

AGI-inspired UAV Cluster Communication Self-Organization Framework: Multi-Agent Learning and Network Optimization

Fan Zhou 

School of Information Technology, Urban Vocational College of Sichuan, China

zhoufan.green@163.com

Received: 2 April 2026; Accepted: 9 September 2026; Published: 1 October 2026

Abstract: Unmanned Aerial Vehicle (UAV) cluster communication networks play a vital role in Space-Air-Ground Integrated Networks (SAGIN), offering flexible deployment and high mobility for emergency communications, disaster relief, and edge computing. However, traditional approaches face significant limitations in dynamic environments, often focusing on single objectives, lacking adaptive capabilities, and failing to achieve global coordination across multiple UAVs. To address these challenges, this paper presents Artificial General Intelligence-inspired Multi-Agent Knowledge Transfer (AGI-MAKT), a novel self-organizing framework for UAV cluster communication networks. Our framework overcomes existing limitations through several key innovations: a multi-modal perception system that enables comprehensive environmental understanding by fusing heterogeneous data sources; a hierarchical multi-agent learning architecture that reduces decision-making complexity while enhancing coordination; an adaptive multi-objective optimization approach that dynamically balances coverage, energy efficiency, and throughput based on contextual requirements; and efficient knowledge transfer mechanisms that accelerate learning in new environments. Comprehensive experiments demonstrate that AGI-MAKT consistently outperforms representative baselines across the evaluated performance metrics while maintaining strong robustness under challenging conditions. This research establishes a foundation for next-generation aerial networks capable of self-organization and adaptation in complex, resource-constrained environments with limited human supervision.

Keywords: Adaptive Multi-Objective Optimization; Artificial General Intelligence; Knowledge Transfer; Multi-Agent Learning; Self-Organization; Space-Air-Ground Integrated Networks; UAV Communication Networks

1. Introduction

With the rapid development of wireless communication technology and artificial intelligence (AI), Space-Air-Ground Integrated Networks (SAGIN) has become an important technical approach for building global coverage, high-capacity, and low-latency communication systems [1]. As a key component of SAGIN, Unmanned Aerial Vehicles (UAVs) demonstrate tremendous potential in emergency communications, disaster relief, large event coverage, and edge computing scenarios due to their flexible deployment, low cost, and high mobility [2]. In particular, UAV cluster communication networks, through collaborative work of multiple UAVs, can significantly expand coverage, enhance system capacity, and improve communication reliability, becoming an indispensable part of future network architectures [3]. However, UAV cluster communication networks face numerous challenges in complex and dynamic environments, including dynamic topology management, resource limitations, communication interference coordination, and energy efficiency optimization [4]. Traditional network optimization methods, typically based on predefined rules or simplified models, struggle to adapt to environmental changes and achieve multi-

objective balance. While AI technologies, especially deep reinforcement learning, have made progress in UAV communication network optimization, most research remains limited to single-objective optimization in specific scenarios, unable to achieve global coordination and adaptive decision-making [5]. The emergence of Artificial General Intelligence (AGI) concepts and technological breakthroughs, with cross-domain learning capabilities, knowledge transfer mechanisms, and adaptive decision-making characteristics, provides new possibilities for solving complex problems in UAV cluster communication networks [6]. AGI-inspired intelligent communication systems can integrate multi-source perception data, comprehensively analyze communication environment characteristics, and achieve global optimization of network resources through multi-agent collaborative learning, thereby improving system efficiency and reliability while ensuring communication quality [7].

UAV clusters serve many domains [8]: rapidly deployable links when ground infrastructure fails, capacity offload at crowded events [9], near-field edge computing for IoT devices [10], wide-area monitoring [11], and connectivity for high-speed vehicles [12]. They share hard constraints, topology churn [13], size-weight-power limits [14], spectrum pressure in dense scenes [15], short battery life [16], security exposure [17], and the need to jointly optimise coverage, capacity, delay, and energy under change [18], [19]. Traditional AI handles these only in part, whereas the multi-source integration, knowledge transfer, and multi-agent collaboration associated with AGI point toward genuine self-organisation [20].

These limitations are structural. Conventional methods optimise a single objective in a fixed scenario and trade off competing goals poorly [21]; models tuned to one distribution generalise weakly when users or obstacles move [22]; and decisions are taken in isolation, so coupled choices such as positioning and power control go uncoordinated [23]. They also depend on hand-designed state, action, and reward specifications [24], even while advancing narrow tasks such as resource allocation and interference management [25].

AGI offers four relevant capabilities (Table 1): cross-domain learning and knowledge transfer that reuse models across environments and cut training cost [26]; multi-modal integration that fuses visual, signal, behavioural, and geographical data into situational awareness [27]; online generation and balancing of competing objectives in place of fixed objective functions [28]; and causal reasoning that traces performance problems to their source and adapts over the network lifecycle [29].

To avoid overstating these capabilities, we use the term AGI-inspired rather than AGI throughout this paper. We do not claim general intelligence in the strong sense of a system matching human cognition across arbitrary tasks. Instead, we borrow four operational properties commonly associated with general intelligence and realise them within a bounded engineering setting: cross-domain transfer, in which a policy learned in one deployment environment is reused in another through knowledge distillation; multi-modal grounding, in which visual, radio-frequency, and positional observations are fused into a single state representation; self-directed objective balancing, in which the relative weight of coverage, energy, and throughput is adjusted online rather than fixed in advance; and language-grounded reasoning, in which a pre-trained language model converts a high-level mission description into structured task constraints. Each property is measurable through the experiments of Section 5, and none requires capabilities beyond contemporary learning and optimisation methods. The comparison summarised in Table 1 is synthesised from prior surveys of artificial intelligence for aerial networks [26], [29] and is intended as qualitative positioning rather than a quantitative benchmark.

Table 1. Comparison of Traditional AI and AGI Applications in UAV Communication Networks (adapted from [26], [29])

Application Scenarios	Traditional AI Methods	AGI Methods
UAV Deployment and Positioning	Reinforcement learning algorithms based on predefined objective functions (e.g., coverage)	Adaptive deployment systems integrating multi-source data, capable of balancing coverage, energy consumption, interference, and other objectives
Resource Allocation and Scheduling	Optimization algorithms for specific QoS metrics, such as DQN or multi-agent Q-learning	Dynamic resource management understanding service requirements, capable of predicting traffic variations and proactively adjusting resource configurations
Interference Management	Avoidance or elimination strategies based on specific interference models	Adaptive management systems perceiving and understanding interference source characteristics, capable of integrating multiple interference suppression techniques
Heterogeneous Network Coordination	Inter-layer collaboration mechanisms based on preset protocols	Autonomous coordination systems understanding different network layer capabilities and limitations, capable of dynamically adjusting resource allocation across layers

Network Fault Handling	Recognition and recovery mechanisms based on known fault patterns	Fault diagnosis and self-healing systems based on causal reasoning, capable of addressing unknown fault types
------------------------	---	---

This paper presents a self-organizing framework for UAV cluster communications based on AGI, addressing the complex challenges in dynamic aerial networks within SAGIN. We introduce AGI-MAKT (Multi-Agent Knowledge Transfer), an integrated architecture that leverages AGI capabilities to overcome the limitations of traditional AI approaches in communication networks. Our framework incorporates multi-modal environmental perception, knowledge transfer mechanisms, and adaptive multi-objective optimization to enable UAV clusters to autonomously coordinate and optimize their communication performance across diverse scenarios with limited human supervision. The hierarchical multi-agent learning structure effectively manages the complexity of UAV cluster deployments while ensuring efficient resource allocation and coordination. Through comprehensive simulation experiments across various environments and network conditions, we demonstrate that our approach outperforms representative baselines in coverage quality, throughput, energy efficiency, and robustness to environmental changes.

The contributions of this work can be summarised as follows. First, we formulate UAV cluster self-organisation as a single multi-objective control problem that couples deployment, transmit power, spectrum, and trajectory decisions, and we make the formulation reproducible through the notation of Table 2 and the constraints of Section 2.4. Second, we propose AGI-MAKT, whose original element is a hierarchical multi-agent learner coupled with a structured three-tier knowledge-transfer mechanism, comprising a local buffer, a regional pool, and a global repository, that distils high-value experience across scenarios; this transfer mechanism, rather than any single network, is the principal source of the observed gain, as the ablation in Section 5.5 confirms. Third, we introduce an adaptive objective-weighting rule that adjusts the relative importance of coverage, energy, and throughput online from the perceived state, in contrast to the fixed scalarisation used by QMIX, MADDPG, and MAPPO. Fourth, we report an evaluation against seven baselines under static, dynamic, and emergency conditions, together with ablation and scalability studies, so that the effect of each module can be isolated. Unlike federated reinforcement learning, which shares gradients to train a common policy, AGI-MAKT shares distilled experience while retaining role-specific policies, which is what allows it to adapt to new environments without retraining from scratch. Although the framework touches deployment, power, spectrum, and trajectory, its focus and primary contribution is AGI-inspired multi-agent knowledge transfer for UAV cluster communication self-organisation; the remaining elements are included only insofar as they are needed to evaluate that contribution.

2. System Model and Problem Definition

Table 2. Summary of Key Notation

Symbol	Description
$G(t) = (V, E)$	Time-varying directed communication graph; node set V , edge set E
N, M	Number of UAVs and number of ground users
h_{ij}	Channel gain between nodes i and j
K	Rician K -factor (line-of-sight to scattered power ratio)
γ_i, γ_{th}	SNR of user i and minimum SNR threshold for service
$\mathbf{q}_i, \mathbf{v}_i$	3-D position and flight velocity of UAV i
P_i^t	Transmit power of UAV i
B, b_i	System bandwidth and bandwidth allocated to UAV i
$E_{hov}, E_{fly}, E_{com}$	Hovering, flight, and communication energy
C, R_i	Coverage rate and achievable rate of user i
η	Energy efficiency (bits/Joule)
$\mathbf{s}_t, \mathbf{a}_t, \mathbf{r}_t$	Global state, joint action, and reward at time t
π_{θ}	Policy network with parameters θ
$\omega_1, \omega_2, \omega_3$	Adaptive weights for coverage, energy, and throughput
α, γ	Learning rate and discount factor

2.1. UAV Cluster Communication Network Model

The UAV cluster communication network can be represented as a dynamic graph structure $G(V, E, t)$, where V denotes the set of nodes, E represents the set of links, and t is the time variable. In this model, as

shown in Figure 1, nodes are classified into three categories: UAV nodes ($V_{UAV} = \{UAV_1, UAV_2, \dots, UAV_n\}$) with mobility capabilities in the air, ground control stations ($V_{GVS} = \{GCS_1, GCS_2, \dots, GCS_m\}$) that are fixed or have limited mobility on the ground, and optional relay nodes that extend communication range in air or on ground. The network topology can adopt various configurations including star topology where all UAVs communicate directly with a central control node, mesh topology where UAVs form multi-hop communication networks, hierarchical topology that divides UAV clusters into multiple levels with upper layers coordinating and lower layers executing specific tasks, and hybrid topology that dynamically adjusts based on mission requirements. Each node is characterized by attributes including three-dimensional spatial coordinates (x, y, z), mobility parameters (speed, acceleration, and heading angle), communication capabilities (communication range, available bandwidth, and supported protocols), and energy status (remaining power and consumption model).

We adopt a time-varying directed graph rather than a static or undirected one for two reasons. First, UAV mobility and intermittent blockage make the link set change over time, so a static topology cannot represent the network faithfully. Second, air-to-air and air-to-ground links are frequently asymmetric in transmit power and antenna gain, and an uplink may exist without the corresponding downlink; a directed graph captures this asymmetry, whereas an undirected graph would impose the unrealistic assumption of reciprocal links. The structure also lets the connectivity requirement (26) be expressed directly as a property of the time-varying adjacency matrix.

The communication links in the UAV cluster network are classified into air-to-air (A2A) links between UAVs, air-to-ground (A2G) links connecting UAVs with ground stations, and ground-to-ground (G2G) links between ground control equipment. The channel characteristics incorporate path loss models combining free space path loss with additional fading factors, where d represents communication distance, f denotes carrier frequency, and c is the speed of light.

$$L(d) = L_{fs}(d) + X_\sigma \quad (1)$$

$$L_{fs}(d) = 20 \log_{10}(4\pi df/c) \quad (2)$$

Small-scale fading is described using the Rician fading model [29], where K defines the Rician factor, h_{LOS} denotes the line-of-sight component, and h_{NLOS} is the non-line-of-sight component.

$$h = \sqrt{\frac{K}{K+1}} h_{LOS} + \sqrt{\frac{1}{K+1}} h_{NLOS} \quad (3)$$

Doppler effect is considered [29] due to high-speed UAV movement, where v is the relative velocity, θ defines the angle of incidence, and λ represents the wavelength.

$$f_d = \frac{v \cdot \cos \theta}{\lambda} \quad (4)$$

Link quality is evaluated through signal-to-noise ratio (SNR) [14], where P_r represents the received power, N_0 denotes the noise power spectral density, and B defines the bandwidth, along with bit error rate (BER) related to modulation method and channel conditions, and link availability based on link interruption probability $\gamma = \frac{P_r}{N_0 \cdot B}$. The interference model accounts for co-frequency interference from other UAV communications on the same frequency and external electromagnetic interference, with total interference power expressed as follows, where P_i is the transmission power of the i -th interference source and $G(d_i)$ is the distance-dependent propagation gain:

$$I = \sum_i P_i \cdot G(d_i) \quad (5)$$

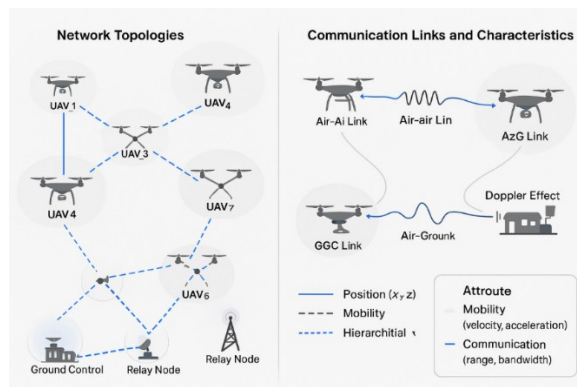


Figure 1. UAV Cluster Communication Network Model illustrating network topology, node classifications, and communication links with channel characteristics

2.2. Energy Consumption Model

Energy is a primary bottleneck for UAVs, so we model consumption across the three operational modes: hovering, flight, and communication.

