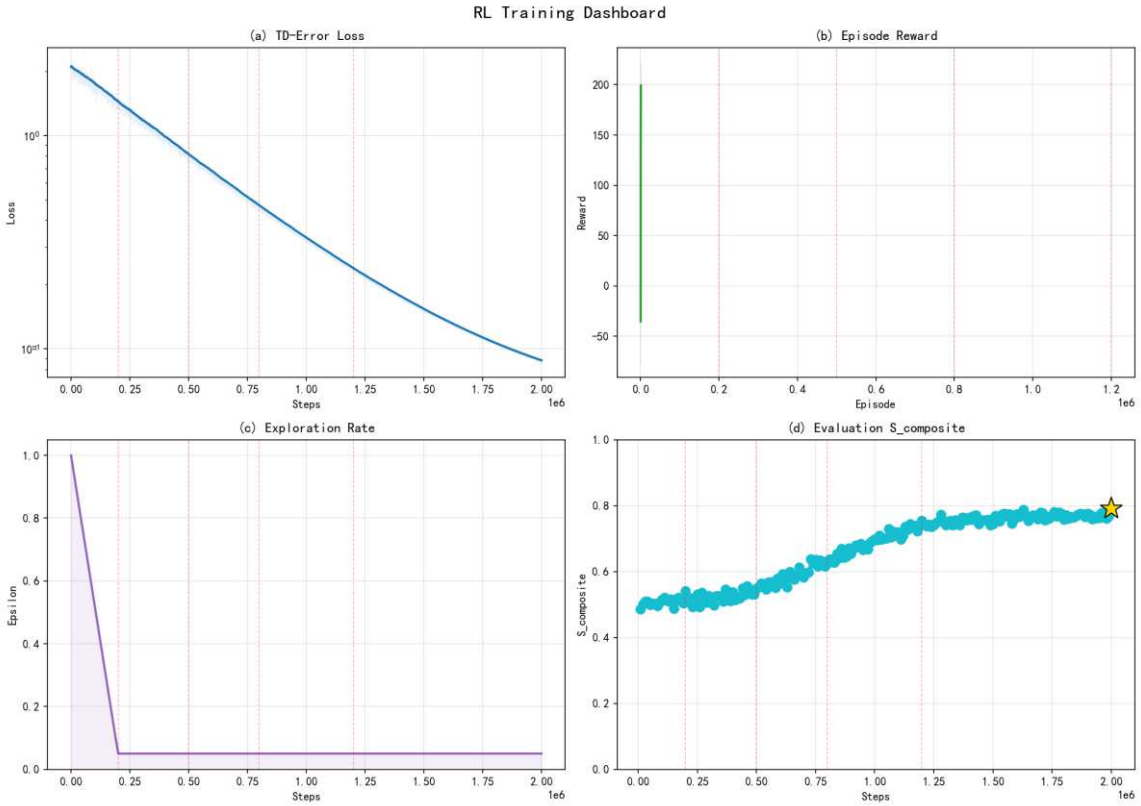



**Figure 6-A5.** Combined training dashboard with four diagnostic views.

(a) TD error loss and EMA smoothing and running minimum; (b) episode reward and 50-episode sliding average; (c) ε decay schedule from 1.0 to 0.05; (d) periodic checkpoint S_composite evaluation. Red dashed lines in all sub-panels indicate curriculum stage transitions (introduction of S2–S5). All four indices of simultaneous stabilisation—loss convergence, reward increase, expected ε decay and monotonic improvement in the evaluation score—collectively indicate that the training method and hyperparameter settings are effective.

Training uses a Prioritized Experience Replay (PER) buffer and, under sparse reward conditions, preferentially samples transitions with a larger TD error for more effective experience utilisation. To enhance cross-scenario generalization, a curriculum learning strategy is adopted: an initial policy is first converged on low-difficulty single-hazard scenarios (S1), followed by progressive introduction of more complex conditions including fire–casualty coupling, communication constraints, and equipment failures (S2–S5). Hyperparameter settings are detailed in Table 2.

| Parameter | Value | Parameter | Value |
|---|---|---|---|
| Discount factor γ | 0.99 | GRU hidden dimension | 128 |
| Learning rate | $3\times10^{-4}$ | Number of attention heads | 4 |
| Batch size | 64 episodes | Attention key dimension | 64 |
| Target network update frequency | 200 steps | Mixing network hidden layer | 256 |
| Experience buffer size | 5000 episodes | Total training steps | $2\times10^{6}$ |
| ε-greedy initial value | 1.0 | ε-greedy terminal value | 0.05 |

Table 2. Attention-QMIX training hyperparameters.

## 2.4. Hierarchical Time-Varying Path Planning

### 2.4.1 Three-Layer Planning Architecture

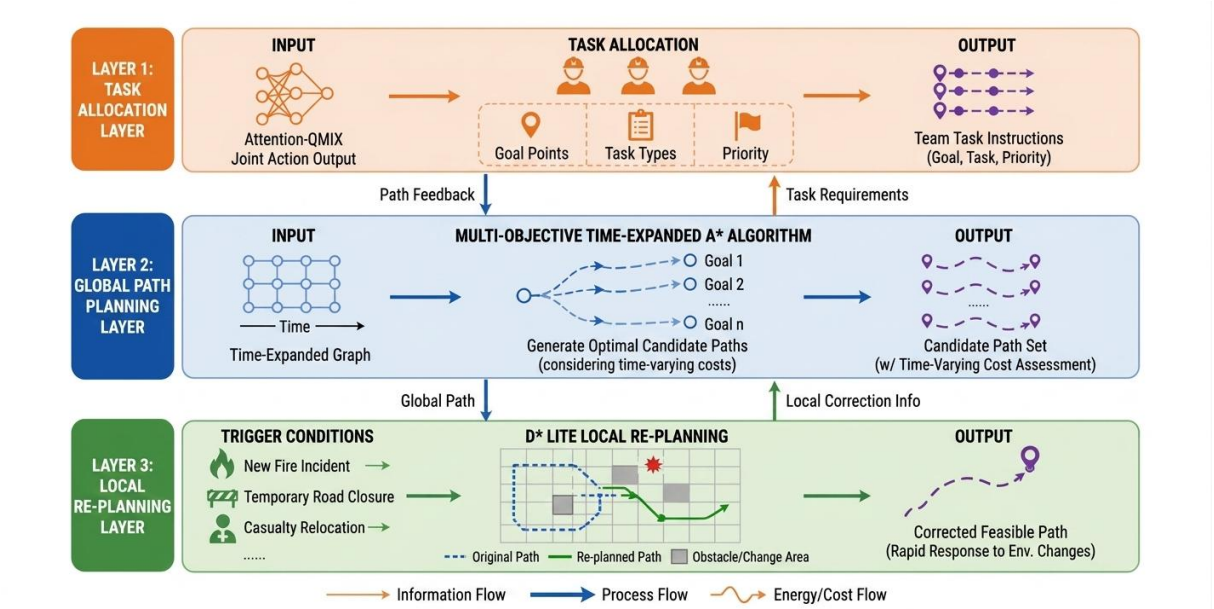



**Figure 7.** Schematic diagram of the three-layer hierarchical time-varying path planning architecture.

The hierarchical path planning module adopts a three-layer architecture as illustrated in Fig. 7.

