applied sciences

MDPI

Article

A Dynamic Spectrum Decision-Making Method for UAV Swarms in Complex Electromagnetic Environments

Yu Chen *, Gang Wang *, Yangqin Wei, Jin Zang and Weijia Wu

Information and Navigation Institute, Air Force Engineering University, Xi'an 710077, China; afeuwyq@163.com (Y.W.); 19591587247@163.com (J.Z.); wuweijia@sina.com (W.W.)
* Correspondence: chenyu13a@alumni.nudt.edu.cn (Y.C.); wglxl@nudt.edu.cn (G.W.)

Abstract

To address frequency conflicts and electromagnetic interference (EMI) for UAV swarms operating in complex electromagnetic environments, this paper conducts a decoupling analysis of channel allocation and power selection in the spectrum decision-making process and proposes a spectrum decision-making method for UAV swarms in such scenarios. A complex electromagnetic scenario model is established by incorporating three dominant interference characteristics, including internal friendly frequency conflicts, adversarial jamming interference, and time-varying background electromagnetic fluctuations. On this basis, performance evaluation metrics are formulated to quantify the command–communication and radar reconnaissance of UAV nodes. Then we propose an Ant Colony Optimization-Double Deep Q-Network (ACO-DDQN) algorithm, where the ACO module implements global channel assignment and the DDQN module conducts adaptive transmit power selection. This decoupled design enables the algorithm to obtain an approximately optimal spectrum decision scheme for coordinated operations. Simulation experiments under different swarm sizes and jamming intensities verify the feasibility and effectiveness of the proposed algorithm.

Keywords: complex electromagnetic environment; unmanned aerial vehicle (UAV) swarms; frequency conflicts; spectrum decision-making

1. Introduction

As a unique strategic resource, the electromagnetic spectrum has become a vital pillar of informatization, intelligent systems, and unmanned operations [1–4]. More efforts should be devoted to the field of the electromagnetic spectrum. It is growing vital for strengthening national power and safeguarding national defense security. Activities in the electromagnetic spectrum span various areas, including command–communications, reconnaissance and sensing, navigation and positioning, jamming and countermeasures, and fire control. Mobile big data, robotics, unmanned aerial vehicles (UAVs), and the Internet of Things (IoT) have advanced rapidly in recent years. In particular, driven by these emerging technologies, frequency usage demands of electromagnetic spectrum devices have grown explosively, rendering the electromagnetic space far more complex [5–8]. In complex electromagnetic environments, limited spectrum resources, numerous spectrum-utilizing devices and diverse frequency usage patterns, together with electromagnetic deception and interference, create substantial challenges for spectrum management and spectrum security. With the advancement of intelligent and unmanned technologies, the use of UAVs has become an increasingly important means [9]. However, UAVs rely heavily on

Check for updates

Academic Editor: Christos Bouras

Received: 7 September 2026
Revised: 9 October 2026
Accepted: 10 October 2026
Published: 14 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

the electromagnetic spectrum. Precise and efficient electromagnetic spectrum management has long served as a prerequisite for UAV swarm coordination. Therefore, exploring effective electromagnetic spectrum management strategies is essential for establishing an orderly spectrum environment and addressing key spectrum competition and interference challenges in UAV swarm operations.