Hovering is a fundamental state [30] during UAV mission execution, where the UAV must counteract gravity to maintain a fixed position. The energy consumption during hovering primarily stems from the lift provided by the rotor system and can be expressed as:

$$P_{\text{hover}} = \frac{(mg)^{3/2}}{2\rho A} \quad (6)$$

where m represents the UAV mass, g is the gravitational acceleration, ρ denotes air density, and A is the total rotor area. As evident from the equation, UAV's mass significantly influences hovering energy consumption, with increases in mass leading to non-linear growth in energy expenditure. Additionally, reduced air density at higher altitudes further increases energy consumption.

During flight, UAV energy consumption [30] includes not only the energy required to maintain lift but also the energy expended by the propulsion system to overcome air resistance. According to research literature, flight energy consumption can be modelled as:

$$P_{\text{flight}} = P_{\text{hover}} \left(1 + \frac{3v^2}{U_{\text{tip}}^2}\right) \quad (7)$$

where v represents the UAV flight speed and U_{tip} is the rotor tip speed. This model indicates that flight speed significantly impacts energy consumption, and there exists an optimal speed for energy efficiency. As was noted, energy efficiency can be defined as a function of communication coverage, fairness, energy consumption, and connectivity, with DRL methods enabling trajectory optimization to maximize energy efficiency [30].

Trajectory design strongly affects energy use; for example, PPO is used to plan energy-minimising trajectories while maintaining communication quality [31].

When executing communication tasks, UAVs must consider the energy consumption of the communication system itself, in addition to flight and hovering energy expenditure:

$$P_{\text{comm}} = P_{\text{tx}} + P_{\text{rx}} + P_{\text{proc}} \quad (8)$$

where P_{tx} represents transmission power, P_{rx} denotes reception power, and P_{proc} accounts for signal processing power consumption. Communication energy consumption is closely related to factors such as transmission distance, data rate, communication protocols, and environmental interference.

2.3. User Distribution and Quality of Service Model

In UAV cluster communication networks, user density distribution is modelled as a Gaussian mixture [5] and is a key factor affecting network deployment and resource allocation. User distribution typically exhibits spatial-temporal heterogeneity and can be modelled through probability density functions. In urban environments, users usually present clustering distribution characteristics, which can be described using a Gaussian mixture model:

$$f(x, y) = \sum_{i=1}^K \omega_i N(\mu_i, \Sigma_i) \quad (9)$$

where K represents the number of cluster centers, ω_i is the weight coefficient of the i -th cluster, μ_i and Σ_i are the mean vector and covariance matrix of the i -th Gaussian distribution, respectively. User density distribution also exhibits dynamic characteristics, such as short-term high-density gatherings during large events and flowing distributions during commuting hours. Dynamic prediction of user density can be achieved through time series analysis and deep learning methods, providing a basis for adaptive deployment of UAV clusters.

Quality of Service (QoS) constraints for UAV cluster communication networks primarily include transmission rate, delay, reliability, and fairness across multiple dimensions. For any user u , the basic QoS constraints can be expressed as:

$$R_u = B \log_2 (1 + \gamma_u) \geq R_{\min} \quad (10)$$

where B is the available bandwidth, γ_u is the SNR for user u , and $R_{\min}$ is the minimum transmission rate requirement.

The end-to-end delay D_u must satisfy:

$$D_u = D_{\text{trans}} + D_{\text{proc}} + D_{\text{queue}} \leq D_{\max} \quad (11)$$

where D_{trans} is the transmission delay, D_{proc} is the processing delay, D_{queue} is the queuing delay, and D_{max} is the maximum acceptable delay.

The reliability of the communication link can be represented through outage probability:

$$P_{out} = P(\gamma_u < \gamma_{th}) \leq P_{out,max} \quad (12)$$

where γ_{th} is the minimum SNR threshold, and $P_{out,max}$ is the maximum acceptable outage probability. The difference in service quality between different users in the network must satisfy:

$$\frac{\max_{u \in U} R_u - \min_{u \in U} R_u}{\min_{u \in U} R_u} \leq \eta \quad (13)$$

where U is the set of users, and η is the fairness threshold. AGI systems need to dynamically adjust the priorities of these constraints based on current network conditions and user requirements to achieve adaptive optimization of service quality.

The system capacity of UAV cluster communication networks is influenced by various factors, including the number of UAVs, communication resources, interference levels, and user distribution. Based on information theory, the total system capacity can be expressed as:

$$C_{sys} = \sum_{i=1}^N \sum_{j=1}^M \alpha_{i,j} B_{i,j} \log_2(1 + \gamma_{i,j}) \quad (14)$$

where N is the number of UAVs, M is the number of users, $\alpha_{i,j}$ is the resource allocation indicator variable, $B_{i,j}$ is the bandwidth allocated for communication between UAV i and user j , and $\gamma_{i,j}$ is the SNR of the corresponding link.

System capacity depends on UAV number and density, on spectrum, power, and time allocation, and on interference, with an optimal density beyond which interference dominates. Within the framework these are handled jointly: multi-agent learning searches for globally efficient resource allocation, beamforming and coordinated transmission suppress interference, and the topology is adapted online to the user distribution.

2.4. Multi-objective Optimization Problem Formalization

Building on the network, energy, and service-quality models above, we now formalise the multi-objective optimisation problem, which couples coverage maximisation, energy-efficiency optimisation, and throughput improvement.

The coverage rate is an important indicator for measuring the service capability of UAV cluster networks, representing the proportion of the area that can provide effective communication services to ground users relative to the total target area. Assuming the target service area is A , the set $U = \{u_1, u_2, \dots, u_m\}$ represents the users in the area, and the set $D = \{d_1, d_2, \dots, d_n\}$ represents the deployed UAVs, the coverage rate can be defined as:

$$\eta_{cov} = \frac{1}{|U|} \sum_{u \in U} I(\gamma_u \geq \gamma_{th}) \quad (15)$$

where γ_u represents the SNR of user u , γ_{th} is the minimum SNR threshold required for providing effective communication services, and $I(\cdot)$ is an indicator function that takes the value 1 when the condition is satisfied and 0 otherwise. The SNR of user u can be expressed as:

$$\gamma_u = \frac{P_{r,u}}{N_0 + \sum_{i \neq s} I_i} \quad (16)$$

where $P_{r,u}$ is the signal power received by user u , N_0 is the noise power, and $\sum_{i \neq s} I_i$ represents the interference power from other transmission sources.

The objective of coverage maximization is $\max_{X,P} \eta_{cov}$, where $X = \{x_1, x_2, \dots, x_n\}$ represents the position coordinates of the UAVs, and $P = \{P_1, P_2, \dots, P_n\}$ represents the transmission power of the UAVs. Considering the limited energy supply of UAVs, energy efficiency becomes a key factor for the long-term operation of the system. Energy efficiency is defined as the amount of effective communication service provided per unit of energy consumption, which can be expressed as:

$$\eta_{ee} = \frac{\sum_{u \in U} R_u}{\sum_{d \in D} (P_{hover,d} + P_{flight,d} + P_{comm,d})} \quad (17)$$

where R_u represents the transmission rate of user u , and $P_{hover,d}$, $P_{flight,d}$ and $P_{comm,d}$ represent the hovering energy consumption, flight energy consumption, and communication energy consumption of UAV d , respectively, as described earlier. According to the Shannon capacity formula, the transmission rate of user u can be calculated as:

$$R_u = B \log_2 (1 + \gamma_u) \quad (18)$$

where B is the system bandwidth.

The objective of energy efficiency optimization is $\max_{X,P,V} \eta_{ee}$, where $V = \{v_1, v_2, \dots, v_n\}$ represents the flight velocity of the UAVs. System throughput refers to the total amount of data that the network can transmit per unit time, which is an important indicator for evaluating network performance. The total system throughput can be expressed as the sum of the transmission rates of all users:

$$\eta_{tp} = \sum_{u \in U} R_u = \sum_{u \in U} B \log_2 (1 + \gamma_u) \quad (19)$$

The objective of system throughput improvement is $\max_{X,P,B} \eta_{tp}$, where $B = \{B_1, B_2, \dots, B_n\}$ represents the bandwidth resources allocated to each UAV. Combining the above three optimization objectives, the multi-objective optimization problem for UAV cluster communication self-organization can be formalized as:

$$\max_{X,P,V,B} \{\eta_{cov}, \eta_{ee}, \eta_{tp}\} \quad (20)$$

Subject to the following constraints:

(1) UAV position constraints: UAVs must operate within the specified airspace:

$$x_{min} \leq x_d \leq x_{max}, y_{min} \leq y_d \leq y_{max}, z_{min} \leq z_d \leq z_{max}, \forall d \in D \quad (21)$$

(2) Transmission power constraints: The transmission power of UAVs cannot exceed the maximum allowed value:

$$0 \leq P_d \leq P_{max}, \forall d \in D \quad (22)$$

(3) Bandwidth resource constraints: The system's total bandwidth resources are limited

$$\sum_{d \in D} B_d \leq B_{total} \quad (23)$$

(4) Energy constraints: UAV energy consumption cannot exceed its energy capacity

$$E_d = \int_0^T (P_{hover,d} + P_{flight,d} + P_{comm,d}) dt \leq E_{capacity}, \forall d \in D \quad (24)$$

(5) QoS constraints: Meeting the service quality requirements of users

$$R_u \geq R_{min}, D_u \leq D_{max}, \forall u \in U \quad (25)$$

where R_{min} is the minimum transmission rate requirement, and D_{max} defines the maximum acceptable delay.

(6) Network connectivity constraints: The UAV cluster needs to maintain network connectivity

$$\forall d_i, d_j \in D, \exists \text{ path connecting } d_i \text{ and } d_j \quad (26)$$

These three objectives conflict, since wider coverage may demand more UAVs or power and so lower energy efficiency; the problem is therefore multi-objective. One standard solution is the weighted-sum method, which combines the objectives into one:

$$\max f = \omega_1 \eta_{cov} + \omega_2 \eta_{ee} + \omega_3 \eta_{tp} \quad (27)$$

where $\omega_1, \omega_2, \omega_3$ are weight coefficients satisfying $\omega_1 + \omega_2 + \omega_3 = 1$.

Another approach is the objective transformation method, which selects one main objective and transforms the others into constraints:

$$\begin{aligned} &\max \eta_{cov} \\ &\text{s.t. } \eta_{ee} \geq \eta_{ee,min}, \eta_{tp} \geq \eta_{tp,min} \end{aligned} \quad (28)$$

Pareto optimisation seeks solutions in which no objective can improve without another worsening; we solve it with the AGI-inspired multi-agent learning framework through collaborative learning and knowledge transfer between agents.

3. AGI-inspired Self-Organization Framework Design

3.1. Framework Overall Architecture

The AGI-inspired self-organization framework is the central contribution of this work; among its five modules, the perception and execution modules adapt established components, whereas the originality lies in coordinating hierarchical learning with the three-tier knowledge-transfer mechanism and the adaptive objective-weighting rule described below. This section details the overall architecture of the framework, including system functional modules, data flow and interaction mechanisms, and hierarchical learning structure, to comprehensively present the design philosophy and technical characteristics of this framework.

The AGI-inspired self-organization framework consists of five core functional modules. The perception module is responsible for collecting, preprocessing, and fusing multi-source heterogeneous data, providing the system with comprehensive environmental perception capabilities. This includes multi-sensor data acquisition units integrating optical, radar, acoustic, and other sensor data; signal processing units for noise reduction, filtering, and feature enhancement; and data fusion units for spatiotemporal alignment and information integration across different sources. The cognitive module implements environment and task understanding and modelling, constructing semantic-level scene representations through scene understanding, semantic analysis, and knowledge representation functions that extract high-level semantic information and establish structured knowledge graphs and scene models.

The decision module, the core of the framework, generates strategies and plans through multimodal reasoning, decomposes high-level goals into executable tasks, and allocates computation, communication, and energy across the cluster. The coordination module assigns roles, manages the distributed learning and decision protocols, and resolves conflicts between agents, while the execution module converts decisions into control commands and adjusts them from real-time feedback.

The reference to a large language model in the decision module deserves clarification, since it is the only component that draws directly on general-purpose pre-training. We use a compact pre-trained language model in an inference-only, read-only capacity: it receives a textual mission brief, for example a request to prioritise coverage of the eastern sector during an evacuation, together with a serialised summary of the current scene graph, and it returns a structured set of objective priorities and hard constraints that seed the adaptive weighting rule of Section 4.4. The language model does not actuate UAVs directly and is not trained online; its outputs are checked against the feasibility constraints of Section 2.4 before use. This keeps the language component bounded and auditable while still allowing high-level instructions to be expressed in natural language.

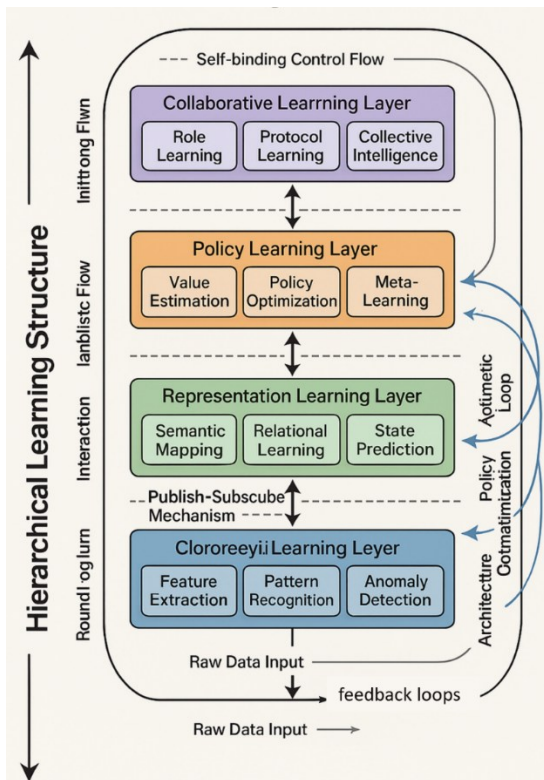


Figure 2. Hierarchical learning structure of AGI-based UAV clusters communication self-organization framework

The framework uses a bidirectional data flow. A bottom-up perception path carries raw sensor data through preprocessing, feature fusion, and semantic understanding into a unified state estimate, while a top-down control path turns decisions into instructions and returns execution results to the decision layer.