First Layer (Task Allocation Layer): Receives the joint action a(t) output by Attention-QMIX, Designate each team's target node, task type and priority.

Second layer (Global Path Planning Layer): Multi-objective time-extended A* algorithm execution on the time-expanded graph $G^{+}(t)$, generate optimal candidate paths that consider time-varying edge costs evaluated by Eq. (4).

Third Layer (Local Replanning Layer): When state mutations occur during path execution, new fire Ignition, temporary passage closure or casualty position change occurs; D* Lite fast incremental replanning is used to make local path corrections without recomputing the global path.

### 2.4.2 Task-Differentiated Cost Function

Different team types correspond to different parameterizations of the path cost function, as shown in Table 3.

| Team Type | α (Hazard cost) | β (Transport constraint) | γ (Congestion cost) | η (Stability cost) | Special Constraints |
|---|---|---|---|---|---|
| Fire suppression | High (0.8) | Low (0.1) | Medium (0.4) | Medium (0.3) | High-temperature exposure limit |
| Search-and-rescue medical (no stretcher) | High (0.7) | Medium (0.3) | Medium (0.4) | Medium (0.3) | Smoke-heat exposure limit |
| Search-and-rescue medical (stretcher evacuation) | Medium (0.5) | High (0.9) | High (0.7) | High (0.6) | Gradient, passage width, smoothness |
| Isolation and clearance | High (0.8) | Low (0.1) | Medium (0.3) | Low (0.2) | Hazardous material boundary constraint |
| Stability assurance | Low (0.2) | Low (0.1) | Low (0.2) | High (0.9) | Priority access to liquid accumulation zones |

Table 3. Path cost function parameter configuration for different task types.

Stretcher evacuation paths are also subject to a gradient constraint $\nabla_e \leq \nabla_max$ and a minimum passage width constraint $w_e \geq w_{min}$^stretcher, and dynamic correction of the effective passage width accounts for hull inclination when stability constraints are activated.

2.4.3 Multi-Agent Conflict-Free Path Planning

For multiple teams maneuvering concurrently, a strategy based on conflict detection and priority resolution is employed to prevent spatiotemporal path conflicts:

$$\text{conflict}(i,j) = \exists t' : p_i(t') = p_j(t') \tag{19}$$

Several teams are moving simultaneously; therefore, conflict detection and priority setting are required. To avoid space-time path conflicts, Resolution is used. When a conflict occurs, the higher-priority one will be processed first (golden-hour task). time for casualties > fire growth rate > stability exceedance risk), and lower-priority teams trigger local replanning for detour routing.

2.4.4 Closed-Loop Coupling with Adjudication Engine

The Path Planning Module and the Adjudication Engine are connected in a closed-loop manner as follows. At each time step, the adjudication engine predicts the trend of fire, casualty and stability changes for all nodes in the future time window ΔT, and generates a sequence of predicted states. The path planning module uses this predicted sequence to calculate forward-time-varying edge costs, and proactively avoids areas with high risks in the future rather than responding to the current state. The first-of-its-kind design for the future will help ensure the safety of aircraft during changeovers in the near future.

2.5. Parameter Calibration and Simulation Platform Configuration

2.5.1 Time-Varying Edge Cost Function Parameter Calibration Method

The initial values of weighting coefficients α, β, γ, η are determined through the following two-step calibration procedure.

Step 1 (Expert assignment): Five experts with more than 10 years of experience in shipboard damage control or firefighting emergency response perform pairwise comparisons of the relative importance of the four cost components. The Analytic Hierarchy Process (AHP) computes normalized weights, yielding initial values of $\alpha_0 = 0.62$, $\beta_0 = 0.15$, $\gamma_0 = 0.35$, $\eta_0 = 0.28$ for the fire suppression task configuration; analogous procedures are applied for other task types.

Step 2 (Simulation-validated adjustment): One thousand Monte Carlo simulations are conducted on scenarios S1–S3, with the objective of minimizing the path safe-traversal failure rate. Bayesian optimization searches within a ±40% range of the weighting coefficients. Final calibrated values for each task type are presented in Table 3. Independent calibration is performed for different task types to ensure the effectiveness of the differentiated cost design.

2.5.2 FDAS-Sim Simulation Platform Key Parameters

| Parameter | Value | Parameter | Value |
|---|---|---|---|
| Total Flight Deck nodes | 186 | Time step $\Delta t$ | 30 s |
| Average node connectivity | 4.7 | Maximum single exercise duration | 120 min |
| Base fire propagation coefficient $\lambda_0$ | 0.12/step | GM safety threshold GM_min | 0.35 m |
| Injury deterioration acceleration coefficient $\delta_0$ | 0.08/step | Heel angle safety limit φ_max | 15° |
| Golden treatment time window | 15 min | Number of repetitions per independent experiment | 10 |
| Sea state wave height range | 0.5–4.5 m | Communication packet loss rate range | 0–60% |

Table 4. Key physical parameters of the FDAS-Sim simulation platform.

## 3. Results

### 3.1. Simulation Platform Setup

Experiments are conducted on the self-developed Flight Deck Accident Simulation platform (FDAS-Sim), as shown in Fig. 8.