Electromagnetic spectrum management entails the rational planning and dynamic allocation of limited spectrum resources, extensive spectrum-dependent devices, and diverse frequency-usage activities. It lays a solid foundation for achieving electromagnetic superiority. Although conventional manual spectrum management methods, including demand aggregation and usage plan formulation, can mitigate spectrum conflicts and standardize spectrum utilization, they suffer from coarse management granularity, low operational efficiency, and poor resource utilization. In particular, when faced with complex and dynamic electromagnetic environments, manual methods struggle to respond effectively. Therefore, intelligent spectrum management techniques are urgently required. To address these limitations, numerous studies have focused on intelligent spectrum management in complex electromagnetic environments. For instance, Reference [10] adopts a game-theoretic framework. It designs spectrum access and utilization strategies based on non-cooperative game theory, and constructs a utility function centered on spectrum utilization efficiency. A sequential Overlapping Coalition Formation (OCF) game model is proposed in Reference [11], which enables UAVs to arrange task execution sequences by task priority and optimize resource allocation. Reference [12] establishes a partial overlapping coalition formation game model. To achieve the interdependence between task assignment and resource decision-making, the authors propose a quantum heuristic algorithm. Drawing on graph theory, reference [13] enables dynamic spectrum access and improves the system's transmission rate and throughput. Artificial intelligence techniques have also been widely used to address spectrum management issues. For instance, in reference [14], multi-agent and deep reinforcement learning techniques are utilized for channel constraint issues. For large-scale UAV swarm systems, reference [15] proposes a distributed, heterogeneous multi-agent reinforcement learning algorithm. The swarm is partitioned into multiple clusters, where each UAV is equipped with an independent learning agent. By combining proximal policy optimization (PPO) for the swarm leader and independent PPO for intra-cluster members, dynamic spectrum sharing across the entire swarm can be realized. Zhao et al. [16] addressed the problem of intelligent coordination in the electromagnetic spectrum by designing a hybrid reward function based on a credit allocation mechanism. This approach ensures overall reward while also taking into account the individual agent payoffs. It improves the system's performance by reducing the number of idle nodes. To address complex electromagnetic interference, reference [17] proposed the Greedy-PPO algorithm for intelligent spectrum-sharing. The algorithm derives an approximately optimal spectrum strategy. Reference [18] targets the well-documented convergence instability that plagues conventional DRL methods in non-stationary jamming conditions. The authors put forward a Curriculum Learning-assisted Multi-Agent Proximal Policy Optimization (CL-MAPPO) scheme built on a Centralized Training with Decentralized Execution (CTDE) framework. This design helps resolve spectrum conflicts that emerge among UAVs operating inside the same swarm. Reference [19] proposes a novel heterogeneous MARL method for joint UAV spectrum optimization. It pairs the framework with a tailored reward function to deliver noticeable performance improvements. Reference [20] developed an orthogonal-gradient-based MARL algorithm for UAV swarm communication scenarios. The method mitigates gradient instability during training and achieves optimized spectrum resource allocation across the swarm. Reference [21] designs a transferable DQN network that builds independent experience pools for both source

and target domains. This method effectively lowers spectrum conflict rates and improves overall spectrum utilization efficiency. In contrast, reference [22] exploits the inherent local spatial filtering properties of CNNs to directly capture signal propagation patterns, circumventing the constraints of traditional path loss models. For UAV air-to-ground communication scenarios, reference [23] employs machine learning methods alongside a novel channel model and applies a clustering algorithm based on a Bayesian–Gaussian mixture model to UAV-to-ground communications. Despite its effective performance in designated scenarios, this method lacks sufficient generalization capability. Reference [24] proposes a self-organized anti-jamming reinforcement learning method. By optimizing UAV deployment and resource allocation, it supports high-quality communication service delivery. References [25,26] carry out in-depth investigations on UAV networking. Targeting cooperative networking of heterogeneous unmanned platforms, they exploit UAVs' rapid deployment capability to set up emergency communication links and guarantee communication reliability for swarm coordination.

Complex electromagnetic conditions and scarce spectrum resources have brought new challenges to UAV swarm operations. Mission requirements are becoming increasingly intricate, spectrum-dependent platforms continue to proliferate, and adversarial jamming grows more severe. These factors make spectrum decision-making far more demanding. It is necessary to adopt more accurate channel allocation and power selection strategies to boost the effectiveness of spectrum decision-making. However, existing studies in this domain have primarily focused on frequency-domain conflicts. Few studies have explored UAV swarm spectrum decision-making under external interference from a joint frequency and power domain perspective. Such a comprehensive scheme needs to account for both internal mutual interference and dynamic external interference simultaneously. From a methodological perspective, existing approaches to this problem can be roughly divided into two categories, each with its own limitations. Heuristic optimization methods can cope with the combinatorial nature of channel assignment, but they rely on iterative population-based search and therefore respond slowly to fast-varying jamming, which limits their real-time anti-jamming capability. Learning-based methods, such as deep and multi-agent reinforcement learning, can respond rapidly once trained; however, the joint channel-power action space grows combinatorially with the number of UAV nodes, which leads to slow convergence and even training failure in large-scale swarms.