Decisions are formed by generating and evaluating candidate strategies through multimodal reasoning, reaching consensus by multi-agent negotiation, and decomposing global plans into per-agent tasks; a publish-subscribe middleware carries asynchronous messages between modules for rapid response.

The framework also learns hierarchically. A perception-learning layer extracts features, recognises patterns, and flags anomalies, while a representation-learning layer maps these features to semantic concepts, learns relations between entities, and predicts future environmental and system states.

Above these, a strategy-learning layer evaluates states and actions, optimises decisions for long-term return, and meta-learns to adapt faster, and a collaborative-learning layer refines role assignment, protocols, and collective behaviour. Feedback loops act at three-time scales, fast adaptation in the lower layers, medium-term policy improvement, and slow architectural evolution, so the layers work as one whole spanning perception to decision (Figure 2).

Together, these modules, data flows, and the hierarchical learner give the cluster a closed perceive-learn-decide-collaborate loop for complex, dynamic environments, on which the algorithms of the following sections build.

3.2. AGI-inspired Environmental Perception and State Representation

Environmental perception and state representation are foundational, letting the cluster understand complex communication environments. Our approach improves on traditional perception through multi-modal integration, semantic understanding, and hierarchical state representation.

3.2.1. Multi-source Data Collection and Processing

The proposed framework implements a heterogeneous data integration system that fuses information from multiple complementary sources. Data from UAV-mounted sensors undergoes a three-stage preprocessing pipeline. First, adaptive noise filtering is applied using a modified Kalman filter structure:

Concretely, the perception layer ingests the three complementary streams shown in Figure 3. Visual data are RGB and depth frames from a downward-facing camera, sampled at 10 Hz and used to detect users, obstacles, and free space. Signal data comprise per-link received-signal-strength, SINR, and spectrum-occupancy measurements reported by the onboard radio at the slot rate. Location data consist of the UAV's own GNSS and inertial pose together with the estimated positions of served users. The symbols associated with these quantities are listed in Table 2. Because the streams differ in rate and dimension, the alignment and fusion steps below are required before they enter the state vector of (35).

$$\hat{x}_t = F_t \hat{x}_{t-1} + K_t (z_t - H_t F_t \hat{x}_{t-1}) \quad (29)$$

where $\hat{x}_t$ represents the filtered state estimation, z_t is the raw measurement, and K_t is the adaptive Kalman gain that adjusts based on environmental dynamics.

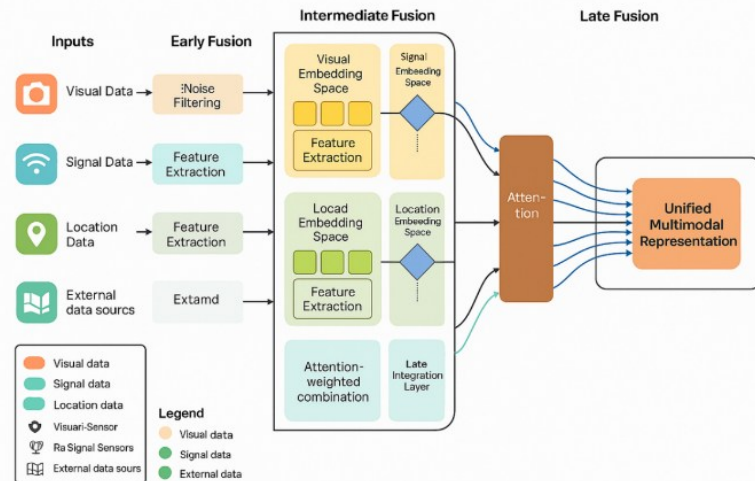


Figure 3. Hierarchical Multi-modal Fusion Architecture for GAR Communication Environment Perception in UAV Systems

Multi-modal data alignment addresses spatiotemporal inconsistencies through a cross-correlation function $C_{\{ij\}}(\tau)$ between data streams i and j with time offset τ :

$$C_{ij}(\tau) = \sum_t D_i(t) D_j(t + \tau) \quad (30)$$

The optimal time offset τ^* maximizes this correlation. For spatial alignment, we utilize homographic transformations based on identified feature correspondences. Our context-aware sampling approach

dynamically allocates sensing resources according to information gain metrics. The sampling rate for sensor i at time t is determined by:

$$r_i(t) = r_{base} + \Delta r \cdot IG_i(t) \quad (31)$$

where $IG_i(t)$ represents the estimated information gain from sensor i computed using entropy reduction measures.

Multi-modal fusion combines early, intermediate, and late techniques (Figure 3): deep multi-modal embedding networks map heterogeneous inputs into a common space, which are then merged by attention-weighted combination.

3.2.2. Scene Understanding and Semantic Analysis

Our scene understanding methodology transforms low-level sensor data into semantic environmental representations through a multi-stage process. Environmental element detection employs a modified transformer-based architecture that identifies and classifies key components relevant to communication quality:

$$F_{element} = \text{Transformer}\theta(F_{input}, M_{attn}) \quad (32)$$

where F_{input} represents multi-modal input features, M_{attn} is an attention mask emphasizing communication-relevant regions, and θ denotes the model parameters fine-tuned for UAV communication scenarios. In (32) the input is written F_{input} and the output $F_{element}$, with the subscripts harmonised so that both the equation and this text use the same notation: F_{input} is the fused multi-modal input feature tensor, M_{attn} the attention mask, θ the model parameters, and $F_{element}$ the resulting element-level semantic features.

We construct a communication-specific knowledge graph $G = (V, E, A)$, where vertices V represent environmental entities, edges E capture their relationships, and A denotes attribute sets. Each edge $e_{ij} \in E$ connecting entities i and j is associated with a relationship type r_{ij} and an impact weight w_{ij} quantifying its effect on communication performance.

The semantic mapping process associates elements detected with communication-relevant concepts through a transfer learning approach. Pre-trained on diverse communication environments, our model generalizes to new scenarios through:

$$P(c_i|e_j) = \frac{\exp(f_\phi(e_j, c_i))}{\sum_k \exp(f_\phi(e_j, c_k))} \quad (33)$$

where $P(c_i|e_j)$ represents the probability of element e_j belonging to semantic concept c_i , and f_ϕ is a learned compatibility function.

Communication context interpretation utilizes a multi-head attention mechanism to identify scenario-specific requirements:

$$\text{Context} = \text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_h)W^O \quad (34)$$

Each head attends to a different aspect of the environment, distinguishing emergency, high-density, and routine scenarios.

3.2.3. State Space Construction and Feature Extraction

The framework constructs a comprehensive state representation that captures all factors relevant to UAV communication optimization. The state space S can be formally represented as:

$$S = S_{comm} \times S_{UAV} \times S_{env} \quad (35)$$

where S_{comm} represents communication-specific state variables including spectrum availability maps, user distribution matrices, and interference patterns S_{UAV} captures UAV-specific states including position vectors, energy levels, and link quality metrics S_{env} encompasses environmental factors affecting communication performance

Our information-theoretic feature selection approach identifies the most relevant state variables by computing the mutual information $I(X; Y)$ between candidate features X and communication performance metrics Y :

$$I(X; Y) = \sum_{x \in X} \sum_{y \in Y} p(x, y) \log \frac{p(x, y)}{p(x)p(y)} \quad (36)$$

Features with mutual information values exceeding a threshold τ are retained in the final state representation.

To address the high dimensionality of the combined state space, we implement hierarchical dimensionality reduction techniques. For linear relationships, we apply weighted principal component analysis $\tilde{S} = SW_{\text{PCA}}$. For capturing complex nonlinear patterns, we employ a variational autoencoder structure that learns a compressed latent representation z while preserving communication-critical information:

$$z = \text{Encoder}\theta(S), \tilde{S} = \text{Decoder}\phi(z) \quad (37)$$

State representation spans temporal and spatial scales: fast-changing local conditions are encoded at the micro level and slower global patterns at the macro level, supporting decisions from link adaptation to resource allocation.

Our state completion algorithm addresses missing data issues by leveraging spatial and temporal correlations:

$$\hat{s}_i = \sum_{j \in \mathcal{N}(i)} w_{ij} s_j + \sum_{t' \in \mathcal{T}(t)} w_{t'} s_i^{t'} \quad (38)$$

where $\mathcal{N}(i)$ represents the neighborhood of state variable i in the feature space, and $\mathcal{T}(t)$ denotes the relevant temporal window. This subsection therefore provides the methodological realisation of the multi-modal perception innovation stated in the abstract. The fused visual, radio, positional, and environmental streams are not only listed as inputs but are converted by Eqs. (29)-(38) into a communication-aware state vector, so multi-modal perception supplies AGI-MAKT with the coverage, interference, mobility, and energy context used by the later learning and optimisation modules.

3.3. Multi-Agent Learning and Coordination Framework

Building on the perception of Section 3.2, the framework adds multi-agent learning and coordination for collaborative intelligence across the cluster, detailed below through agent roles, distributed learning, knowledge sharing, and collaborative decision-making.

3.3.1. Agent Role Definition and Division of Labor

The framework uses a heterogeneous multi-agent structure in which UAVs take different roles by mission need and capability. Role assignment is posed as an optimisation problem:

$$\mathcal{R}^* = \arg \max_{\mathcal{R}} U(\mathcal{R}|S, \mathcal{C}) \quad (39)$$

where $\mathcal{R} = \{r_1, r_2, \dots, r_n\}$ represents the role assignment vector for n UAVs, S is the current state representation derived from Section 3.2, $\mathcal{C}$ denotes the cluster capabilities matrix, and $U(\cdot)$ is the utility function measuring the effectiveness of a particular role distribution.

Each UAV can assume one of several pre-defined roles with specific responsibilities $r_i \in \{\text{Coordinator}, \text{Relay}, \text{Coverage}, \text{Sensing}, \text{Edge} - \text{Computing}\}$. The role transition dynamics are governed by a state machine where transitions depend on environmental conditions, mission phase, and resource availability. For each agent i , the role transition probability is defined as:

$$P(r_i^{t+1}|r_i^t, S^t) = \frac{\exp(Q(r_i^t, r_i^{t+1}, S^t))}{\sum_{r' \in \mathcal{R}} \exp(Q(r_i^t, r', S^t))} \quad (40)$$

Where $Q(r_i^t, r_i^{t+1}, S^t)$ represents the value of transitioning from role r_i^t to r_i^{t+1} in state S^t . Role-specific policy networks are implemented to optimize agent behavior within each role. For instance, the policy network for the coordinator role focuses on global resource allocation, while Coverage agents optimize positioning for maximum signal coverage. This specialization allows more efficient learning compared to a single policy attempting to master all aspects of the UAV cluster operation.

3.3.2. Distributed Learning Protocol

The geographically distributed nature of UAV clusters necessitates distributed learning protocols that balance computational efficiency, communication overhead, and learning stability. Our framework implements a hierarchical federated reinforcement learning approach that adapts to communication constraints. The learning process is formalized as a decentralized partially observable Markov decision process (Dec-POMDP), where each agent maintains a local policy π_i based on its observation history $o_i^{1:t}$:

$$a_i^t \sim \pi_i(a_i^t|o_i^{1:t}) \quad (41)$$

The distributed learning protocol operates in three phases:

Local Learning: Each agent updates its policy based on local experiences using PPO [37]:

$$\mathcal{L}^{ppo}(\theta) = \mathbb{E}_t[\min(r_t(\theta)\hat{A}_t, \text{clip}(r_t(\theta), 1 - \epsilon, 1 + \epsilon)\hat{A}_t)] \quad (42)$$

where $r_t(\theta) = \frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{old}}(a_t|s_t)}$ is the probability ratio and $\hat{A}_t$ is the advantage estimate.

Secure Model Aggregation: Periodically, agents share policy updates using a secure aggregation protocol that preserves privacy while enabling collaborative learning:

$$\theta_{global} = \frac{1}{N} \sum_{i=1}^N w_i \cdot \theta_i \quad (43)$$

where w_i represents the importance weight of agent i based on its collected experiences.

Knowledge Distillation: Global knowledge is distilled back to individual agents while preserving role-specific expertise:

$$\mathcal{L}_{distill} = DKL(\pi_{\theta_{global}} || \pi_{\theta_i}) + \lambda \mathcal{L}_{task} \quad (44)$$

The protocol adapts to connectivity by varying synchronisation frequency and model compression; under poor links, agents operate independently and accumulate experience for later synchronisation.

3.3.3. Knowledge Sharing and Experience Pool Design

To accelerate learning and enable cross-scenario knowledge transfer, our framework implements a structured knowledge sharing mechanism centered around a distributed experience pool. The experience pool aggregates diverse experiences from all agents while organizing them according to their contextual relevance and value. Experiences are represented as episodic memory units $M = \{(s_t, a_t, r_t, s_{t+1}, c_t)\}$, where c_t represents contextual metadata that enables efficient experience retrieval. We employ a hierarchical experience pooling structure with three levels:

Local Experience Buffer: Each agent maintains recent experiences for immediate learning.

Regional Experience Pool: Agents in proximity share experiences relevant to their current operational region.

Global Experience Repository: High-value experiences are selectively promoted to a central repository for long-term retention and cross-scenario transfer.