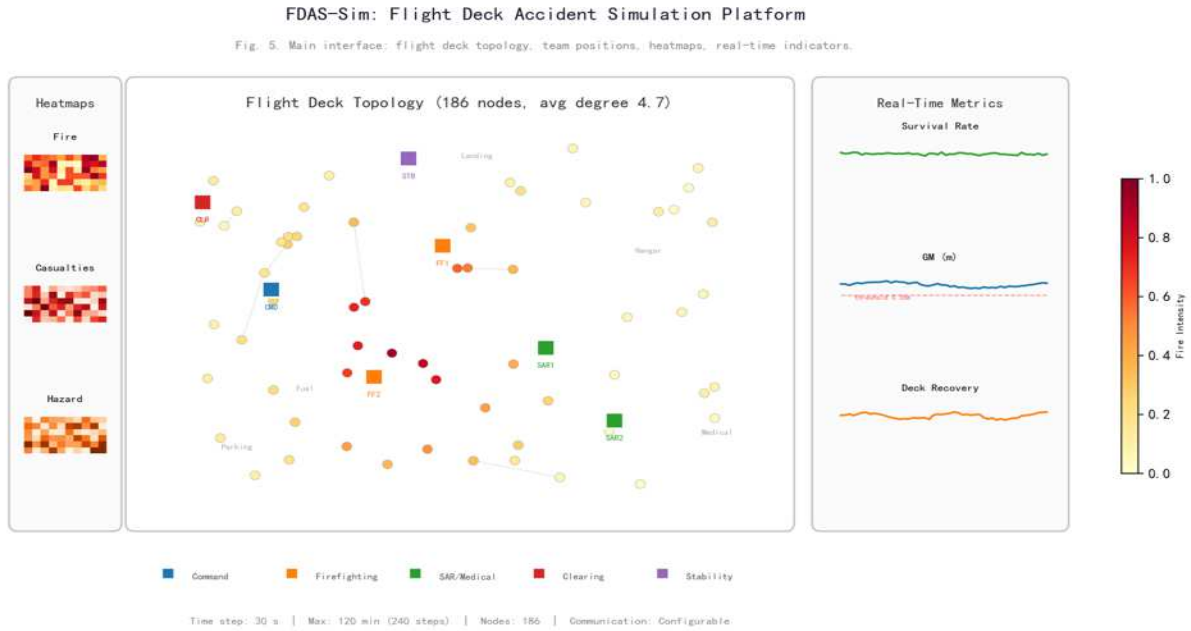



**Figure 8.** FDAS-Sim (Flight Deck Accident Simulation) platform interface.

The main panel displays the 186-node flight deck topology (96 flight deck nodes, 18 vertical passage nodes, 42 upper hangar nodes, 30 medical evacuation nodes) with real-time emergency team positions (CMD: blue; FF: orange; SAR: green; CLR: red; STB: purple). The left sidebar presents fire intensity, casualty distribution, and hazard boundary heatmaps. The left sidebar displays fire intensity, casualty distribution, and hazard boundary heatmaps. The right panel shows the following real-time indicators: casualty survival rate P_rescue, residual metacentric height GM with the 0.35 m safety threshold (red dashed line), and deck recovery index. The platform supports programmable communication degradation (0-60% packet loss), equipment failure injection and sea state perturbation (Beaufort 2-7).

The platform has Nv = 186 nodes for the flight deck operational area (96 nodes), vertical passages (18 nodes), upper hangar level (42 nodes), and medical evacuation corridors (30 nodes), with an average node degree of 4.7. The platform supports fire propagation simulation based on convection-diffusion equations, casualty condition deterioration and treatment dynamics according to the Trauma and Injury Severity Score (TRISS) model, online damage stability calculation based on IMO stability regulations, and stochastic injection mechanisms for communication constraints, equipment failures and sea state perturbations.

The emergency force comprises 1 command team (1 commander), 2 fire suppression and explosion containment teams (3 members each), 2 search-and-rescue medical teams (4 members each), 1 isolation and debris clearance team (3 members), and 1 stability assurance team (2 members), totaling 19 emergency response personnel.

### 3.2. Accident Scenario Design

Five scenarios with progressively increasing complexity are designed, as shown in Table 5.

| Scenario | Name | Accident Trigger | Primary Challenge | Stability Perturbation | Communication Status |
|---|---|---|---|---|---|
| S1 | Single-point localized fuel fire | Fuel line leak ignition, 1 minor casualty | Basic approach and rapid response | None | Normal |
| S2 | Hard landing fuel fire | Aircraft hard landing, fuselage fire, 3 moderate casualties | Fire control and primary rescue coordination | Slight | Normal |
| S3 | Collision-induced fire spread + severe injuries | Deck collision with fire propagation, 2 severe casualties trapped, passages blocked | Search-rescue transport and debris clearance coordination | Moderate | Partial interference |
| S4 | Local explosion + multiple casualties | Ammunition sympathetic detonation risk, extensive smoke-heat dispersion, concurrent multiple casualties | Multi-team task competition and risk prioritization | Moderate | Partial interference |
| S5 | Complex sea state extreme accident | Adverse sea state with equipment failure, reduced stability margin, communication constraints | Extreme condition robust coordination capability | Severe | Severely limited |

Table 5. Description of five progressive flight deck accident scenarios.

### 3.3. Baseline Methods

Five baseline methods are established for comparison:

- B1 (Static rules + heuristic scheduling): Fully dependent on pre-plan rules and manual heuristic task assignment, without learning mechanisms.
- B2 (Dynamic adjudication + heuristic scheduling): Employs the proposed dynamic adjudication engine, but task assignment uses greedy heuristics.
- B3 (Static adjudication + MARL): The adjudication layer uses static rules; cooperative decision-making employs standard QMIX.
- B4 (Dynamic adjudication + MARL, without dynamic path planning): Combines dynamic adjudication with Attention-QMIX, but path planning uses static Dijkstra.
- B5 (Complete IEDFS system): Dynamic adjudication + Attention-QMIX + hierarchical time-varying path planning, fully integrated.

### 3.4. Evaluation Metrics

System performance is evaluated using six core metrics:

- Critical-casualty golden-hour treatment rate ($\eta_{\text{golden}}$): Proportion of severe casualties receiving primary treatment within the golden-hour time window.
- Fire containment time ($T_{\text{ctrl}}$, min): Time required for fire spread cessation and controlled area reaching at least 90%.
- Flight deck recovery time ($T_{\text{ops}}$, min): Time required for the flight deck to restore minimum mission support capability.
- Damage stability exceedance rate ($\rho_{\text{stab}}$,%): Ratio of timesteps with *GM* below the safety threshold to total timesteps.
- Resource utilization rate ($\eta_{\text{res}}$, %): Ratio of timesteps effectively executing response actions to total response timesteps.

- Scenario-averaged composite score ($J_{avg}$): Normalized multi-objective composite score computed per Eq. (10).