This paper addresses the spectrum decision-making problem for UAV swarms in complex electromagnetic environments. Taking both frequency and power domains into account, we propose an ACO-DDQN spectrum decision-making scheme to optimize channel allocation and power selection for spectrum-using devices under constrained spectrum resources, in which channel allocation and power selection are decoupled in a hierarchical manner: the ACO module performs global channel assignment, while the DDQN module conducts adaptive power selection conditioned on the channel scheme. In this way, the proposed scheme combines the global search ability of heuristic optimization with the fast response capability of reinforcement learning, and it differs from both categories of existing methods in how the two decision dimensions are handled. Experiments confirm that the proposed ACO-DDQN scheme converges effectively in complex UAV swarm mission scenarios. Compared with traditional heuristic optimization algorithms and multi-agent reinforcement learning schemes, it delivers superior anti-jamming capability and enhanced spectrum decision performance. The main contributions of this paper are summarized as follows:

- A red–blue adversarial scenario model for UAV swarm spectrum decision-making in complex electromagnetic environments, together with a corresponding set of evaluation metrics;

- An ACO-DDQN spectrum decision algorithm. It decouples channel allocation and power selection for UAV swarms to handle spectrum decision tasks in complex electromagnetic environments;
- Simulation experiments carried out under different jamming intensities and varying UAV swarm scales. Comparisons with heuristic and multi-agent reinforcement learning baselines, including the recent MAPPO-based approaches, verify the feasibility and effectiveness of our proposed approach.

2. Problem Formulation

2.1. Spectrum Usage Scenarios

UAV swarms typically consist of multiple unmanned aerial platforms networked together to perform reconnaissance, strikes, relay communications, long-range delivery, and jamming and countermeasures. Frequency usage patterns vary noticeably across the swarm's mission phases, namely the takeoff-and-landing phase, cruise phase, and mission execution phase. During takeoff and landing, spectrum demand is dominated by takeoff support systems. For the cruise phase, frequencies for navigation and flight path maintenance account for most frequency usage. In the mission phase, spectrum usage centers on mission payloads. However, command and communication links remain active across all stages. To streamline the analysis, this paper primarily focuses on the spectrum usage issues related to command and communication as well as reconnaissance and detection during UAV swarm reconnaissance missions. During mission execution, each UAV node maintains a data link with the ground control center. These links carry mission commands dispatched from the ground and return real-time, high-fidelity reconnaissance data from the aerial nodes. Any interference to these links disrupts the UAV nodes' ability to respond to mission commands, putting mission continuity at serious risk. On-board reconnaissance radars operate within specific frequency bands to image target areas and collect precise measurements of distance, speed, and position. If the sensors fail due to electromagnetic jamming, the UAV's capabilities for reconnaissance, identification, and tracking will be compromised, which will similarly affect mission execution.

To enable a rigorous investigation of the spectrum for command–communications, as well as reconnaissance and detection during UAV swarm reconnaissance missions, this paper constructs a scenario for aerial electromagnetic countermeasures featuring UAV swarms, in which the red and blue sides engage in electromagnetic countermeasures. The red team is primarily equipped with fixed and mobile jammers, which deliver electronic jamming to suppress the blue team's equipment and disrupt its key electromagnetic missions—command–communications, reconnaissance, and detection. The blue team primarily utilizes UAV swarms to conduct reconnaissance and surveillance missions. Table 1 lists all scenario variables and corresponding symbolic parameters [16].

2.2. Complex Electromagnetic Environment

2.2.1. Team Blue Frequency Conflicts

The blue team's primary assets include ground control centers and a deployed aerial UAV swarm. The frequency allocation module primarily addresses two dimensions: command–communications and radar reconnaissance. The command–communications module facilitates command, control, and information transmission between the ground control center and individual UAV nodes. For radar reconnaissance, on-board sensor modules enable the UAVs to conduct detection and reconnaissance across the mission area. A total of m non-overlapping orthogonal sub-bands $F \in \{f_1, f_2, \cdots, f_m\}$ are defined in the scenario. Frequency competition exists not only among different functional modules of a single UAV, but also co-channel interference between different UAVs.

Table 1. Nomenclature of scenario parameter symbols.