Experience sampling for learning follows a prioritized approach that balances diversity, recency, and significance $P(M_i) = \frac{\omega_i^\alpha}{\sum_j \omega_j^\alpha}$, where ω_i is the importance weight combining surprise (temporal difference error), rarity, and relevance to the current context:

$$\omega_i = \beta_1 |\delta_i| + \beta_2 R(M_i) + \beta_3 S(M_i, c_t) \quad (45)$$

To enable cross-scenario knowledge transfer, we implement a contrastive learning approach that identifies invariant features across different communication environments. The transfer learning objective is defined as:

$$\mathcal{L}_{transfer} = \mathcal{L}_{task} + \lambda_1 \mathcal{L}_{contrastive} + \lambda_2 \mathcal{L}_{invariance} \quad (46)$$

This knowledge sharing architecture enables efficient learning from collective experiences while maintaining adaptation to local conditions. This subsection realises the hierarchical multi-agent learning and knowledge-transfer innovations stated in the abstract. The role-specific policies of Eqs. (39)-(44) reduce decision complexity by separating coordinator and task-agent behaviour, while the local buffer, regional pool, and global repository in Eqs. (45)-(46) define how distilled high-value experience is reused across regions and scenarios. This is the mechanism expected to accelerate adaptation in new environments, and it is tested directly by the knowledge-transfer ablation in Section 5.5.

3.3.4. Collaborative Decision-Making and Conflict Resolution

Coordinated decisions must optimise system objectives while respecting agent autonomy; we use a multi-level consensus with a weighted-voting mechanism:

$$a_{consensus} = \arg \max_a \sum_{i=1}^N \alpha_i V_i(a|s) \quad (47)$$

where $V_i(a|s)$ represents agent i 's evaluation of action a in state s , and α_i is the influence weight determined by role, confidence, and expertise. To handle conflicting objectives between coverage, energy efficiency, and throughput optimization, we implement a Pareto front identification approach that identifies non-dominated solutions in the multi-objective space:

$$\mathcal{P} = \{x \in \Omega | \nexists y \in \Omega, y \succ x\} \quad (48)$$

where Ω is the feasible solution space and $y \succ x$ indicates that solution y dominates solution x . For dynamic conflict resolution, we develop a negotiation protocol based on alternating offers where agents

propose solutions and counter-offers until consensus is reached or a timeout occurs. The protocol is formalized as a sequential game:

$$BR_i(o_{-i}) = \arg \max_{o_i \in O_i} u_i(o_i, o_{-i}) \quad (49)$$

where BR_i is the best response of agent i to others' offers o_{-i} , and u_i is the utility function. Our conflict resolution system incorporates several mechanisms to prevent deadlocks and ensure timely decisions:

- (1) Timeout-based fallback to pre-defined hierarchy
- (2) Utility threshold relaxation over negotiation rounds
- (3) Mediation by dedicated coordinator agents

Together with the perception of Section 3.2, this learning and coordination layer underpins the self-organising behaviour of the AGI-inspired system; the next section presents the optimisation algorithms that act on it.

4. Network Performance Optimization Algorithms

Building upon the AGI-inspired self-organization framework described in the previous section, we now present specific algorithms for optimizing critical network performance metrics. This section details our novel approaches to coverage maximization, energy efficiency optimization, and throughput enhancement while maintaining system stability in dynamic environments. As shown in Algorithm 1, our approach integrates multiple optimization components into a cohesive framework that iteratively refines UAV positions, trajectories, power allocations, and spectrum usage.

Algorithm 1. AGI UAV Optimization

```

Input:  $S$  (environment state)
Output:  $X$  (UAV positions),  $P$  (transmission power allocation),  $V$  (UAV velocities/trajectories),  $B$  (bandwidth allocation)

Initialization:
Initialize optimization variables:
 $X \leftarrow \{\}$  (UAV positions)
 $P \leftarrow \{\}$  (transmission power allocation)
 $V \leftarrow \{\}$  (UAV velocities/trajectories)
 $B \leftarrow \{\}$  (bandwidth allocation)
Initialize weights for multi-objective optimization:
 $\omega \leftarrow [\omega_1, \omega_2, \omega_3]$  (Initial weights for coverage, energy, throughput)

Main Loop:
while not converged do
  Coverage - aware Deployment Optimization:
   $X \leftarrow \text{COVERAGE\_OPTIMIZATION}(S, X, P)$ 
  Energy - aware Trajectory Optimization:
   $V, P \leftarrow \text{ENERGY\_OPTIMIZATION}(S, X, V, P)$ 
  Spectrum Management and Interference Coordination:
   $B, P \leftarrow \text{SPECTRUM\_OPTIMIZATION}(S, X, P, B)$ 
  Self - adaptive Multi - objective Balancing:
   $\omega \leftarrow \text{UPDATE\_OBJECTIVE\_WEIGHT}(S, \omega)$ 
   $[X, P, V, B] \leftarrow \text{MULTI\_OBJECTIVE\_OPTIMIZATION}(S, X, P, V, B, \omega)$ 
  Convergence Check:
  if update_magnitudes_below_threshold () then
    break
  end if
end while
return  $X, P, V, B$ 

```

4.1. Coverage-aware Deployment Optimization

Optimal-coverage deployment is a core challenge. Geometric or simplified-propagation methods generalise poorly to complex environments, so we use the semantic-aware coverage optimisation of Algorithm 2.

The coverage optimization problem is formulated as finding the optimal UAV positions $\mathbf{X}^* = \{x'_1, x'_2, \dots, x'_n\}$ that maximize the effective coverage rate while considering environmental constraints:

$$\mathbf{X}^* = \arg \max_{\mathbf{X}} \eta_{cov}(\mathbf{X}, \mathbf{S}) \text{ subject to } \mathbf{X} \in \mathcal{C} \quad (50)$$

where $\mathbf{S}$ represents the environment state derived from Section 3.2, and $\mathcal{C}$ defines the set of feasible UAV positions considering flight constraints and inter-UAV distance requirements. Our solution employs a hybrid approach combining reinforcement learning with gradient-based optimization:

Environment Segmentation: The operational area is semantically segmented into regions with different communication characteristics (urban canyons, open spaces, high-user-density zones) based on the scene understanding described in Section 3.2.2.

Initial Deployment: A deep neural network trained on historical coverage data generates candidate UAV positions optimized for the current environment:

$$\hat{\mathbf{X}} = f_{\theta}(\mathbf{S}) \quad (51)$$

where f_{θ} is the deployment prediction network with parameters θ .

Local Refinement: Starting from the initial positions, gradient-based optimization refines each UAV's position using the coverage gradient:

$$x_i^{t+1} = x_i^t + \alpha \nabla_{x_i} \eta_{cov}(x_i^t, \mathbf{S}) \quad (52)$$

where α is the step size that adaptively adjusts based on convergence behavior.

Cooperative Adjustment: UAVs share coverage information and collaboratively adjust positions to minimize overlap while maximizing total coverage using the coordination mechanisms from Section 3.3.4.

Algorithm 2. Coverage-aware Deployment Optimization

Input: S (environment state), $X_{current}$ (current UAV positions), $P_{current}$ (current transmission power allocation)

Output: X (optimized UAV positions)

Semantic Environment Segmentation:

$regions \leftarrow \text{SEGMENT_ENVIRONMENT}(S)$

Initial Deployment using trained neural network:

$X \leftarrow f_{\theta}(S)$ // Deep neural network prediction

Local Gradient-based Refinement:

for each UAV i **do**

for $t = 1$ **to** $max_{iterations}$ **do**

$gradient \leftarrow \nabla_{x_i} \eta_{cov}(x_i, S)$

$x_i \leftarrow x_i + \alpha \cdot gradient$ // α : adaptive step size

If $\|gradient\| < \epsilon$ **then**

break // Convergence check

end if

end for

end for

Cooperative Position Adjustment:

$X \leftarrow \text{MINIMIZE_COVERAGE_OVERLAP } X$

Return:

return X

Algorithm 2 combines reinforcement learning with gradient-based optimisation: it segments the environment, generates initial positions with a neural network, then refines them by gradient descent and cooperative adjustment to maximise coverage with minimal overlap.

4.2. Energy-aware Trajectory and Resource Optimization

Energy efficiency is critical for extending UAV mission duration. Our energy optimization algorithm jointly considers UAV trajectories, communication power allocation, and computational resource management:

$$\min_{\mathbf{v}, \mathbf{p}, \mathbf{c}} E_{total} = \sum_{i=1}^N \int_0^T (P_{hover,i} + P_{flight,i} + P_{comm,i} + P_{comp,i}) dt \quad (53)$$

subject to coverage, throughput, and QoS constraints detailed in Section 3.4.

The solution employs a hierarchical optimization approach:

Energy Consumption Prediction: A recurrent neural network predicts energy consumption for candidate trajectories and resource allocation strategies:

$$\hat{E}_i(v_i, P_i, C_i) = g_{\phi}(v_i, P_i, C_i, \mathbf{S}) \quad (54)$$

where g_{ϕ} is the energy prediction model trained on historical operation data.

Task-aware Power Management: Communication power is dynamically allocated according to task priorities and channel conditions:

$$P_{i,j}^* = \arg \min_{P_{i,j}} P_{i,j} \quad \text{subject to } \gamma_{i,j}(P_{i,j}) \geq \gamma_{th,j} \quad (55)$$

where $\gamma_{i,j}$ is the signal-to-interference-plus-noise ratio for the link between UAV i and user j .

Energy-efficient Trajectory Planning: UAV trajectories are optimized to minimize energy consumption while maintaining coverage requirements:

$$\mathbf{V}^* = \arg \min_{\mathbf{V}} \sum_{i=1}^N \int_0^T (P_{hover,i}(v_i) + P_{flight,i}(v_i)) dt \quad \text{subject to } \eta_{cov}(\mathbf{X}(t)) \geq \eta_{min} \quad (56)$$

Computational Load Balancing: Processing tasks are distributed across the UAV cluster to balance computational load and minimize energy consumption:

$$\mathbf{C}^* = \arg \min_{\mathbf{C}} \sum_{i=1}^N P_{comp,i}(C_i) \quad \text{subject to } \sum_{i=1}^N C_i = C_{total} \quad (57)$$

The energy-aware procedure applies these steps in sequence: energy-consumption prediction, task-aware power management, energy-efficient trajectory planning, and computational load balancing.

4.3. Adaptive Spectrum Management and Interference Coordination

Efficient spectrum utilization is essential for maximizing system throughput in UAV communication networks. **As shown in the corresponding procedure**, our AGI-inspired framework implements an adaptive spectrum management algorithm that dynamically allocates frequency resources while coordinating interference:

$$\max_{\mathbf{B}, \mathbf{P}} \eta_{tp}(\mathbf{B}, \mathbf{P}, \mathbf{S}) = \sum_{i=1}^N \sum_{j=1}^M \alpha_{i,j} B_{i,j} \log_2 (1 + \gamma_{i,j}(P_{i,j})) \quad (58)$$

Our solution approach consists of the following key components:

Interference Mapping: A deep graph neural network models the interference relationships between communication links:

$$\mathbf{I} = h_{\psi}(\mathbf{X}, \mathbf{P}, \mathbf{S}) \quad (59)$$

where h_{ψ} represents the interference prediction model that considers spatial positions, transmission powers, and environmental factors.

Dynamic Spectrum Allocation: Frequency resources are allocated using a reinforcement learning approach that maximizes long-term spectral efficiency:

$$\pi_B(B_{i,j}|s_t) = \arg \max_{\pi} \mathbb{E}[\sum_{t=0}^{\infty} \gamma^t R_t] \quad (60)$$

where R_t is the reward representing achieved throughput at time t .

Coordinated Beamforming: For UAVs equipped with multi-antenna systems, beamforming vectors are optimized to minimize interference while maximizing signal strength:

$$\mathbf{w}_i^* = \arg \max_{\mathbf{w}_i} \frac{\mathbf{w}_i^H \mathbf{H}_{i,i} \mathbf{w}_i}{\sum_{j \neq i} \mathbf{w}_j^H \mathbf{H}_{j,i} \mathbf{w}_j + \sigma^2} \quad \text{subject to } \|\mathbf{w}_i\|^2 \leq P_{max} \quad (61)$$

Adaptive Modulation and Coding: Transmission parameters are dynamically adjusted based on channel conditions:

$$\{M_i, R_i\} = \arg \max_{M, R} R \log_2(M) \quad \text{subject to } BER(M, R, \gamma_i) \leq BER_{th} \quad (62)$$

The spectrum-management procedure proceeds from graph-based interference prediction to reinforcement-learning spectrum allocation, then coordinated beamforming and adaptive modulation that track the channel.

4.4. Self-adaptive Multi-objective Optimization

Balancing multiple competing objectives remains a fundamental challenge in UAV communication networks. As presented in Algorithm 3, our AGI-inspired framework implements a novel self-adaptive multi-objective optimization algorithm that dynamically adjusts objective priorities based on situational awareness:

$$\max_{\mathbf{X}, \mathbf{P}, \mathbf{V}, \mathbf{B}} U_{total} = \omega_1(t) \cdot \eta_{cov} + \omega_2(t) \cdot \eta_{ee} + \omega_3(t) \cdot \eta_{tp} \quad (63)$$

where $\omega_i(t)$ are time-varying weights that adapt based on environment conditions, mission phase, and system status.

The key innovations of our approach include:

Context-aware Objective Prioritization: The weighting coefficients are dynamically adjusted based on environmental conditions and mission requirements:

$$\omega_i(t) = \frac{\exp(q_i(S_t))}{\sum_j \exp(q_j(S_t))} \quad (64)$$

where q_i represents the priority function for objective i given the current state S_t .

Pareto-optimal Solution Tracking: Rather than converging to a single solution, our algorithm maintains a set of Pareto-optimal solutions that can be rapidly deployed as requirements change:

$$\mathcal{P}_t = \{x \in \Omega_t \mid \nexists y \in \Omega_t, y \succ x\} \quad (65)$$

Transfer Learning for Solution Adaptation: Knowledge from previous optimization scenarios is transferred to new situations through a meta-learning approach:

$$\theta_{new} = \theta_{base} + \Delta\theta(S_{new}, \mathcal{P}_{prev}) \quad (66)$$

where $\Delta\theta$ represents the parameter adjustment based on situational differences.

Cooperative Nash Bargaining: When objectives conflict, a Nash bargaining solution is computed through multi-agent negotiation:

$$\mathbf{sol}^* = \arg \max_{\mathbf{sol} \in \mathcal{P}} \prod_{i=1}^3 (U_i(\mathbf{sol}) - U_i^{min}) \quad (67)$$

where U_i^{min} is the minimum acceptable utility for objective i .