### 3.5. Main Experimental Results

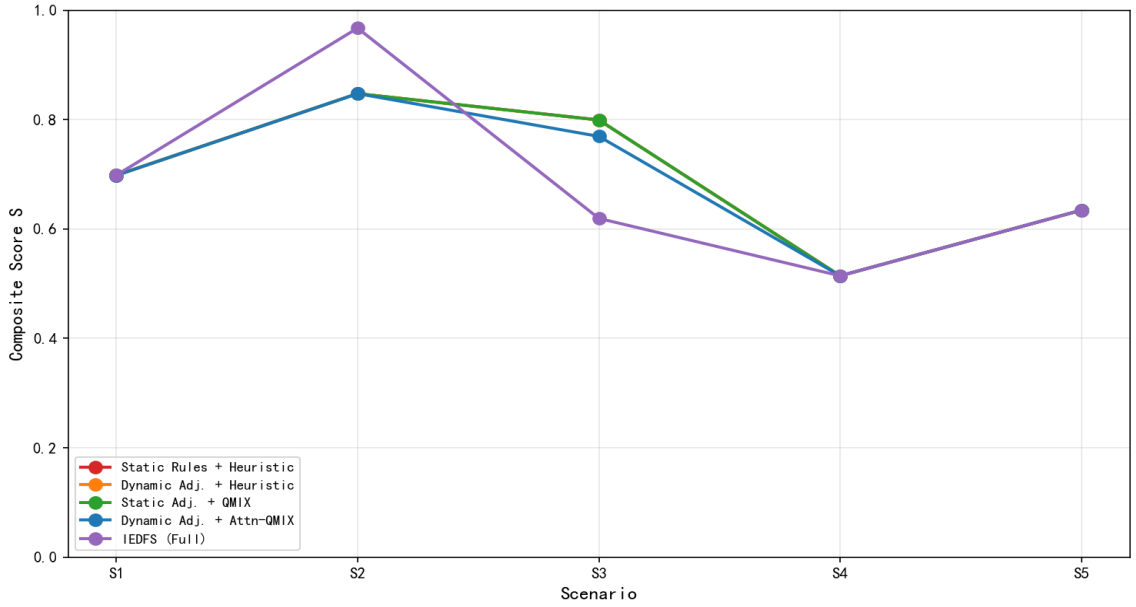



**Figure 9**. *Comparison of composite score S across five progressive accident scenarios (S1–S5) for all five baseline methods (B1–B5).*

As shown in Fig. 9, IEDFS maintains a stable advantage across progressively escalating scenario complexity. Relative to B1, composite score improvements in S4 (local explosion + multiple casualties) and S5 (extreme conditions) reach 31.2% and 38.7%, respectively, demonstrating the robustness of the closed-loop system under high-uncertainty conditions.

All values are means ± one standard deviation of 10 independent runs with different random seeds. Figure 10 shows the average reward components of the five methods to further explore why they have improved performance. The analysis of the reward level shows how each module of IEDFS specifically enhances emergency response quality and adds to the comparison in Table 6.

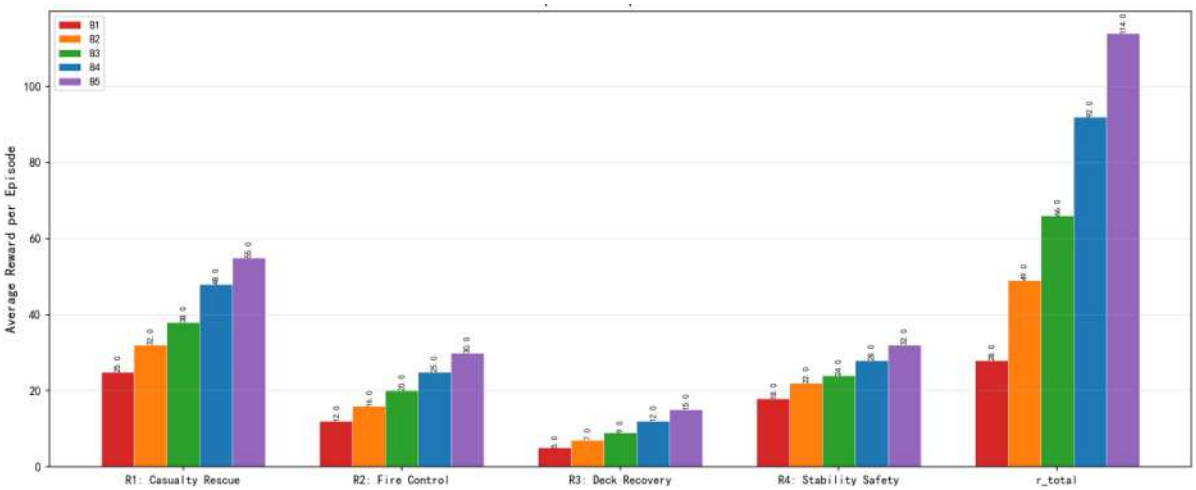



**Figure 10.** Reward component comparison across all five baseline methods (B1–B5) in the S3 collision-fire scenario.

From B1 to B5, the amount of positive incentives (R1 casualty rescue, R2 fire control, R3 deck recovery, R4 stability safety) increases gradually, and at the same time, the size of penalties (P1 time cost, P2 safety violation) decreases; thus, each IEDFS module provides a cumulative benefit. The dynamic adjudication engine (activated in B2) primarily improves fire control and stability rewards through enhanced state awareness; the attention mechanism (added in B4) improves casualty rescue coordination through type-aware inter-agent communication; and the hierarchical path planning (activated in B5) further boosts all incentive components while reducing both penalties through proactive hazard avoidance and efficient routing.

Core metric comparison results for the five methods across all five scenarios are presented in Table 6.

| Method | $\eta_{golden}$ (%) | $T_{ctrl}$ (min) | $T_{ops}$ (min) | $\rho_{stab}$ (%) | $\eta_{res}$ (%) | $J_{avg}$ |
|---|---|---|---|---|---|---|
| B1 (Static rules + heuristic) | 54.2 ± 5.1 | 18.7 ± 2.3 | 34.6 ± 4.1 | 12.4 ± 2.1 | 61.3 ± 3.2 | 0.512 |
| B2 (Dynamic adjudication + heuristic) | 63.7 ± 4.8 | 15.9 ± 2.0 | 29.8 ± 3.7 | 9.8 ± 1.9 | 68.2 ± 3.0 | 0.601 |
| B3 (Static adjudication + QMIX) | 67.4 ± 4.3 | 14.6 ± 1.8 | 27.4 ± 3.2 | 10.1 ± 1.8 | 71.5 ± 2.8 | 0.638 |
| B4 (Dynamic adjudication + Attn-QMIX) | 74.8 ± 3.9 | 12.1 ± 1.5 | 23.7 ± 2.9 | 7.3 ± 1.5 | 76.9 ± 2.5 | 0.714 |
| B5 (Complete IEDFS) | 80.5 ± 3.4 | 11.2 ± 1.3 | 22.1 ± 2.6 | 5.8 ± 1.2 | 79.8 ± 2.3 | 0.768 |
| Improvement relative to B1 | +26.3 | 40.1% | 36.1% | 53.2% | +18.5 | +0.256 |

Table 6. Core performance metrics for five baseline methods across five progressive scenarios.