Symbol	Description
F	Non-overlapping orthogonal sub-bands
Con	Set of Team Blue ground control centers
$Node$	Set of Team Blue UA
T	Index of mission time slot
Jam	Set of Team Red jammers
$(x_{i,t}, y_{i,t})$	Coordinates of the node i in time slot t
$R_{\max}$	Maximum detection range for radar reconnaissance
$P_{t,f}^{jam}(j,n)$	Jamming power generated by the jammer j at node n in time slot t on sub-band f
$P_{bg,t}$	Background electromagnetic interference in time slot t
$P_{tx,p}$	Team Blue radar pulse power
$P_{t,f}^{\mathrm{c}}(n',n)$	Command and communication power received at node n from node n' in time slot t
$P_{t,f}^{\mathrm{r}}(n',n)$	Radar reconnaissance power received at node n from node n' in time slot t
P_{th}	Detection threshold power value of the Team Blue radar reconnaissance module
R_{th}	Detection threshold range of the Team Blue radar reconnaissance module
$SINR_{th}$	Threshold value of signal-to-interference-plus-noise ratio
α	Weighting coefficient for the Team Blue radar reconnaissance module
β	Weighting coefficient for the Team Blue command and communication module

The set of the blue team's ground control centers is denoted as $Con \in \{con_1, con_2, \cdots, con_n\}$, and the set of UAVs is denoted as $Node \in \{node_1, node_2, \cdots, node_3\}$. To meet the demands of distributed mission scenarios, the UAV swarm is deployed in a locally concentrated manner. For this reason, the UAV swarm is split into multiple clusters. Each cluster is assigned a ground control center, which commands and controls intra-cluster UAVs and receives radar detection data relayed by those aerial nodes. For analytical simplicity, this paper focuses on the spectrum decision-making for UAV swarms during the hovering reconnaissance phase.

We model the entire mission as a sequence of $T \in \{1, 2, \cdots, n\}$ time slots. In each slot, every UAV can simultaneously activate its command–communications module and radar reconnaissance module, yet on different frequency bands. The time slots of an episode form one decision period. At the beginning of each period, every node is assigned its command–communication channel and radar reconnaissance channel, and these channels remain unchanged until the period ends. The transmit power of each module, however, is adjusted in every time slot. This setting reflects engineering practice: changing the channel requires retuning and link re-establishment, so it is carried out on a slower timescale than power adjustment.

2.2.2. Red Team Jamming

The red team executes electromagnetic jamming via a mix of fixed and mobile jammers. The set of jammers is denoted as $\mathrm{Jam} \in \{jam_1, jam_2, jam_3, jam_4\}$, with jam_1, jam_2 for the fixed jammers and jam_3, jam_4 for the mobile jammers. Each jammer delivers jamming suppression on one sub-band f in each slot, keeping the jamming sub-band and power unchanged during that slot. Fixed jammers are deployed at the boundaries of the mission

area and remain stationary throughout the mission. Mobile jammers move along the pre-planned mission path, as specified by Equation (1).

$$\begin{cases} x_{i,t} = x_{i,t-1} + \Delta x_{i,t} \\ y_{i,t} = y_{i,t-1} + \Delta y_{i,t} \end{cases} \tag{1}$$

where $x_{i,t} \in [0, 120KM]$, $y_{i,t} \in [0, 120KM]$, $\Delta x_{i,t}$ and $\Delta y_{i,t}$ represent the position offsets of the red team's mobile jammers at time slot t. At each slot, each jammer first scans all available channels at its current position to measure received power levels. It automatically locks onto the frequency band with the highest detected signal strength, then delivers jamming suppression with rated power.

2.2.3. Background Electromagnetic Interference

Alongside the self-interference and mutual conflicts among the blue team's assets mentioned above, plus fixed and mobile jamming from the red team, background electromagnetic interference persists throughout the entire mission. The background electromagnetic environment consists of civilian frequency-using devices and natural electromagnetic noise. To realistically simulate the dynamic changes in civilian emitters and natural electromagnetic noise during the mission, the entire mission area is divided into multiple sub-regions in our scenario. Background interference differs notably across sub-regions, while the interference magnitude in each sub-region varies randomly over time.

2.3. Spectrum Decision Evaluation Metrics

In line with spectrum decision requirements and drawing on the metric definitions presented in References [16,17], we develop dedicated performance evaluation metrics for the radar reconnaissance and command-and-communication modules of the blue team's ground control centers and UAV swarm. Based on these module-level metrics, we further derive an overall effectiveness evaluation metric for spectrum decision-making.