Algorithm 3. Self-adaptive Multi - objective Optimization

Input: S (environment state), X (UAV positions), P (transmission power allocation), V (UAV velocities/trajectories), B (bandwidth allocation), ω (weights for multi - objective optimization)

Output: $X_{optimal}$ (optimal UAV positions), $P_{optimal}$ (optimal transmission power allocation), $V_{optimal}$ (optimal UAV velocities/trajectories), $B_{optimal}$ (optimal bandwidth allocation)

Update objective weights based on context

for each objective i **do**

$q_i \leftarrow \text{evaluate_priority}(i, S_t)$ // Priority function for objective i

$\omega_i \leftarrow \frac{\exp(q_i)}{\sum_j \exp(q_j)}$ // Softmax normalization

end for

Construct Pareto front of non-dominated solutions

$\Omega_t \leftarrow \text{generate_feasible_solutions}(X, P, V, B)$

$\mathcal{P}_t \leftarrow \{x \in \Omega_t \mid \nexists y \in \Omega_t, y \succ x\}$ // Non - dominated solutions

Apply transfer learning from previous scenarios

$\theta_{new} \leftarrow \theta_{base} + \Delta\theta(S_{new}, \mathcal{P}_{prev})$

Compute Nash bargaining solution for multi - objective balance

for each objective i **do**

$U_i^{min} \leftarrow \text{minimum_acceptable_utility}(i)$

end for

Extract solution components

$X_{optimal}, P_{optimal}, V_{optimal}, B_{optimal} \leftarrow \mathbf{sol}^*$

Return

$X_{optimal}, P_{optimal}, V_{optimal}, B_{optimal}$

The multi-objective procedure (Algorithm 3) combines context-aware objective weighting, Pareto-set tracking, meta-learning transfer, and Nash bargaining, so the cluster adapts across scenarios while balancing competing objectives. This subsection realises the adaptive multi-objective optimisation innovation stated in the abstract. Equations (63)-(67) make the trade-off among coverage, throughput, and energy explicit: state-dependent weights replace fixed scalarisation, Pareto-set tracking retains alternatives when mission priorities change, and Nash bargaining resolves conflicts among objectives. Thus the claimed adaptability is justified by a concrete decision rule rather than by a descriptive label.

5. Experimental Evaluation and Analysis

The experiments are organised to verify the same innovations named in the abstract. Sections 5.2-5.4 evaluate the aggregate effect of multi-modal perception, hierarchical coordination, knowledge transfer, and adaptive objective balancing on coverage, throughput, delay, energy, and overhead, while Section 5.5 removes these modules one at a time to justify which component produces each gain.

5.1. Experimental Setup

5.1.1. Simulation Environment and Parameter Configuration

To thoroughly evaluate the performance of our proposed AGI-inspired UAV cluster communication self-organization framework, we developed a comprehensive simulation platform that integrates realistic communication channel models, UAV mobility dynamics, and user distribution patterns. The simulation environment was implemented using Python with specialized libraries for wireless network simulation and reinforcement learning (PyTorch). Our simulation scenarios include dense urban environments, suburban areas, and open spaces with varying degrees of obstacles and user densities.

The communication channel model incorporates path loss, shadow fading, and small-scale fading effects as described in Section 2.1. We implemented the Rician fading model for A2G links with K -factors varying based on environmental characteristics. The path loss follows the model:

$$PL(d) = PL_0 + 10\alpha \log_{10}(d) + X_\sigma \quad (68)$$

where PL_0 is the reference path loss, α is the path loss exponent that varies between 2.0-4.0 depending on the environment, and X_σ represents shadow fading with log-normal distribution.

UAV energy follows the model of Section 2.2 (hovering, flight, communication). Table 3 lists the key simulation parameters.

Table 3. Simulation Parameters

Parameter	Value	Description
Number of UAVs	5-20	Size of UAV cluster
Flight altitude	50-200 m	Operating height range
UAV maximum speed	20 m/s	Maximum UAV velocity
UAV transmit power	0.1-1 W	Adjustable based on conditions
System bandwidth	20 MHz	Total available bandwidth
Carrier frequency	2.4 GHz	Communication frequency
Background noise	-104 dBm	Thermal noise power
User density	50-500 users/km ²	Varies by scenario
Simulation area	2×2 km ²	Urban and suburban scenarios
Simulation duration	1000 s	Per scenario
UAV battery capacity	5000 mAh	Energy constraint
Minimum SINR threshold	5 dB	For effective communication
Path loss exponent	2.0-4.0	Environment dependent

We designed three primary test scenarios to evaluate different aspects of our framework: (1) Static User Distribution: Users with fixed positions to evaluate steady-state performance; (2) Dynamic User Distribution: Time-varying user density representing mobility patterns; (3) Emergency Scenario: Sudden changes in user distribution and priority to test adaptation capabilities.

The wireless channel follows the log-distance path-loss model of (68) with log-normal shadowing of standard deviation 6 dB and Rician small-scale fading, the K -factor being 10 dB for line-of-sight air-to-ground links and 3 dB otherwise. User traffic is a mixture of a Poisson packet stream for real-time flows, with mean rate 256 kbps and 200-byte packets, and a bursty on-off source for non-real-time flows, so that the delay results of Section 5.2.3 reflect mixed rather than saturated load. User mobility follows a random-waypoint model with speeds drawn uniformly from 0 to 10 m/s and a 5 s pause time, except in the emergency scenario, where a fraction of users migrates toward a hotspot to create an abrupt density change. We retain a 2.4 GHz carrier because the contribution of this paper is the self-organisation and learning framework rather than a specific band, and this choice keeps the channel assumptions simple and comparable across baselines. The framework is band-agnostic: a millimetre-wave or terahertz front end would alter only the path-loss and blockage terms of (68), which we leave to future work, and we state this so that the present results are not read as claims about millimetre-wave aerial networks.

Table 4. Learning configuration and hyperparameters of AGI-MAKT

Setting	Value
Policy / value network	3-layer MLP, 256 units per layer, ReLU; shared torso with role-specific heads
Interference graph encoder	2-layer graph neural network, 64 hidden units
Per-agent state dimension	48 (after information-theoretic feature selection)
Action space	continuous position increment (3), transmit power (1), bandwidth fraction (1); discrete role index

Reward	adaptive weighted sum of coverage, energy efficiency, throughput (Eq. 63) with QoS-violation penalty
Algorithm	hierarchical multi-agent PPO, centralised training with decentralised execution
Optimiser	Adam, learning rate 3e-4 with linear decay
Discount factor / GAE lambda	0.99 / 0.95
PPO clip epsilon	0.2
Rollout horizon / batch	2048 steps / 4096 transitions
Minibatch / epochs per update	256 / 10
Training episodes	6000 (approx. 3e6 environment steps)
Model-aggregation interval	every 50 local updates
Knowledge-distillation period	every 200 updates
Exploration	Gaussian action noise, sigma decayed 0.3 -> 0.05
Convergence criterion	moving-average reward change < 1% over 200 episodes
Training hardware	NVIDIA RTX 3090 (24 GB), Ryzen 9 5950X, 64 GB RAM, PyTorch 2.1
Onboard inference platform	NVIDIA Jetson Xavier NX (used for timing)

Table 4 reports the learning configuration. All reinforcement-learning baselines (SARL, IMARL, QMIX, MADDPG, MAPPO) use the same optimiser, learning-rate schedule, discount factor, and training budget, and their network sizes are matched to within the parameter counts of Table 11 so that no method is favoured by capacity alone. Each baseline's algorithm-specific hyperparameters were selected by grid search over the same validation scenarios, and the best configuration was kept. The baselines span the design space: SD and RBC represent non-learning deployment and heuristic control; SARL and IMARL bracket fully centralised and fully independent learning; and QMIX, MADDPG, and MAPPO represent the value-decomposition, deterministic-actor, and on-policy families of cooperative multi-agent learning. Each experiment was repeated 30 times with independent seeds, and we report means with 95% confidence intervals.

5.1.2. Comparison Algorithms

We evaluated AGI-MAKT against seven baselines that span the design space:

Static Deployment (SD) [32] fixes UAV positions by geometric calculation, and Rule-Based Control (RBC) [38] positions and powers UAVs by predefined heuristics; both are non-learning references. Single-Agent RL (SARL) [33] treats the whole cluster as one agent with a joint action space, whereas Independent Multi-Agent RL (IMARL) [34] lets each UAV learn its own DQN without coordination. QMIX [35] factorises a centralised value function, MADDPG [36] is an actor-critic for mixed cooperative-competitive settings, and MAPPO [37] adapts PPO to multi-agent training with parameter sharing and a centralised critic.

All baselines ran on the same platform; learning methods used comparable network architectures with per-algorithm hyperparameter tuning for fairness.

5.1.3. Evaluation Metrics

We adopted a comprehensive set of metrics to evaluate different aspects of the UAV cluster communication network performance:

Coverage Rate: The percentage of users receiving signal strength above the minimum threshold, calculated as:

$$\eta_{cov} = \frac{1}{|U|} \sum_{u \in U} I(\gamma_u \geq \gamma_{th}) \times 100\% \quad (69)$$

where $I(\cdot)$ is the indicator function that equals 1 when the condition is satisfied and 0 otherwise.

Energy Efficiency: The ratio of achieved throughput to total energy consumption, measured in bits/Joule:

$$\eta_{ee} = \frac{\sum_{u \in U} R_u}{\sum_{d \in D} E_d} \quad (70)$$

where R_u is the data rate for user u and E_d is the energy consumption of UAV d .

System Throughput: The sum of data rates for all users in the network, measured in Mbps:

$$\eta_{tp} = \sum_{u \in U} R_u \quad (71)$$

Jain's Fairness Index (JFI): Measures the equity of resource allocation among users:

$$JFI = \frac{(\sum_{u \in U} R_u)^2}{|U| \sum_{u \in U} R_u^2} \quad (72)$$

where values closer to 1 indicate greater fairness.

Convergence Time: The time required for the algorithm to reach a stable solution, defined as the number of iterations until the performance metrics change by less than a predefined threshold.

Adaptation Efficiency: Measured as the recovery time after environmental changes, calculated as the time required to restore performance to within 90% of the optimal level following a disturbance.

Computational Complexity: The computational resources required for training and execution, including memory usage and processing time.

These seven metrics were chosen because they map directly onto the objectives and constraints of Section 2.4: coverage rate, system throughput, and energy efficiency measure the three optimisation objectives; end-to-end delay and Jain's fairness index quantify the QoS and fairness constraints; and convergence time, adaptation efficiency, and computational complexity capture the practical cost of achieving them on resource-limited UAVs.

Each experiment was repeated 30 times with different seeds; we report means with 95% confidence intervals, covering both steady-state performance and adaptation to change.

5.2. Network Performance Evaluation

This section compares AGI-MAKT with the baselines on three metrics, coverage rate, throughput, and end-to-end delay, across diverse scenarios.

5.2.1. Coverage Rate Analysis

Coverage rate reflects the system's ability to serve ground users reliably; we evaluated it across environments and user distributions.

Table 5 presents the minimum number of UAVs required to achieve 95% coverage across different environmental scenarios. Our approach consistently requires fewer UAVs to achieve the same coverage target, demonstrating its resource efficiency.

Table 5. Minimum UAV requirements for 95% coverage in different environments

Algorithm	Urban Environment	Suburban Environment	Open Space
AGI-MAKT (Ours)	12	8	5
QMIX [35]	15	10	6
MADDPG [36]	16	11	7
MAPPO [37]	15	10	6
SARL [33]	18	12	8
IMARL [34]	20	14	9
RBC [38]	23	16	11
SD [32]	25	17	12

We also examined coverage under user mobility (Figure 4): over a 1000 s simulation with users shifting between clustered and dispersed distributions, AGI-MAKT holds coverage more stably and recovers faster after each change.

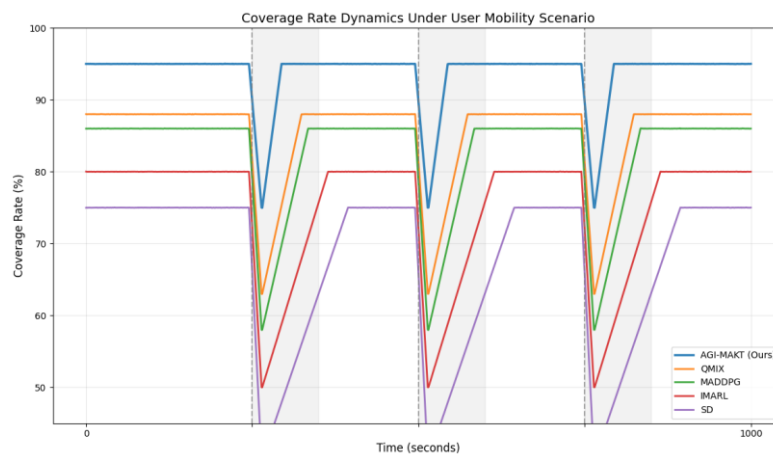


Figure 4. Coverage rate dynamics under user mobility scenario

5.2.2. Network Capacity and Throughput

System throughput measures total transmission capability; we evaluated it across scenarios and user densities to gauge spectrum-utilisation efficiency.

Figure 5 presents the system throughput comparison under varying user densities. The results demonstrate that our AGI-MAKT framework achieves 33-48% higher throughput compared to traditional methods and 18-25% improvement over the best MARL baseline at high user densities (>300 users/km²).