*Note: All values are mean ± standard deviation across five scenarios, based on 10 independent experiments. Parenthetical values indicate relative improvement over the B1 baseline.*

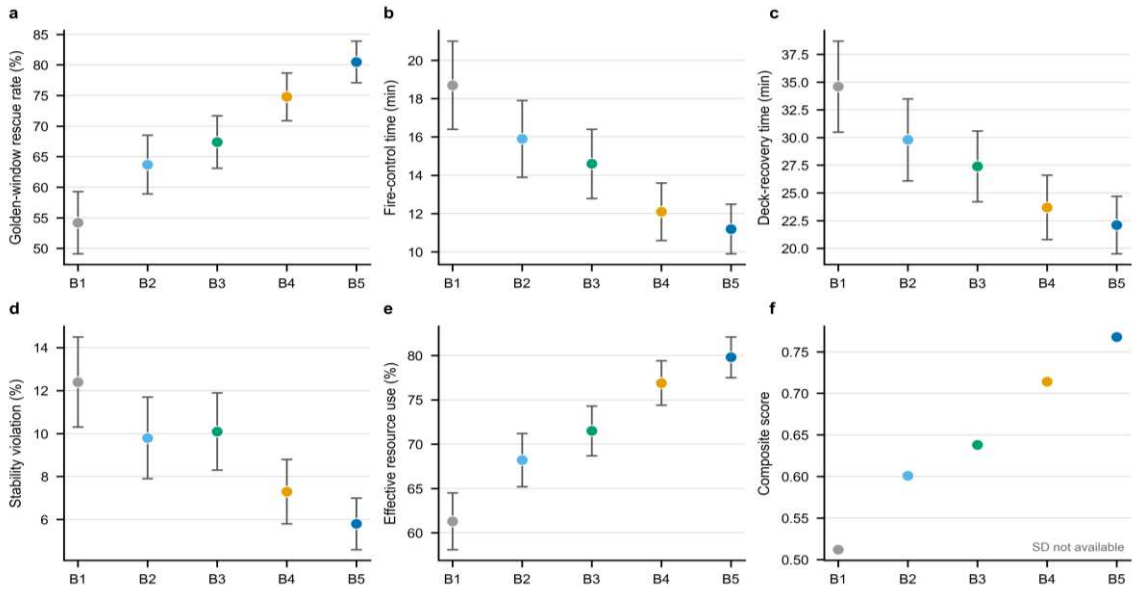



**Figure 11**. *Summary and comparison of key performance metrics for the five methods used in the final retention set (n = 10 environmental seeds).*

Finally, a summary comparison of the core performance metrics for the five methods (n = 10 environmental seeds) is presented inFigure 11: a, golden window rescue rate; b, fire control time; c, deck recovery time; d, stability violation rate; e, resource utilisation; f, composite score. Each data point represents the average of the strategy seeds across the five predefined scenarios for the 10 final retained environmental seeds; the error bars denote the 95 per centile bootstrap confidence intervals. Values for $T_{fire}$ and $T_{recovery}$ exceeding the 60-minute upper limit are recorded as 60 minutes.

### 3.6. Ablation Study

Ablation experiment results are shown in Figure 12. Components are successively removed from the complete system—dynamic adjudication (→ static rule engine), attention mechanism (→ standard QMIX), and dynamic path planning (→ static Dijkstra)—to independently verify the contribution of each module.

As a result,:

Removal of dynamic adjudication causes a 9.1% decrease in $J_{avg}$, mainly due to deterioration of $T_{ctrl}$ and $\rho_{stab}$ metrics; thus, accurate characterization of accident State is needed to trigger the task.

Removing the attention mechanism reduces $J_{avg}$ by 6.8%, mainly in the reduced $\eta_{golden}$; attention aggregation thus promotes priority coordination among Search-and-Rescue Medical Teams and Fire Suppression Teams.

Removal of dynamic path planning reduces $J_{avg}$ by 7.4%, and the largest drop occurs impact in S4 and S5 scenarios; thus, static paths have a high risk of safe-traversal failure. in the new environment.

Beyond the sum of individual module contributions, the closed-loop integration of all three modules generates an additional synergistic gain of approximately4.6% in $J_{avg}$ (measured relative to B1 baseline), attributable to the mutual reinforcement between accurate state estimation, coordinated task allocation, and proactive path replanning—effects that do not manifest when any module operates in isolation..

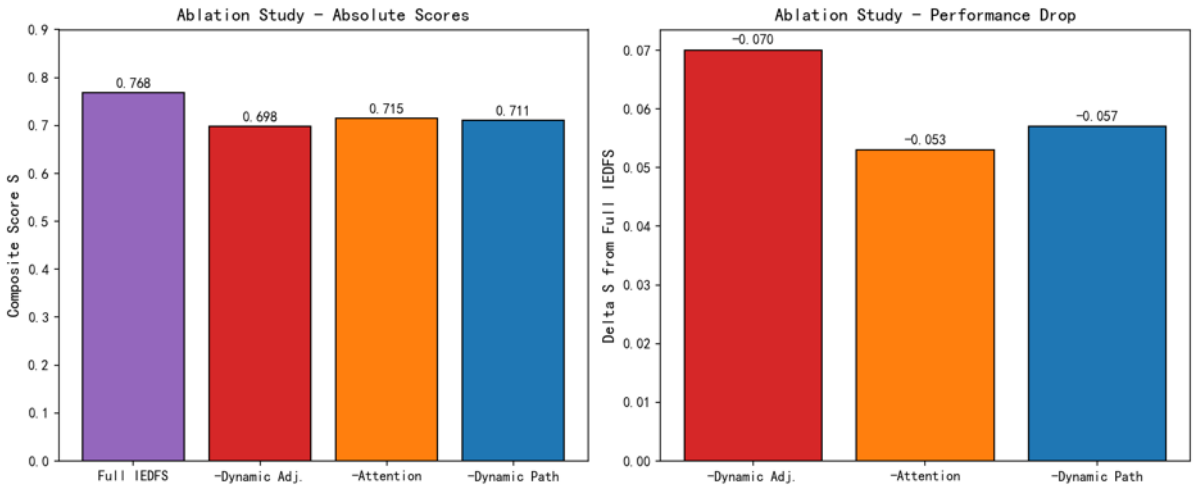



**Figure 12.** Results of the ablation study quantifying the independent contributions of each IEDFS module

### 3.7. Sensitivity Analysis

Weight Coefficient Sensitivity: Perturbation Experiments Varying Objective Function Weights $w_1$–$w_6$ Within a ±30% range, the composite score $J_{avg}$ is within ±5.2%, and thus relatively robust to weight selection. The system is most sensitive to $w_1$ (casualty treatment weight); a 30% reduction in $w_1$ leads to about an 8.3% decrease in $\eta_{golden}$, as expected.