2.3.1. Radar Reconnaissance Metrics

The radar reconnaissance module is primarily dedicated to detection and surveillance. Its core task is to detect and monitor potential threat targets within the mission area. The maximum detection distance $R_{\max}$ and detection probability P_d are often used as performance metrics for radar reconnaissance.

$$\begin{cases} R_{\max} = \sqrt[4]{\frac{P_{tx,p} A_{eff}^2 \sigma}{4\pi\lambda^2 P_{th}}} \\ P_d = Q_1(\sqrt{2SINR}, \tau) \end{cases} \tag{2}$$

where A_{eff} is the effective receiving aperture area of the radar antenna; σ represents the radar cross-section; $P_{tx,p}$ denotes the peak transmission power of the radar pulse; λ stands for the wavelength of the radar signal; P_{th} is the minimum detectable power of the radar reconnaissance module's receiver; τ refers to the detection threshold.

$$P_{th} = SINR_{\min} \cdot \left(P_N + I_{t,n,f} \right) \tag{3}$$

where $SINR_{\min}$ represents the minimum detectable signal-to-interference-plus-noise ratio of the radar receiver; P_N is the thermal noise power of the radar receiver; $I_{t,n,f}$ represents the total power received at node n over frequency band f. It refers to the superposition of power received at node n at time slot t, including radar reconnaissance and command–

communication modules of the other nodes n', jamming power generated by each red team jammer Jam, and the background environmental electromagnetic interference.

$$I_{t,n,f} = \sum_{n' \neq n} P^{c}_{t,f}(n', n) + \sum_{n' \neq n} P^{r}_{t,f}(n', n) + \sum_{j=1,2,3,4} P^{jam}_{t,f}(j, n) + P_{bg,t} \tag{4}$$

where $P^{r}_{t,f}(n', n)$ and $P^{c}_{t,f}(n', n)$ represent radar reconnaissance and command–communication power received at node n from node n' at time slot t; $P^{jam}_{t,f}(j, n)$ denotes the jamming power received at node n at time slot t; $P_{bg,t}$ stands for the background environmental electromagnetic interference at time slot t.

Accordingly, a node is classified as an effective detection node when its radar reconnaissance module meets two criteria: its detection distance reaches the threshold R_{th}, and its detection probability P_d attains the predefined reference level. Building on this classification, the evaluation function for the detection and perception performance of each node's radar reconnaissance module is defined as follows:

$$F^{r}_{t,n,f} = \begin{cases} 1, & R^{t,n,f}_{\max} > R_{\text{th}} \cap P_d > P_{d,ref} \\ \min(\frac{R^{t,n,f}_{\max}}{R_{\text{th}}}, 1) \cdot \min(\frac{P_d}{P_{d,ref}}, 1), & otherwise \end{cases} \tag{5}$$

where $R^{t,n,f}_{\max}$ is the maximum reconnaissance detection distance of node n over frequency band f at time slot t, and $P_{d,ref}$ refers to the reference detection probability.

2.3.2. Command–Communication Metrics

The command–communication modules of each UAV node handle command, control, and information transmission between the ground control center and the swarm. The wireless channel is simplified to a line-of-sight transmission model. The path loss is calculated as follows:

$$PL_{\text{LoS}} = 20 \log_{10}(d) + 20 \log_{10}(f) - 27.55 \tag{6}$$

where the path loss PL_{LoS} is measured in decibels (dB), the transmission distance d is measured in kilometers (km), and the carrier center frequency f is measured in megahertz (MHz).

The Shannon channel capacity is adopted as the evaluation metric for the command–communication module. For a given channel bandwidth B, the channel capacity between the ground control center con and the UAV node n over frequency band f at time slot t is expressed as

$$\begin{cases} C_{con,t,f} = B \log_2(1 + SINR_{\text{con},t,f}) \\ SINR_{\text{con},t,f} = \frac{P^{c}_{t,f}(n,con)}{\sum\limits_{n' \neq n} P^{c}_{t,f}(n',con) + \sum\limits_{n} P^{r}_{t,f}(n,con) + \sum\limits_{j=1,2,3,4} P^{jam}_{t,f}(j,con) + P_{bg,t}} \end{cases} \tag{7}$$

C_{th} denotes the threshold of channel capacity. When the channel capacity between the ground control center and the UAV node meets or exceeds the threshold C_{th}, the command–communication between that control center and the corresponding UAV is regarded as valid. The performance evaluation function for the command–communication module of each control center is formulated as follows:

$$F^{c}_{t,n,f} = \begin{cases} 1, & C_{con,t,f} \geq C_{th} \\ \frac{C_{con,t,f}}{C_{th}}, & C_{con,t,f} < C_{th} \end{cases} \tag{8}$$

2.3.3. Overall Metrics for Spectrum Decision-Making

In complex electromagnetic environments, the spectrum decision-making effectiveness of the blue team is affected by multiple factors, including red team jamming, co-frequency mutual interference, and background electromagnetic interference.