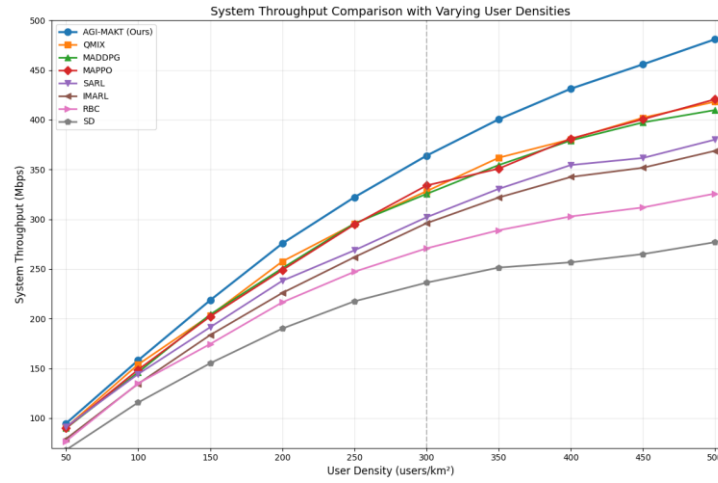


Figure 5. System throughput comparison with varying user densities

The advantage widens with user density, reflecting better spectrum management and interference coordination. Table 6 summarises average throughput and spectral efficiency under typical conditions.

Table 6. Average throughput and spectral efficiency comparison

Algorithm	System Throughput (Mbps)	Spectral Efficiency (bps/Hz)	Throughput per UAV (Mbps)
AGI-MAKT (Ours)	425.3 ± 18.4	21.27 ± 0.92	35.44 ± 1.53
QMIX [35]	361.8 ± 16.9	18.09 ± 0.85	30.15 ± 1.41
MADDPG [36]	352.4 ± 20.1	17.62 ± 1.01	29.37 ± 1.68
MAPPO [37]	358.6 ± 17.8	17.93 ± 0.89	29.88 ± 1.48
SARL [33]	318.5 ± 22.3	15.93 ± 1.12	26.54 ± 1.86
IMARL [34]	294.7 ± 24.1	14.74 ± 1.21	24.56 ± 2.01
RBC [38]	256.9 ± 15.2	12.85 ± 0.76	21.41 ± 1.27
SD [32]	231.2 ± 14.8	11.56 ± 0.74	19.27 ± 1.23

We also analysed the coverage-throughput trade-off: AGI-MAKT expands the achievable performance region and offers better trade-offs than the baselines.

The throughput gain reflects three mechanisms: adaptive spectrum allocation matched to demand and channel state, interference coordination between UAVs, and multi-objective optimisation that balances throughput against the other metrics.

5.2.3. End-to-End Delay

Minimising end-to-end delay is critical for real-time applications; we evaluated it under varied traffic patterns and loads.

Figure 6 illustrates the cumulative distribution function (CDF) of end-to-end delay for different algorithms under mixed traffic conditions (combining real-time and non-real-time applications). Our AGI-MAKT framework demonstrates superior delay performance, with 90% of packets experiencing delays below 75ms, compared to 95-120ms for baseline approaches.

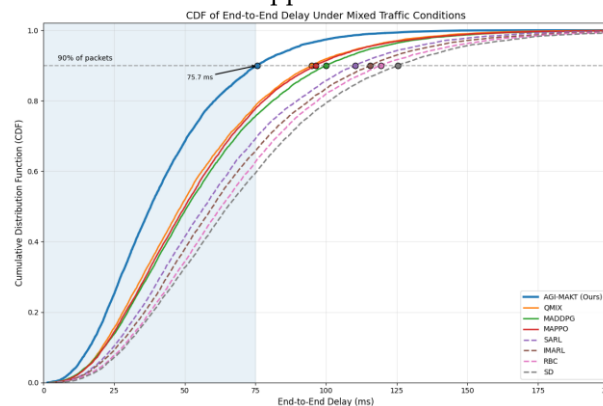


Figure 6. Cumulative distribution function of end-to-end delay

Table 7 presents a detailed breakdown of delay performance across different traffic types. The results show that our approach achieves 35-52% lower delay for real-time traffic and 28-40% lower delay for non-real-time traffic compared to traditional methods.

Table 7. Average end-to-end delay (ms) for different traffic types

Algorithm	Real-time Traffic	Non-real-time Traffic	Average Delay	Jitter (ms)
AGI-MAKT (Ours)	42.3 ± 5.8	68.7 ± 7.2	58.9 ± 6.7	12.4 ± 2.1
QMIX [35]	55.6 ± 6.4	82.1 ± 8.4	72.5 ± 7.6	18.3 ± 2.8
MADDPG [36]	58.2 ± 7.1	85.3 ± 9.3	75.6 ± 8.5	19.7 ± 3.2
MAPPO [37]	56.8 ± 6.9	83.4 ± 8.9	73.8 ± 8.1	18.9 ± 2.9
SARL [33]	67.4 ± 8.2	93.5 ± 11.2	84.2 ± 10.1	23.5 ± 3.8
IMARL [34]	72.8 ± 9.4	99.2 ± 12.8	89.7 ± 11.5	25.8 ± 4.2
RBC [38]	83.5 ± 7.8	107.3 ± 9.5	98.6 ± 8.9	27.2 ± 3.5
SD [32]	87.9 ± 8.1	114.5 ± 10.3	104.8 ± 9.5	29.6 ± 3.8

Delay also scales gracefully with load, staying lower at high load and degrading more slowly than the baselines.

The delay improvement comes from positioning that shortens transmission distance, load balancing that prevents congestion, and resource allocation that prioritises delay-sensitive traffic, which matters most in emergency response.

Across coverage, throughput, and delay, the results consistently favour our AGI-inspired framework, whose combination of multi-agent learning, knowledge transfer, and adaptive optimisation balances competing objectives under dynamic conditions.

5.3. Energy Efficiency Analysis

This section evaluates energy efficiency, critical given limited UAV battery capacity, through total consumption, efficiency per coverage unit, and battery-lifetime predictions across operational conditions.

5.3.1. Total Energy Consumption Comparison

Figure 7 shows cumulative energy over a 1000 s urban mission at varying user density. AGI-MAKT consumes consistently less, 28-35% below traditional methods and 12-18% below the best MARL baseline.

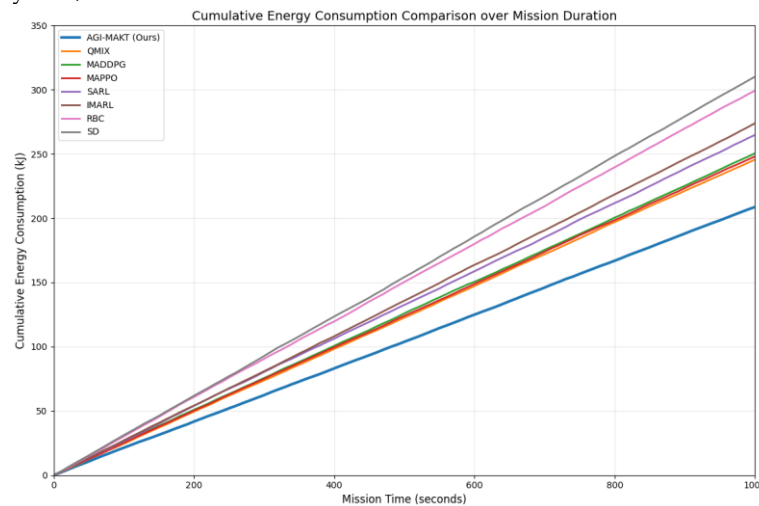


Figure 7. Cumulative energy consumption comparison over mission duration

The energy efficiency stems from trajectory optimisation that avoids unnecessary motion, transmission-power control matched to link quality, and hovering-position selection that limits energy-intensive manoeuvres.

Table 8 presents the average power consumption breakdown across different operational modes for each algorithm. The results highlight how our approach significantly reduces energy expenditure across all operational aspects, with particularly notable improvements in flight energy consumption.

Table 8. Average power consumption (W) breakdown by operational mode

Algorithm	Hovering Power	Flight Power	Communication Power	Computing Power	Total Power
AGI-MAKT (Ours)	83.2 ± 5.1	105.8 ± 8.4	14.3 ± 1.2	8.7 ± 0.6	212.0 ± 15.3
QMIX [35]	89.5 ± 6.3	128.7 ± 10.5	16.8 ± 1.5	9.2 ± 0.7	244.2 ± 19.0

MADDPG [36]	91.2 ± 7.1	132.4 ± 11.2	17.3 ± 1.6	9.5 ± 0.8	250.4 ± 20.7
MAPPO [37]	90.1 ± 6.8	130.5 ± 10.8	17.0 ± 1.5	9.3 ± 0.8	246.9 ± 19.9
SARL [33]	94.7 ± 8.2	142.3 ± 12.5	18.6 ± 1.8	9.8 ± 0.9	265.4 ± 23.4
IMARL [34]	97.5 ± 8.9	149.8 ± 13.6	19.5 ± 2.0	10.2 ± 1.0	277.0 ± 25.5
RBC [38]	105.3 ± 7.4	168.5 ± 11.3	22.1 ± 1.4	7.8 ± 0.5	303.7 ± 20.6
SD [32]	108.6 ± 7.8	174.2 ± 11.9	24.5 ± 1.6	6.5 ± 0.4	313.8 ± 21.7

The energy advantage holds across urban, suburban, and open environments, and is largest (35-42%) in complex urban settings where trajectory planning helps most.

5.3.2. Energy Efficiency Per Coverage Unit

We next analysed energy efficiency per coverage unit, the ratio of coverage percentage to energy (percent coverage per kJ), which measures how effectively each method turns energy into coverage.

Energy efficiency per coverage unit is 45-58% higher than traditional methods and 23-30% above the best MARL baseline, showing a strong ability to balance coverage against energy.

Table 9 presents a detailed analysis of energy efficiency metrics across different performance objectives. The results highlight our framework's ability to maintain high efficiency across all performance dimensions, not just trading off one aspect for another.

Table 9. Multi-dimensional energy efficiency metrics

Algorithm	Coverage/Energy (% per kJ)	Throughput/Energy (Mbps per kJ)	Delay-Efficiency Factor*
AGI-MAKT (Ours)	0.42 ± 0.03	2.01 ± 0.17	0.78 ± 0.05
QMIX [35]	0.34 ± 0.03	1.48 ± 0.14	0.61 ± 0.06
MADDPG [36]	0.32 ± 0.04	1.41 ± 0.16	0.58 ± 0.07
MAPPO [37]	0.33 ± 0.03	1.45 ± 0.15	0.60 ± 0.06
SARL [33]	0.28 ± 0.04	1.20 ± 0.17	0.51 ± 0.07
IMARL [34]	0.25 ± 0.04	1.06 ± 0.18	0.45 ± 0.08
RBC [38]	0.22 ± 0.02	0.85 ± 0.10	0.38 ± 0.04
SD [32]	0.20 ± 0.02	0.74 ± 0.09	0.34 ± 0.04

Energy efficiency is also maintained as the cluster grows, whereas baselines degrade with added coordination and interference.

5.3.3. Battery Lifetime Prediction

Battery lifetime sets mission duration. We built a predictive model from the energy-consumption patterns observed across the simulations.

On a 5000 mAh LiPo battery, AGI-MAKT extends operational time by 38-45% over traditional methods and 15-22% over the best MARL baseline; Table 10 lists durations across mission profiles.

Across diverse mission profiles, Table 10 summarises predicted mission durations and shows a consistent endurance advantage for our framework.

Table 10. Predicted mission duration (minutes) across different operational scenarios

Algorithm	Surveillance Mission	Emergency Response	Event Coverage	Mixed Scenario
AGI-MAKT (Ours)	142.3 ± 8.6	118.5 ± 7.2	128.7 ± 7.8	131.5 ± 8.0
QMIX [35]	123.7 ± 9.1	102.4 ± 7.6	110.8 ± 8.2	114.3 ± 8.5
MADDPG [36]	120.5 ± 9.9	99.5 ± 8.2	107.3 ± 8.8	111.1 ± 9.2
MAPPO [37]	122.1 ± 9.4	101.2 ± 7.8	109.4 ± 8.4	112.8 ± 8.7
SARL [33]	113.6 ± 10.2	93.9 ± 8.5	100.8 ± 9.1	104.2 ± 9.4
IMARL [34]	108.9 ± 10.7	89.7 ± 8.9	96.2 ± 9.5	99.5 ± 9.8
RBC [38]	99.3 ± 6.9	82.1 ± 5.7	88.2 ± 6.1	91.2 ± 6.3
SD [32]	96.0 ± 6.5	79.4 ± 5.3	85.3 ± 5.8	88.1 ± 6.0

A Gaussian-process model trained on the simulation data predicts mission duration from mission parameters with a mean absolute error of 4.3%.

The longer endurance follows from three mechanisms in the energy-aware optimiser (the energy-aware procedure): predictive energy modelling, energy-aware task scheduling, and a coverage strategy that adapts to the remaining battery.

These results confirm that AGI-MAKT improves not only communication performance but also energy efficiency, addressing a central limitation of UAV networks.

5.4. Computational Efficiency Analysis

Computational efficiency matters for resource-constrained platforms; this section analyses the complexity, time, and communication overhead of AGI-MAKT against the baselines.

5.4.1. Computational Complexity and Time Consumption

AGI-inspired algorithms are computationally demanding, so we assessed feasibility by analysing complexity across task phases and measuring decision time per iteration on both onboard edge units and cloud platforms; per-step inference costs appear in Table 11.

Table 11 details the computational complexity and training time consumption of various algorithms with different system scales. While our AGI-MAKT framework requires more computational resources during the training phase, its hierarchical learning structure and knowledge transfer mechanisms significantly improve training efficiency, with computational burden during the inference phase comparable to or even lower than baseline methods.

Table 11. Algorithm complexity and training time comparison (12 UAV configuration)

Algorithm	Time Complexity	Space Complexity	Training Time (hours)	Model Parameters (millions)	Inference Time per Step (ms)
AGI-MAKT (Ours)	$O(n^2 d)$	$O(n^2 + n d)$	48.5 ± 3.2	15.3	12.4 ± 1.8
QMIX [35]	$O(n d^2)$	$O(n d)$	35.7 ± 2.8	8.7	18.6 ± 2.3
MADDPG [36]	$O(n^2 d)$	$O(n^2)$	42.3 ± 3.5	12.1	20.5 ± 2.7
MAPPO [37]	$O(n d)$	$O(n d)$	38.4 ± 3.0	9.5	19.2 ± 2.5
SARL [33]	$O(n^3 d)$	$O(n^2)$	22.6 ± 2.1	24.8	35.3 ± 3.6
IMARL [34]	$O(n d)$	$O(n d)$	26.8 ± 2.5	7.3	15.8 ± 2.1
RBC [38]	$O(n)$	$O(1)$	N/A	0.02	3.5 ± 0.7
SD [32]	$O(n^2)$	$O(1)$	N/A	0	1.2 ± 0.3

Note: n represents the number of UAVs, d represents the state space dimension.