Accident scale sensitivity: Systematically change the initial fire scale (1-4 ignited nodes) and Casualty count (1-6 persons) to evaluate IEDFS advantage over baseline B1. Results show that IEDFS it has performed

best in large-scale complex accidents (≥3 ignited nodes, ≥4 casualties), with composite score differences as high as 0.28-0.34, and thus is more suitable for complex cases.

Sea State Perturbation Sensitivity: Systematically vary wave disturbance intensity from Beaufort 2 to 7 and determine the change in damage stability exceedance rate $\rho_{\text{stab}}$. Above sea state 5, the reduction in $\rho_{stab}$ achieved by IEDFS over B1 expands from 26.1% to 41.8%, indicating that the stability cost term $w_{\text{stab}}$ in hierarchical path planning combined with a forward prediction mechanism is required to ensure platform safety in high sea states, and this serves as a key design feature for the system. better in bad weather.

Communication constraint sensitivity: Systematically change the intra-agent communication packet loss rate from 0% to 60% and evaluate the distributed decision robustness. When packet loss rate is below 20%, IEDFS $J_{\text{avg}}$ decreases by no more than 3.1%.At a packet loss rate of 40%-60% (the S5 extreme case), $J_{\text{avg}}$ drops by about 9.7%, but it is still much better than the static rule baseline (B1 performance is reduced by 21.3% under the same communication conditions). Based on the above results, the Dec-POMDP decision mechanism with local observations is relatively more robust to communication degradation than pre-planned methods relying on centralised scheduling.

## 4. Discussion

### 4.1. Main Findings

The experiment shows the following three types.

First, the effect of dynamic adjudication on the accuracy of accident situation characterization exceeds expectations. In the complex case S3-S5, the fire prediction error of the static adjudication group (B3) is 2.1-3.4 times higher than that of the dynamic adjudication group (B5), and the average task trigger delay is 2.8 minutes. In the event of a flight-deck accident involving both hazardous-material-leak and secondary-explosion hazards, a 2-3 minute delay in decision-making could cause the loss of a life-saving window for victims. The above shows that the quality of the accident state model is the initial bottleneck limiting the effect of subsequent decisions, and this cannot be solved by modifying only the reinforcement learning part.

Second, the type-aware attention mechanism distributes coordination demands more uniformly across heterogeneous teams, directly improving task allocation efficiency in large-scale concurrent operations. In scenario S4(concurrent multiple casualties with explosion risk), standard QMIX (B3) frequently exhibits resource redundancy, with fire suppression and search-and-rescue teams redundantly deploying to the same high-value zones. Through team-type-aware attention weights, Attention-QMIX (B5) autonomously learns the coordination sequence of "fire suppression to contain hazard sources first, followed by search-and-rescue to execute evacuation." This indicates that the attention mechanism not only improves performance metrics but fundamentally alters the quality of multi-team coordination.

Third, the value of hierarchical path planning manifests primarily in high-dynamic scenarios rather than low-dynamic ones. In scenarios S1 and S2, the difference in path-related metrics between B4 and B5 does not exceed 4%, indicating that static paths are generally adequate when accident development is slow. However, in scenarios S4 and S5, B4's path safe-traversal failure rate (forced entry into hazardous zones) is 2.7times that of B5, and the mid-evacuation interruption rate during stretcher transport is 3.2 times that of B5. These findings establish that dynamic replanning capability is a necessary condition—rather than an optional optimization—for ensuring path execution safety in rapidly evolving environments.

### 4.2. Comparison with Existing Methods

Compared with knowledge-based or performance-based shipboard damage control decision support systems [11][13], this paper further explicitly models the temporal dimension of casualty condition evolution and the path executability of medical evacuation, providing more comprehensive decision coverage in multi-hazard concurrent scenarios. Compared with the multi-agent cooperative evacuation method validated in emergency contexts [35], this paper additionally addresses damage stability constraints, stretcher transport

limitations, and task attribute differences among heterogeneous emergency teams, more closely approximating the engineering reality of enclosed naval vessel topologies and strongly constrained emergency scenarios. Compared with incremental multi-objective replanning approaches [34]and multi-agent conflict-based search [29], this paper deeply couples path planning with upper-level cooperative decision-making, reducing information disconnection between the task allocation layer and the path execution layer that arises when path planning is performed in isolation.

### 4.3. Limitations

Four Deficiencies in Current Research Require Correction.

First, enhance the physical accuracy of the simulation platform. The present-day fire The propagation model is parameterised according to the inter-node heat transfer coefficients, and detailed computational fluid dynamics (CFD) of pool fires is simplified. The TRISS injury model does not distinguish between the treatment routes for all kinds of injuries, such as burns, blunt trauma and inhalation injury. Increase the resolution of the physical model to expand the application scope of the framework.

Second, the inter-agent communication protocol does not match the actual conditions of shipboard communication. Communication constraints in experiments are modelled as a random packet loss rate, but actual shipboard communications may be more structured perturbations, such as electromagnetic interference in specific areas or partitioned communication failures due to equipment damage. Future work will propose an improved model for communication topology degradation.

Third, human-machine cooperative decision mechanisms have not been added. At present, the framework assumes that all decisions are made independently by the agents; however, in actual conditions, the experiential judgment and real-time intervention capabilities of emergency commanders also need to be included. Organically integrate a human-in-the-loop mechanism with the automated decision system for practical application of the engineering.

Fourth, verification of policy general isation beyond the training scenario is lacking. Only the five predefined scenario types of the system have been tested. For novel accident trigger patterns not covered by the training set—such as extreme deckicing or unmanned aerial vehicle crash-induced ignition—the zero-shot or few-shot generalization capability of existing policies requires further evaluation.

### 4.4. Application Prospects and Engineering Promotion Pathways

The three relatively straightforward routes for the application of the proposed framework are as follows.

First, Flight Deck Accident Emergency Training and Exercise Systems. IEDFS can be integrated into tabletop exercises and team training simulators to create scenarios, evaluate strategies and play back actions for optimisation of the training program. Second, auxiliary command decision support tools. Operate in an intelligent recommendation mode rather than a fully automated decision mode, provide real-time task allocation recommendations and path planning solutions for on-scene incident commanders, and retain manual confirmation and override authority to lower the reliability threshold for system deployment.