It is necessary not only to mitigate mutual interference caused by frequency band sharing among UAV nodes, but also to ensure the ability to make decisions effectively under interference. When making spectrum decisions, nearby UAV nodes should adopt different frequency bands to minimize mutual interference. Meanwhile, to counter the red team dynamic jamming, specific anti-jamming decision-making must be considered to achieve dynamic and effective spectrum management. Accordingly, this paper formulates the key research problem as maximizing the effectiveness of the UAV swarm's radar reconnaissance and command–communication modules in complex electromagnetic environments, so as to enable dynamic spectrum decision-making. The optimization model for the spectrum decision-making objective is defined as follows:

$$\begin{aligned} &\max O(t,n,f) = \alpha \sum_{t}\sum_{n} F^{r}_{t,n,f} + \beta \sum_{t}\sum_{n} F^{c}_{t,n,f} \\ &\text{s.t.} \begin{cases} \sum\limits_{f} x_{t,n,f} \le 1, \quad \forall n \quad (9a) \\ \sum\limits_{f} \varphi_{t,n,f} \le 1, \quad \forall n \quad (9b) \\ x_{t,n,f} + \varphi_{t,n,f} \le 1, \quad \forall n, f \quad (9c) \\ P^{c}_{t,f} \in P^{c} \quad (9d) \\ P^{r}_{t,f} \in P^{r} \quad (9e) \end{cases} \end{aligned} \tag{9}$$

where $x_{t,n,f} \in \{0,1\}$ is the binary assignment indicator variable for the command–communication module: it takes the value 1 if the command–communication module of node n is assigned to frequency band f at time slot t, and 0 otherwise. $\varphi_{t,n,f} \in \{0,1\}$ is the binary assignment indicator variable for the radar reconnaissance module: it takes the value 1 if the radar reconnaissance module of node n is assigned to frequency band f at time slot t, and 0 otherwise. Constraint (9a) indicates that each node n may select only one frequency band for command–communication per time slot. Constraint (9b) means that node n may select one frequency band for radar reconnaissance per time slot. Constraint (9c) specifies that the UAV node n's command–communication and radar reconnaissance modules must operate on different frequency bands within a single time slot. Constraints (9d) and (9e) represent the allowable transmission power ranges for the command–communication module and the radar reconnaissance module of each UAV node. α and β represent the performance weighting ratios of the radar reconnaissance and the command–communication modules. $F^{r}_{t,n,f}$ is the performance evaluation value of the command–communication module when node n is assigned to frequency band f at time slot t, and $F^{c}_{t,n,f}$ is the performance evaluation value of the radar reconnaissance module when node n is assigned to frequency band f at time slot t. Both values vary dynamically with the assigned channel f, reflecting the channel dependence of the objective function.

3. ACO-DDQN Spectral Decision Algorithm

3.1. Decision-Making Algorithm Framework

For large-scale UAV nodes, the joint actions of frequency band allocation and power selection suffer from the "curse of dimensionality". Coupled with the dynamically varying jamming imposed by the red team, this high-dimensional complexity makes it difficult for centralized optimization methods to deliver effective anti-jamming decisions. To address