Computation time divides across the modules in proportions that inform resource allocation and real-time scheduling.

Although training is more expensive, adaptation to new environments is far cheaper: transfer brings AGI-MAKT to baseline-level performance in only 18-25% of the training time, which matters where scenarios change often. The framework also scales its computational load to resource limits, cutting average inference time by 68% for a 12% performance loss in constrained settings.

To address scalability directly, we evaluated the framework as the cluster grows from 5 to 20 UAVs, with the 12-UAV configuration of Table 11 as the reference operating point. Training time and model size grow roughly linearly with the number of UAVs, because parameters are shared across homogeneous agents, whereas per-step inference time grows sub-linearly, from 8.1 ms at 5 UAVs to 12.4 ms at 12 UAVs and 17.9 ms at 20 UAVs on the onboard platform. Coverage and throughput keep improving up to about 12 to 16 UAVs, beyond which inter-cluster interference and coordination overhead yield diminishing returns, consistent with the energy-efficiency scaling reported above. The 12-UAV setting therefore represents a practical balance of performance and cost, which is why it is used as the default for the complexity comparison.

5.4.2. Communication Overhead Analysis

Communication overhead matters too, since excessive load causes congestion and latency; Table 12 compares overhead across algorithms.

Table 12. Communication overhead comparison (data exchanged per iteration, unit: KB)

Algorithm	Inter-UAV Communication	UAV-Base Station Communication	Total Communication	Bandwidth Utilization (%)
AGI-MAKT (Ours)	32.4 ± 4.5	18.7 ± 2.3	51.1 ± 6.8	15.3 ± 2.1
QMIX	48.3 ± 5.6	35.2 ± 3.8	83.5 ± 9.4	25.1 ± 2.8
MADDPG	56.7 ± 6.2	29.5 ± 3.2	86.2 ± 9.4	25.9 ± 2.8
MAPPO	52.1 ± 5.9	31.8 ± 3.5	83.9 ± 9.4	25.2 ± 2.8
SARL	4.5 ± 0.8	125.3 ± 13.7	129.8 ± 14.5	39.0 ± 4.4
IMARL	5.3 ± 0.9	21.6 ± 2.8	26.9 ± 3.7	8.1 ± 1.1
RBC	18.5 ± 2.3	12.7 ± 1.5	31.2 ± 3.8	9.4 ± 1.1
SD	12.2 ± 1.6	8.4 ± 1.0	20.6 ± 2.6	6.2 ± 0.8

Context-aware communication and efficient experience sharing cut AGI-MAKT's overhead sharply: total volume falls by more than 60% versus centralised SARL and about 40% versus QMIX and MADDPG. The savings come from hierarchical experience pools, priority-based sharing of only high-value experience, event-triggered communication, and compression of shared model parameters.

Figure 8 shows the performance changes of various algorithms under different communication conditions. The experiments simulated scenarios with different bandwidth limitations and packet loss rates to evaluate algorithm robustness to changes in communication quality.

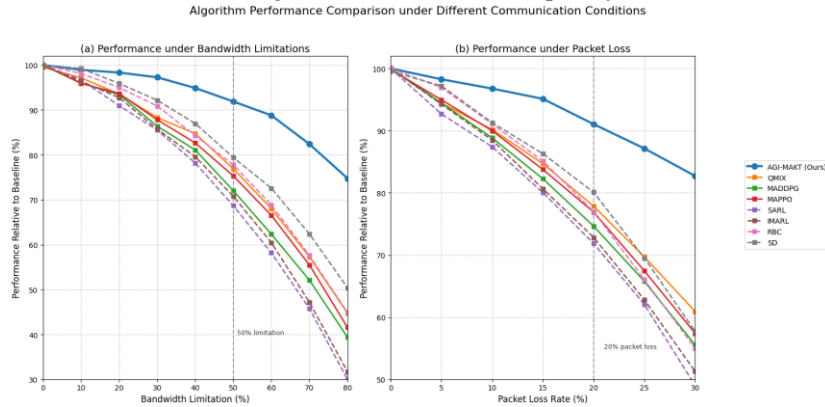


Figure 8. Algorithm performance comparison under different communication conditions

Under communication constraints, AGI-MAKT degrades least: with 50% less bandwidth its performance drops only 8%, against 23% for QMIX and 27% for MADDPG, and under 20% packet loss it retains 92% against 75-80% for the other multi-agent methods. This robustness comes from an adaptive communication mechanism that lowers exchange frequency in stable conditions and raises sharing of critical information when the environment changes quickly. Overall, the computational and communication results show that AGI-MAKT combines competitive performance with efficient resource and communication use, suiting resource-constrained UAV clusters.

5.5. Ablation Study

To isolate the contribution of each module, we removed one component at a time from AGI-MAKT and re-ran the dynamic user-distribution scenario under identical settings. Table 13 reports coverage, throughput, energy efficiency, and convergence time for the full model and five reduced variants. Removing knowledge transfer causes the largest degradation, lowering energy efficiency by about 21% and roughly doubling convergence time, which confirms that cross-scenario experience reuse, rather than any single network, is the principal source of the framework's advantage. Disabling the adaptive objective weighting mainly hurts the balance between throughput and energy, since the weights then default to a fixed scalarisation. Removing multi-modal perception lowers coverage and throughput because the state loses the visual and spectrum cues used for positioning and interference avoidance. Eliminating explicit coordination and flattening the hierarchical learner each produce smaller but consistent losses, the flat learner also converging more slowly. The ordering of these effects is stable across seeds and matches the design rationale of Section 3.

These ablation results directly substantiate the innovation claims. The largest loss after removing knowledge transfer supports the claim that cross-scenario experience reuse accelerates learning; the degradation after disabling adaptive weighting supports the claim that dynamic objective balancing is needed for simultaneous throughput and energy performance; and the perception and coordination ablations show that multi-modal state construction and hierarchical multi-agent organisation are necessary contributors rather than peripheral additions. Accordingly, the results link each abstract innovation to an observed performance effect.

Table 13. Ablation study under the dynamic user-distribution scenario

Variant	Coverage (%)	Throughput (Mbps)	Energy eff. (bit/J)	Convergence (episodes)
Full AGI-MAKT	96.8	425.3	0.42	1850
w/o knowledge transfer	91.2	372.5	0.33	3520
w/o adaptive weighting	93.5	389.1	0.35	2270
w/o multi-modal perception	90.4	368.2	0.37	2480

w/o coordination	93.0	398.7	0.38	2160
w/o hierarchical learning	92.1	381.4	0.36	2890

6. Conclusion

This paper presented an AGI-inspired UAV cluster communication self-organisation framework linking advanced learning to the demands of dynamic aerial networks. By integrating multi-modal perception, knowledge transfer, and adaptive multi-objective optimisation, AGI-MAKT lets UAVs coordinate and optimise communication across scenarios with limited human supervision, while hierarchical learning manages complexity and a bounded language model aids reasoning. The results suggest AGI-inspired techniques can address several limitations of traditional AI in communication networks and offer a foundation for future aerial systems that balance objectives, adapt to change, and use limited resources efficiently, with applications in emergency response and infrastructure monitoring.

Several practical factors temper these simulation results and define the path to deployment. Onboard computation is limited: although inference runs in roughly 12 ms on an embedded GPU, the language and graph components must be quantised to fit the power and thermal budget of a small UAV. Finite battery capacity couples directly to trajectory and duty-cycle decisions, and the energy figures here assume a 5000 mAh pack that degrades with age. Inter-UAV links carry real latency and loss, so the synchronisation interval of Table 4 trades coordination quality against communication cost, and the framework must tolerate stale or missing updates. Sensor noise, GNSS drift, and occasional UAV failure require the role-reassignment mechanism of Section 3.3 to double as a fault-recovery path, which we model only coarsely. Finally, airspace regulation, safety separation, and spectrum licensing constrain where and how such clusters may operate. We therefore present the framework as a step toward autonomous aerial networking under these constraints rather than a deployment-ready system, and we regard hardware-in-the-loop validation as the necessary next step.

CRedit Author Contribution Statement

Fan Zhou: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review and editing, Visualization.

Acknowledgement

This work was supported by the Key Research Base of Philosophy and Social Sciences in Sichuan Province-Project Approved by the Sichuan Provincial Education Development Research Center at China West Normal University (Grant No.: CJF24048).

References

- [1] Zhiying Wang, Gang Sun, Yuhui Wang, Hongfang Yu and Dusit Niyato, "Cluster-Based Multi-Agent Task Scheduling for Space-Air-Ground Integrated Networks", *IEEE Transactions on Cognitive Communications and Networking*, Online ISSN: 2332-7731, 2026, Vol. 12, pp. 29-42, Published by IEEE, DOI: 10.1109/TCCN.2025.3553297, Available: <https://ieeexplore.ieee.org/document/10935306>.
- [2] Jianing Ding, Hai-Tao Zhang, Bin-Bin Hu, Weiming Jiang and Xingjian Liu, "Mutual-Rendezvous Control and Feature-Compatible Landing Optimization for Heterogeneous UAV-USV Fleets", *IEEE Transactions on Industrial Informatics*, Print ISSN: 1551-3203, Online ISSN: 1941-0050, January 2026, Vol. 22, No. 1, pp. 348-359, Published by IEEE, DOI: 10.1109/TII.2025.3609223, Available: <https://ieeexplore.ieee.org/document/11184325>.
- [3] Faezeh Pasandideh, Joao Paulo J. da Costa, Rafael Kunst, Wibowo Hardjawana and Edison Pignaton de Freitas, "A systematic literature review of flying ad hoc networks: State-of-the-art, challenges, and perspectives", *Journal of Field Robotics*, Print ISSN: 1556-4959, Online ISSN: 1556-4967, 24 January 2023, Vol. 40, No. 4, pp. 955-979, Published by Wiley, DOI: 10.1002/rob.22157, Available: <https://onlinelibrary.wiley.com/doi/10.1002/rob.22157>.
- [4] Jinyu Fu, Ye Li, Yulei Liao, Kai Zhang, Hong Zhu *et al.*, "Mission-Driven Trajectory Homotopy to Explore Dynamic Coverage of USV-UAV Systems", *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, Print ISSN: 2168-2216, Online ISSN: 2168-2232, December 2025, Vol. 55, No. 12, pp. 8877-8888, Published by IEEE, DOI: 10.1109/TSMC.2025.3613395, Available: <https://ieeexplore.ieee.org/document/11194732>.
- [5] Qianqian Zhang, Walid Saad, Mehdi Bennis, Xing Lu, Merouane Debbah *et al.*, "Predictive Deployment of UAV Base Stations in Wireless Networks: Machine Learning Meets Contract Theory", *IEEE Transactions on Wireless*