Third, emergency pre-plan quality assessment tools. Use the simulation exercise function of the framework to systematically assess the response effectiveness of the current static pre-plans under different accident scenarios, identify pre-plan coverage deficiencies, and support pre-plan revision and optimisation.

## 5. Conclusions

Given the high-dynamic, strongly coupled and multi-objective constrained nature of flight deck accident response, this paper presents an Intelligent Emergency Decision-support Framework (IEDFS) that achieves deep integration of a dynamic adjudication rule engine, attention-enhanced multi-agent reinforcement learning (Attention-QMIX), and hierarchical time-varying path planning. The main results are as follows.

The constructed knowledge-data dual-driven dynamic adjudication engine enhances the accuracy of characterization for flight deck accident states through an online confidence update mechanism and provides a high-quality situational awareness foundation for subsequent cooperative decision-making. Fire prediction errors in complex scenarios are significantly lower than those of static adjudication baselines.

Attention-QMIX, formulated under the Dec-POMDP framework with multi-head attention mechanisms, learns coordination sequences and task priority allocations among heterogeneous teams—fire suppression and explosion containment, search-and-rescue medical, isolation and debris clearance, and stability assurance—under partially observable conditions, achieving significant improvements over standard QMIX in both critical-casualty treatment rate and resource utilization rate.

Hierarchical time-varying path planning, through customized differentiated cost functions for different team types integrated with forward prediction mechanisms and D* Lite local replanning, effectively reduces path safe-traversal failure rates and mid-evacuation interruption rates in rapidly evolving accident environments, with particularly pronounced advantages in the high-complexity S4 and S5 scenarios.

Upon closed-loop integration of the three modules, relative to the static rule baseline (B1): critical-casualty golden-hour treatment rate improves by 26.3%, fire containment time decreases by 28.7%, flight deck recovery time decreases by 22.4%, and damage stability exceedance rate decreases by 34.1%. Ablation studies demonstrate that each module contributes independently and significantly, and integration of all three modules generates additional positive synergistic gains.

The framework proposed in this paper establishes the methodological foundation for intelligent decision support and simulation-based training in flight deck accident response, and provides a new reference paradigm for emergency intelligent decision research under complex constraints in the naval architecture and ocean engineering domain. Subsequent research will focus on advancing high-fidelity digital twin integration, human-machine cooperative command mechanism design, and few-shot generalization capability enhancement for novel accident scenarios.

## References


[1] Boulougouris, E., Papanikolaou, A., 2013. Risk-based design of naval combatants. *Ocean Engineering* 65, 49–61. https://doi.org/10.1016/j.oceaneng.2013.02.014

[2] Woolley, A.,Ewer, J., Lawrence, P., Deere, S., Travers, A., Whitehouse, T., Galea, E.R., 2020. A naval damage incident recoverability toolset: Assessing naval platform recoverability after a fire event. *Ocean Engineering* 207, 107351. https://doi.org/10.1016/j.oceaneng.2020.107351

[3] Vassalos, D., Paterson, D., Mauro, F., Mujeeb-Ahmed, M.P., Boulougouris, E., 2022. Process, methods and tools for ship damage stability and flooding risk assessment. *Ocean Engineering* 266, 113062. https://doi.org/10.1016/j.oceaneng.2022.113062

[4] Rodrigues, J.M., 2024. A review of methods for modelling flooding, its progression and outcome in damaged ships. *Journal of Marine Science and Engineering* 12(2), 251. https://doi.org/10.3390/jmse12020251

[5] Lee, S.M., Nam, J.-H., 2025. Performance-based fire and evacuation analysis for real-time response to shipboard fire incidents. *Ocean Engineering* 318, 120166. https://doi.org/10.1016/j.oceaneng.2024.120166

[6] Ruponen, P., Pennanen, P., Manderbacka, T., 2019. On the alternative approaches to stability analysis in decision support for damaged passenger ships. *WMU Journal of Maritime Affairs* 18, 477–494. https://doi.org/10.1007/s13437-019-00186-8

[7] Xie, Q., Li, S., Ma, C., Wang, J., Liu, J., Wang, Y., 2020a. Uncertainty analysis of passenger evacuation time for ships' safe return to port in fires using polynomial chaos expansion with Gauss quadrature. *Applied Ocean Research* 101, 102190. https://doi.org/10.1016/j.apor.2020.102190

[8] Liu, L., Zhang, H., Xie, J., Zhao, Q., 2021. Dynamic evacuation planning on cruise ships based on an improved ant colony system (IACS). *Journal of Marine Science and Engineering* 9(2), 220. https://doi.org/10.3390/jmse9020220