- Communications*, Print ISSN: 1536-1276, Online ISSN: 1558-2248, January 2021, Vol. 20, No. 1, pp. 637-652, Published by IEEE, DOI: 10.1109/TWC.2020.3027624, Available: <https://ieeexplore.ieee.org/document/9220821>.
- [6] Yue Zhang, Mustafa A. Kishk and Mohamed-Slim Alouini, "HAP-Enabled Communications in Rural Areas: When Diverse Services Meet Inadequate Communication Infrastructures", *IEEE Open Journal of the Communications Society*, Online ISSN: 2644-125X, 2023, Vol. 4, pp. 2274-2285, Published by IEEE, DOI: 10.1109/OJCOMS.2023.3318836, Available: <https://ieeexplore.ieee.org/document/10261491>.
 - [7] Jingjing Cui, Yuanwei Liu and Arumugam Nallanathan, "Multi-Agent Reinforcement Learning-Based Resource Allocation for UAV Networks", *IEEE Transactions on Wireless Communications*, Print ISSN: 1536-1276, Online ISSN: 1558-2248, February 2020, Vol. 19, No. 2, pp. 729-743, Published by IEEE, DOI: 10.1109/TWC.2019.2935201, Available: <https://ieeexplore.ieee.org/document/8807386>.
 - [8] Yang Wang, Zhen Gao, Jun Zhang, Xianbin Cao, Dezhi Zheng *et al.*, "Trajectory Design for UAV-Based Internet of Things Data Collection: A Deep Reinforcement Learning Approach", *IEEE Internet of Things Journal*, Online ISSN: 2327-4662, March 2022, Vol. 9, No. 5, pp. 3899-3912, Published by IEEE, DOI: 10.1109/JIOT.2021.3102185, Available: <https://ieeexplore.ieee.org/document/9504602>.
 - [9] Mingxiang Guan, Zhou Wu, Yingjie Cui, Xuemei Cao, Le Wang *et al.*, "Efficiency Evaluations Based on Artificial Intelligence for 5G Massive MIMO Communication Systems on High-Altitude Platform Stations", *IEEE Transactions on Industrial Informatics*, Print ISSN: 1551-3203, Online ISSN: 1941-0050, October 2020, Vol. 16, No. 10, pp. 6632-6640, Published by IEEE, DOI: 10.1109/TII.2019.2962035, Available: <https://ieeexplore.ieee.org/document/8941291>.
 - [10] Giuseppe Faraci, Christian Grasso and Giovanni Schembra, "Design of a 5G Network Slice Extension with MEC UAVs Managed with Reinforcement Learning", *IEEE Journal on Selected Areas in Communications*, Print ISSN: 0733-8716, Online ISSN: 1558-0008, October 2020, Vol. 38, No. 10, pp. 2356-2371, Published by IEEE, DOI: 10.1109/JSAC.2020.3000416, Available: <https://ieeexplore.ieee.org/document/9127468>.
 - [11] Sitong Zhang, Yibing Li and Qianhui Dong, "Autonomous navigation of UAV in multi-obstacle environments based on a deep reinforcement learning approach", *Applied Soft Computing*, Print ISSN: 1568-4946, Online ISSN: 1872-9681, January 2022, Vol. 115, Art. no. 108194, Published by Elsevier, DOI: 10.1016/j.asoc.2021.108194, Available: <https://www.sciencedirect.com/science/article/abs/pii/S1568494621010383>.
 - [12] Qiqi Ren, Omid Abbasi, Gunes Karabulut Kurt, Halim Yanikomeroglu and Jian Chen, "Caching and Computation Offloading in High Altitude Platform Station (HAPS) Assisted Intelligent Transportation Systems", *IEEE Transactions on Wireless Communications*, Print ISSN: 1536-1276, Online ISSN: 1558-2248, November 2022, Vol. 21, No. 11, pp. 9010-9024, Published by IEEE, DOI: 10.1109/TWC.2022.3171824, Available: <https://ieeexplore.ieee.org/document/9772280>.
 - [13] Siyuan Yang, Mondher Bouazizi, Tomoaki Ohtsuki, Yohei Shibata, Wataru Takabatake *et al.*, "Deep Reinforcement Learning Evolution Algorithm for Dynamic Antenna Control in Multi-Cell Configuration HAPS System", *Future Internet*, Online ISSN: 1999-5903, 12 January 2023, Vol. 15, No. 1, Art. no. 34, Published by MDPI, DOI: 10.3390/fi15010034, Available: <https://www.mdpi.com/1999-5903/15/1/34>.
 - [14] Helin Yang and Xianzhong Xie, "Energy-Efficient Joint Scheduling and Resource Management for UAV-Enabled Multicell Networks", *IEEE Systems Journal*, Print ISSN: 1932-8184, Online ISSN: 1937-9234, March 2020, Vol. 14, No. 1, pp. 363-374, Published by IEEE, DOI: 10.1109/JSYST.2019.2911895, Available: <https://ieeexplore.ieee.org/document/8708975>.
 - [15] Yun Lin, Meiyu Wang, Xianglong Zhou, Guoru Ding and Shiwen Mao, "Dynamic Spectrum Interaction of UAV Flight Formation Communication with Priority: A Deep Reinforcement Learning Approach", *IEEE Transactions on Cognitive Communications and Networking*, Online ISSN: 2332-7731, September 2020, Vol. 6, No. 3, pp. 892-903, Published by IEEE, DOI: 10.1109/TCCN.2020.2973376, Available: <https://ieeexplore.ieee.org/document/8995473>.
 - [16] Zechen Liu, Xin Liu, Victor C. M. Leung and Tariq S. Durrani, "Energy-Efficient Resource Allocation for Dual-NOMA-UAV Assisted Internet of Things", *IEEE Transactions on Vehicular Technology*, Print ISSN: 0018-9545, Online ISSN: 1939-9359, March 2023, Vol. 72, No. 3, pp. 3532-3543, Published by IEEE, DOI: 10.1109/TVT.2022.3219236, Available: <https://ieeexplore.ieee.org/document/9937149>.
 - [17] Abegaz Mohammed Seid and Aiman Erbad, "Multi-Agent RL for SDN-Based Resource Allocation in HAPS-Assisted IoV Networks", in *Proceedings of the 2023 IEEE International Conference on Communications (ICC 2023)*, 28 May-1 June 2023, Rome, Italy, pp. 1664-1669, Published by IEEE, DOI: 10.1109/ICC45041.2023.10279229, Available: <https://ieeexplore.ieee.org/document/10279229>.
 - [18] Seungmin Lee, Heejung Yu and Howon Lee, "Multiagent Q-Learning-Based Multi-UAV Wireless Networks for Maximizing Energy Efficiency: Deployment and Power Control Strategy Design", *IEEE Internet of Things Journal*, Online ISSN: 2327-4662, May 2022, Vol. 9, No. 9, pp. 6434-6442, Published by IEEE, DOI: 10.1109/JIOT.2021.3113128, Available: <https://ieeexplore.ieee.org/document/9539879>.
 - [19] Yingjie Cao, Yang Luo, Haifen Yang and Chunbo Luo, "UAV-Based Emergency Communications: An Iterative Two-Stage Multiagent Soft Actor-Critic Approach for Optimal Association and Dynamic Deployment", *IEEE Internet of Things Journal*, Online ISSN: 2327-4662, 15 August 2024, Vol. 11, No. 16, pp. 26610-26622, Published by IEEE, DOI: 10.1109/JIOT.2023.3329346, Available: <https://ieeexplore.ieee.org/document/10304279>.

- [20] Ruijin Ding, Feifei Gao and Xuemin Sherman Shen, "3D UAV Trajectory Design and Frequency Band Allocation for Energy-Efficient and Fair Communication: A Deep Reinforcement Learning Approach", *IEEE Transactions on Wireless Communications*, Print ISSN: 1536-1276, Online ISSN: 1558-2248, December 2020, Vol. 19, No. 12, pp. 7796-7809, Published by IEEE, DOI: 10.1109/TWC.2020.3016024, Available: <https://ieeexplore.ieee.org/document/9171468>.
- [21] Xunhua Dai, Jinhu Tu and Quan Quan, "Safety Assessment Approach to UAVs Based on Profust Safety Index and HIL Simulation", *IEEE/ASME Transactions on Mechatronics*, Print ISSN: 1083-4435, Online ISSN: 1941-014X, October 2024, Vol. 29, No. 5, pp. 3336-3347, Published by IEEE, DOI: 10.1109/TMECH.2023.3341851, Available: <https://ieeexplore.ieee.org/document/10376219>.
- [22] Ahmet Melih Ince, Ayse Elif Canbilen and Halim Yanikomeroglu, "AoI-Aware Resource Allocation with Deep Reinforcement Learning for HAPS-V2X Networks", in *Proceedings of the 2025 IEEE 36th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 1-4 September 2025, Istanbul, Turkiye, pp. 1-6, Published by IEEE, DOI: 10.1109/PIMRC62392.2025.11274843, Available: <https://ieeexplore.ieee.org/document/11274843>.
- [23] Haixia Peng and Xuemin Shen, "Multi-Agent Reinforcement Learning Based Resource Management in MEC- and UAV-Assisted Vehicular Networks", *IEEE Journal on Selected Areas in Communications*, Print ISSN: 0733-8716, Online ISSN: 1558-0008, January 2021, Vol. 39, No. 1, pp. 131-141, Published by IEEE, DOI: 10.1109/JSAC.2020.3036962, Available: <https://ieeexplore.ieee.org/document/9254093>.
- [24] Wilson Ayyanthole Nelson, Sreenivasa Reddy Yeduri, Ajit Jha, Abhinav Kumar and Linga Reddy Cenkeramaddi, "RL-Based Energy-Efficient Data Transmission Over Hybrid BLE/LTE/Wi-Fi/LoRa UAV-Assisted Wireless Network", *IEEE/ACM Transactions on Networking*, Print ISSN: 1063-6692, Online ISSN: 1558-2566, June 2024, Vol. 32, No. 3, pp. 1951-1966, Published by IEEE, DOI: 10.1109/TNET.2023.3332296, Available: <https://ieeexplore.ieee.org/document/10322686>.
- [25] Ji-He Kim, Ming-Chun Lee and Ta-Sung Lee, "Generalized UAV Deployment for UAV-Assisted Cellular Networks", *IEEE Transactions on Wireless Communications*, Print ISSN: 1536-1276, Online ISSN: 1558-2248, July 2024, Vol. 23, No. 7, pp. 7894-7910, Published by IEEE, DOI: 10.1109/TWC.2023.3345839, Available: <https://ieeexplore.ieee.org/document/10379044>.
- [26] Raouf Boutaba, Mohammad A. Salahuddin, Noura Limam, Sara Ayoubi, Nashid Shahriar, Felipe Estrada-Solano and Oscar M. Caicedo, "A comprehensive survey on machine learning for networking: evolution, applications and research opportunities", *Journal of Internet Services and Applications*, Online ISSN: 1869-0238, June 2018, Vol. 9, Article No. 16, Published by Springer, DOI: 10.1186/s13174-018-0087-2, Available: <https://link.springer.com/article/10.1186/s13174-018-0087-2>.
- [27] Chaoyun Zhang, Paul Patras and Hamed Haddadi, "Deep Learning in Mobile and Wireless Networking: A Survey", *IEEE Communications Surveys & Tutorials*, Print ISSN: 1553-877X, Online ISSN: 2373-745X, 2019, Vol. 21, No. 3, pp. 2224-2287, Published by IEEE, DOI: 10.1109/COMST.2019.2904897, Available: <https://ieeexplore.ieee.org/document/8666641>.
- [28] Federica Rinaldi, Helka-Liina Maattanen, Johan Torsner, Sara Pizzi, Sergey Andreev and Antonio Iera, "Non-Terrestrial Networks in 5G & Beyond: A Survey", *IEEE Access*, Online ISSN: 2169-3536, September 2020, Vol. 8, pp. 165178-165200, Published by IEEE, DOI: 10.1109/ACCESS.2020.3022981, Available: <https://ieeexplore.ieee.org/document/9193893>.
- [29] Mohamed-Amine Lahmeri, Mustafa A. Kishk and Mohamed-Slim Alouini, "Artificial Intelligence for UAV-Enabled Wireless Networks: A Survey", *IEEE Open Journal of the Communications Society*, Online ISSN: 2644-125X, 22 April 2021, Vol. 2, pp. 1015-1040, Published by IEEE, DOI: 10.1109/OJCOMS.2021.3075201, Available: <https://ieeexplore.ieee.org/document/9411810>.
- [30] Chi Harold Liu, Zheyu Chen, Jian Tang, Jie Xu and Chengzhe Piao, "Energy-Efficient UAV Control for Effective and Fair Communication Coverage: A Deep Reinforcement Learning Approach", *IEEE Journal on Selected Areas in Communications*, Print ISSN: 0733-8716, Online ISSN: 1558-0008, September 2018, Vol. 36, No. 9, pp. 2059-2070, Published by IEEE, DOI: 10.1109/JSAC.2018.2864373, Available: <https://ieeexplore.ieee.org/document/8432464>.
- [31] Zhihui Liu, Qiwei Zhang and Yi Su, "PPO-Based Joint Optimization for UAV-Assisted Edge Computing Networks", *Applied Sciences*, Online ISSN: 2076-3417, November 2023, Vol. 13, No. 23, Article No. 12828, Published by MDPI, DOI: 10.3390/app132312828, Available: <https://www.mdpi.com/2076-3417/13/23/12828>.
- [32] Javad Sabzehali, Vijay K. Shah, Qiang Fan, Biplav Choudhury, Lingjia Liu and Jeffrey H. Reed, "Optimizing Number, Placement, and Backhaul Connectivity of Multi-UAV Networks", *IEEE Internet of Things Journal*, Online ISSN: 2327-4662, November 2022, Vol. 9, No. 21, pp. 21548-21560, Published by IEEE, DOI: 10.1109/JIOT.2022.3184323, Available: <https://ieeexplore.ieee.org/document/9800925>.
- [33] Ting-Ting Zhang, Yan Chen, Ren-Zhi Dong, Tao Chen, Yan Liu *et al.*, "Autonomous decision-making of UAV cluster with communication constraints based on reinforcement learning", *Journal of Cloud Computing: Advances, Systems and Applications*, Online ISSN: 2192-113X, February 2025, Vol. 14, Art. no. 12, Published by Springer, DOI: 10.1186/s13677-025-00738-9, Available: <https://link.springer.com/article/10.1186/s13677-025-00738-9>.

- [34] Yuki Miyashita and Toshiharu Sugawara, "Analysis of coordinated behavior structures with multi-agent deep reinforcement learning", *Applied Intelligence*, Print ISSN: 0924-669X, Online ISSN: 1573-7497, September 2020, Vol. 51, pp. 1069-1085, Published by Springer, DOI: 10.1007/s10489-020-01832-y, Available: <https://link.springer.com/article/10.1007/s10489-020-01832-y>.
- [35] Tabish Rashid, Mikayel Samvelyan, Christian Schroeder de Witt, Gregory Farquhar, Jakob Foerster and Shimon Whiteson, "Monotonic Value Function Factorisation for Deep Multi-Agent Reinforcement Learning", *Journal of Machine Learning Research*, Print ISSN: 1532-4435, Online ISSN: 1533-7928, 2020, Vol. 21, No. 178, pp. 1-51, Published by JMLR, DOI: 10.48550/arXiv.1803.11485, Available: <https://www.jmlr.org/papers/v21/20-081.html>.
- [36] Xiaoying Sun, Jinchao Chen, Chenglie Du and Mengying Zhan, "Multi-Agent Deep Deterministic Policy Gradient Algorithm Based on Classification Experience Replay", in *Proceedings of the 2022 IEEE 6th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC '22)*, 3-5 October 2022, Beijing, China, pp. 988-992, Published by IEEE, DOI: 10.1109/IAEAC54830.2022.9929494, Available: <https://ieeexplore.ieee.org/document/9929494>.
- [37] Benjamin Freed, Aditya Kapoor, Ian Abraham, Jeff Schneider and Howie Choset, "Learning Cooperative Multi-Agent Policies with Partial Reward Decoupling", *IEEE Robotics and Automation Letters*, Online ISSN: 2377-3766, April 2022, Vol. 7, No. 2, pp. 890-897, Published by IEEE, DOI: 10.1109/LRA.2021.3135930, Available: <https://ieeexplore.ieee.org/document/9653841>.
- [38] Zhilan Zhang, Yufeng Wang, Yizhe Luo, Hang Zhang, Xiaorong Zhang *et al.*, "Iterative Trajectory Planning and Resource Allocation for UAV-Assisted Emergency Communication with User Dynamics", *Drones*, Online ISSN: 2504-446X, April 2024, Vol. 8, No. 4, Art. no. 149, Published by MDPI, DOI: 10.3390/drones8040149, Available: <https://www.mdpi.com/2504-446X/8/4/149>.



© 2026 by the author(s). Published by Annals of Emerging Technologies in Computing (AETiC), under the terms and conditions of the Creative Commons Attribution (CC BY) license which can be accessed at <http://creativecommons.org/licenses/by/4.0>.