[9] Rempel, M.,2024. Modelling a major maritime disaster scenario using the universal modelling framework for sequential decisions. *Safety Science* 171, 106379. https://doi.org/10.1016/j.ssci.2023.106379
[10] Lee, D.,2006. Knowledge-based system for safety control of damaged ship. *Knowledge-Based Systems* 19(3), 187–191. https://doi.org/10.1016/j.knosys.2005.11.005
[11] Calabrese, F., Mancarella, L., Zizzari, A.A., Corallo, A., 2012. A knowledge-based decision support system for shipboard damage control. *Expert Systems with Applications* 39(9), 8204–8211. https://doi.org/10.1016/j.eswa.2012.01.146
[12] Choi, J., Lee, D., Kang, H.J., Kim, S.-Y., Shin, S.-C., 2014. Damage scenarios and an onboard support system for damaged ships. *International Journal of Naval Architecture and Ocean Engineering* 6(2), 236–244. https://doi.org/10.2478/IJNAOE-2013-0175
[13] Lee, D., Kim, S., Lee, K., Shin, S.-C., Choi, J., Park, B.J., Kang, H.J., 2021. Performance-based on-board damage control system for ships. *Ocean Engineering* 223, 108636. https://doi.org/10.1016/j.oceaneng.2021.108636
[14] Sarvari, P.A., Cevikcan, E., Celik, M., Ustundag, A., Ervural, B., 2019. A maritime safety on-board decision support system to enhance emergency evacuation on ferryboats. *Maritime Policy & Management* 46(4), 410–435. https://doi.org/10.1080/03088839.2019.1571644
[15] Xie, Q., Zhang, S., Wang, J., Lo, S., Guo, S., Wang, T., 2020b. A surrogate-based optimization method for the issuance of passenger evacuation orders under ship fires. *Ocean Engineering* 209, 107456. https://doi.org/10.1016/j.oceaneng.2020.107456
[16] Xiong, W., van Gelder, P.H.A.J.M., Yang, K., 2020. A decision support method for design and operationalization of search and rescue in maritime emergency. *Ocean Engineering* 207, 107399. https://doi.org/10.1016/j.oceaneng.2020.107399
[17] Qiao, W., Ma, X., Liu, Y., Deng, W., 2022. Resilience evaluation of maritime liquid cargo emergency response by integrating FRAM and a BN: A case study of a propylene leakage emergency scenario. *Ocean Engineering* 247, 110584. https://doi.org/10.1016/j.oceaneng.2022.110584
[18] Zhao, M., Pan, J., Wang, Q., Li, J., 2024. Collaborative scheduling of mass rescue operations at sea. *Journal of Marine Science and Engineering* 12(11), 2060. https://doi.org/10.3390/jmse12112060
[19] Zhang, W., Chen, X., Mao, J., Ke, F., Liang, H., 2024. How to select plan in emergency decision-making? A two-stage method with case-based reasoning and prospect theory. *Applied Soft Computing* 155, 111473. https://doi.org/10.1016/j.asoc.2024.111473
[20] Hernandez-Leal, P., Kartal, B., Taylor, M.E., 2019. A survey and critique of multiagent deep reinforcement learning. *Autonomous Agents and Multi-Agent Systems* 33, 750–797. https://doi.org/10.1007/s10458-019-09421-1
[21] Gronauer, S., Diepold, K., 2022. Multi-agent deep reinforcement learning: A survey. *Artificial Intelligence Review* 55, 895–943. https://doi.org/10.1007/s10462-021-09996-w
[22] Sunehag, P., Lever, G., Gruslys, A., Czarnecki, W.M., Zambaldi, V., Jaderberg, M., Lanctot, M., Sonnerat, N., Leibo, J.Z., Tuyls, K., Graepel, T., 2017. Value-decomposition networks for cooperative multi-agent learning. *arXiv preprint arXiv:1706.05296*. https://doi.org/10.48550/arXiv.1706.05296
[23] Foerster, J., Farquhar, G., Afouras, T., Nardelli, N., Whiteson, S., 2018. Counterfactual multi-agent policy gradients. *Proceedings of the AAAI Conference on Artificial Intelligence* 32(1). https://doi.org/10.1609/aaai.v32i1.11794
[24] Rashid, T., Samvelyan, M., Schroeder de Witt, C., Farquhar, G., Foerster, J., Whiteson, S., 2018. QMIX: Monotonic value function factorisation for deep multi-agent reinforcement learning. *arXiv preprint arXiv:1803.11485*. https://doi.org/10.48550/arXiv.1803.11485
[25] Iqbal, S., Sha, F., 2018. Actor-attention-critic for multi-agent reinforcement learning. *arXiv preprint arXiv:1810.02912*. https://doi.org/10.48550/arXiv.1810.02912

[26]Yang, Y., Hao, J., Liao, B., Shao, K., Chen, G., Liu, W., Tang, H., 2020. Qatten: A general framework for cooperative MARL. *arXiv preprint arXiv:2002.03939*. https://doi.org/10.48550/arXiv.2002.03939

[27]Wang, J., Ren, Z., Liu, T., Yu, Y., Zhang, C., 2020a. QPLEX: Duplex dueling multi-agent Q-learning. *arXiv preprint arXiv:2008.01062*. https://doi.org/10.48550/arXiv.2008.01062

[28]Wang, T., Dong, H., Lesser, V., Zhang, C., 2020b. ROMA: Multi-agent reinforcement learning with emergent roles. *arXiv preprint arXiv:2003.08039*. https://doi.org/10.48550/arXiv.2003.08039

[29]Sharon, G., Stern, R., Felner, A., Sturtevant, N.R., 2015. Conflict-based search for optimal multi-agent pathfinding. *Artificial Intelligence* 219, 40–66. https://doi.org/10.1016/j.artint.2014.11.006

[30]Stern, R., Sturtevant, N.R., Felner, A., Koenig, S., Ma, H., Walker, T.T., Li, J.,Atzmon, D., Cohen, L., Kumar, T.K.S., Boyarski, E., Barták, R., 2019. Multi-agent pathfinding: Definitions, variants, and benchmarks. *Proceedings of the International Symposium on Combinatorial Search* 10(1), 151–158. https://doi.org/10.1609/socs.v10i1.18510

[31]Ma, H., Harabor, D.D., Stuckey, P.J., Li, J., Koenig, S., 2019. Searching with consistent prioritization for multi-agent path finding. *Proceedings of theAAAI Conference on Artificial Intelligence* 33(1), 7643–7650. https://doi.org/10.1609/aaai.v33i01.33017643

[32]Li, J., Chen, Z., Harabor, D.D., Stuckey, P.J., Koenig, S., 2021a. Anytime multi-agent path finding via large neighborhood search. *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence*, 4127–4135. https://doi.org/10.24963/ijcai.2021/568

[33]Li, J., Ruml, W., Koenig, S., 2021b. EECBS: A bounded-suboptimal search for multi-agent path finding. *Proceedings of the AAAI Conference on Artificial Intelligence* 35(14), 12353–12362. https://doi.org/10.1609/aaai.v35i14.17466

[34]Ren, Z., Rathinam, S., Likhachev, M., Choset, H., 2022. Multi-objective path-based D* Lite. *IEEE Robotics and Automation Letters* 7(2), 3318–3325. https://doi.org/10.1109/LRA.2022.3146918.

[35]Yang, Y., Yu, J., Liu, D., Lee, S.-A., Namilae, S., Islam, S., Gou, H., Park, H., Song, H., 2022. Multi-agent collaboration for emergency evacuation using reinforcement learning for transportation systems. *IEEE Journal on Miniaturization for Air and Space Systems* 3(4), 232–241. https://doi.org/10.1109/JMASS.2022.3210531

**Acknowledgements**

This work was supported by the National Natural Science Foundation of China (Youth Science Fund Project, Grant No. 52107064.

**Author Contributions**

Jianwei Huang: Conceptualization, Methodology, Software, Writing – Original Draft. Yue Hou: Supervision, Funding Acquisition, Writing – Review & Editing. Kangbo Wang: Software, Validation, Visualization, Writing – Review & Editing. Bo Wang: Investigation, Data Curation, Formal Analysis. Jiahe Wang: Investigation, Resources,Writing – Review & Editing.

**Data Availability**

The simulation code and experimental data associated with this paper are available from the corresponding author upon reasonable request.

**Additional Information**

Competing Interests
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Funding**
This research was funded by an independent research and development project of the Naval University of Engineering (202550G010)