this challenge, this paper proposes an ACO-DDQN dynamic spectrum decision-making algorithm. The ACO-DDQN algorithm integrates the heuristic Ant Colony Optimization (ACO) algorithm with the DDQN algorithm. Specifically, the ACO component optimizes the channel allocation in spectrum decision-making, while the DDQN module determines power levels. Leveraging the concept of hierarchical decoupling, the proposed algorithm achieves rational allocation of limited spectrum resources and alleviates frequency contention among UAV swarms. It also delivers robust anti-jamming capability in complex electromagnetic environments and satisfies the dynamic spectrum response requirements of mission scenarios. The design of the ACO-DDQN architecture follows directly from the heterogeneous structure of the decision problem in Equation (9), rather than from a simple juxtaposition of two existing algorithms. On one hand, channel allocation is a swarm-level combinatorial assignment problem governed by the coupling constraints (9a)–(9c), whereas power selection is a node-level decision that must react to jamming within each time slot. Pheromone-based global search is well suited to the former, and value-based reinforcement learning is well suited to the latter, which motivates the pairing of ACO with channel allocation and DDQN with power selection. This decoupling also substantially reduces the learning burden: with the candidate sets listed in Table 2, a joint channel-power decision contains $6 \times 5 \times 5 \times 3 = 450$ discrete combinations per node, whereas after decoupling the DDQN module only needs to choose among 15 power combinations per node, with the combinatorial complexity absorbed by the ACO module. On the other hand, the two modules do not operate independently. The channel assignment generated by the ACO module is encoded as one-hot features and concatenated into the observation of the DDQN module, so that each power decision is conditioned on the current channel scheme. Conversely, the environmental feedback of the joint channel-power decision is returned to update the pheromones of the ACO module, so that the channel assignment evolves according to the quality of the joint decisions. Both modules are driven by the same global reward defined in Equation (12), which is consistent with the optimization objective in Equation (9).

Table 2. Experimental parameter settings.

Parameter	Value
Mission area size	120 km × 120 km
Number of time slots per episode	20
Central frequency (GHz)	[0.6, 0.9, 1.2, 1.5, 1.8, 2.3]
Radar reconnaissance transmit power (kW)	[0.5, 0.75, 1]
Command–communication transmit power (W)	[20, 40, 60, 80, 100]
Jamming Power (W)	[50, 100, 200]
Weighting coefficient α	0.5
Weighting coefficient β	0.5
Number of jammers	4
$SINR_{th}$	1
R_{th}(km)	15
Number of background-noise sub-regions	4
Background environmental noise $P_{bg,t}$(W)	$[1 \times 10^{-13}, 5 \times 10^{-13}, 1 \times 10^{-11}, 2 \times 10^{-11}]$
Radar duty cycle	0.03
Channel bandwidth B	20 MHz
Reference detection probability	0.9
Radar cross-section σ (m^2)	20

- Action space: the spectrum decision of UAV node n at time-slot t consists of the power levels and channel selections for each module. Accordingly, the action space is defined as follows:

$$\begin{cases} A = \{a_t^1, a_t^2, \cdots, a_t^n | n \in Node\} \\ a_t^n = [f_{t,n}^r, f_{t,n}^c, p_{t,n}^r, p_{t,n}^c] \end{cases} \tag{10}$$

  where a_t^n denotes the actions of the UAV node n at time slot t, encompassing the channel selection and power configuration for radar reconnaissance and command–communication modules;
- Observation state: the observation state of a UAV node n at time slot t is composed of the distance from the node to its cluster's ground control station, as well as the total interference power on each frequency band. This interference stems from other co-frequency interfering nodes, various jammers, and background electromagnetic interference. Accordingly, the state space is defined as follows:

$$O_n(t) = \{dis_{n,t}, I_{n,t}\} \tag{11}$$

- Reward function: the reward function is consistent with the overall performance evaluation metrics for radar reconnaissance and command–control communication modules. It is computed based on the action of each node at time slot t:

$$R(t) = \alpha \sum_n F_{t,n}^{\mathrm{r}} + \beta \sum_n F_{t,n}^c \tag{12}$$

The global reward in Equation (12) is decomposed at the node level. The individual reward of node n at time slot t is defined as. The individual rewards are used both for the per-node updates of the DDQN module and for the pheromone increments of the ACO module, so that both modules remain consistent with the overall objective in Equation (9).

3.2. Algorithm Design Procedure

The ACO algorithm is an optimization algorithm that simulates the foraging behavior of ants. It utilizes a simulated ant colony to perform random exploration in the task space and employs a pheromone-based positive feedback mechanism to guide the ants' collaborative behavior. Given limited spectrum resources, node-level local optimization alone cannot effectively mitigate co-frequency interference among nodes. Instead, overall optimization must be performed at the global level. Frequency band assignment is mapped to the ants' path construction process, where multiple ants generate candidate solutions in parallel. As pheromones accumulate on high-quality frequency band combinations, subsequent ants are guided to prioritize these high-pheromone paths, obtaining an approximately optimal frequency band assignment scheme. The evaporation coefficient ρ controls how quickly outdated experience is forgotten, while the elite-preserving reset retains half of the accumulated pheromones across episodes; together they balance experience retention against adaptability to the evolving jamming strategy. The specific pheromone updating rules are expressed as follows:

$$\begin{cases} \tau_{ij}(t+1) = (1-\rho)\tau_{ij}(t) + \Delta\tau_{ij} \\ \Delta\tau_{ij} = \sum_k \Delta\tau_{ij}^k \end{cases} \tag{13}$$

where $\rho(\rho \in (0,1))$ denotes the pheromone evaporation coefficient, and $\Delta\tau_{ij}$ denotes the total amount of pheromone released along the path ij in this iteration. At the beginning of each episode, 50 candidate channel assignments are constructed: 30% of them are sampled according to the pheromone distribution and the heuristic function, while the remaining 70%

are generated randomly to preserve diversity. The historically best assignment is reused with a probability of 0.3. The fitness of each candidate is the global reward in Equation (12), so that the fitness reflects the quality of the channel assignment itself. After the joint decision is executed, the pheromone increment of each node is proportional to its individual reward, and an additional reinforcement is granted to the best assignment found so far. The pheromone values are clipped to a bounded interval to avoid premature convergence.

In the proposed ACO-DDQN hybrid framework, the frequency band assignment for UAV nodes is treated as the decision output of the ACO component. According to the observation states of all nodes $O_n(t)$, the ACO module derives frequency band assignments for each UAV node. Based on this allocation result, the frequency bands of each node together with the global observation states are fed into the DDQN module to perform power level selection. Each node is equipped with an independent DDQN agent. The network takes as input the global observation in Equation (11) concatenated with the one-hot encoding of the current channel assignment, and it contains two hidden layers with 128 units and ReLU activations. The output layer has 15 units, corresponding to the combinations of the five command–communication power levels and the three radar power levels in Table 2. The network is trained with the Adam optimizer and the MSE loss, and gradient clipping is applied to stabilize training. The detailed implementation procedure is illustrated in Figure 1, where the neural network parameters are updated through gradient backpropagation.

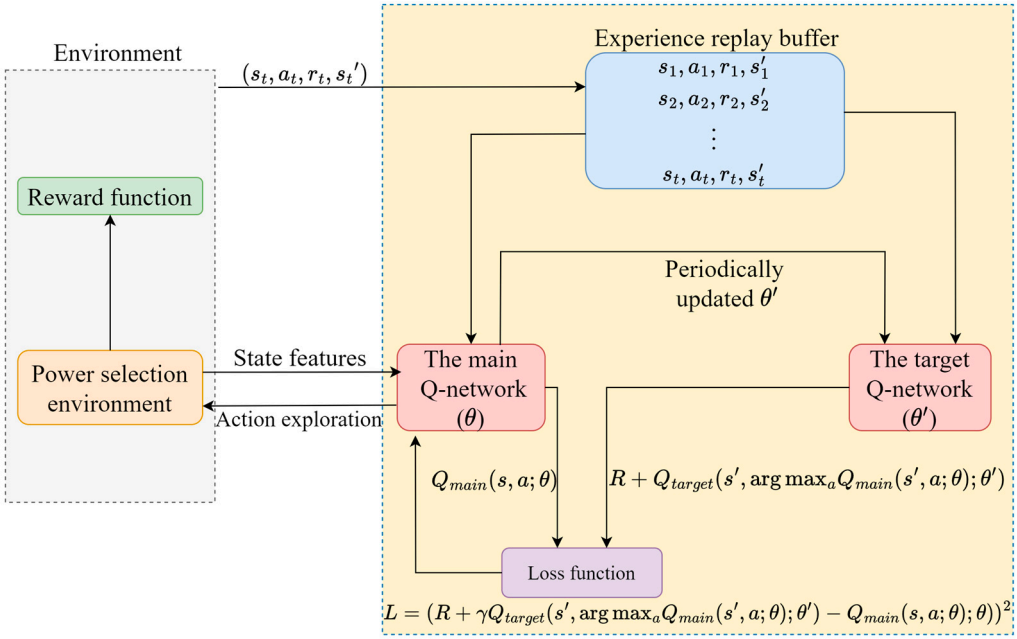


Figure 1. DDQN power selection.

By this hierarchical decoupling design, the ACO and the DDQN modules respectively perform frequency band allocation and power selection, which together produce the joint action **A**. The power level selected by the DDQN is jointly influenced by the global environmental state and the frequency allocation output by the ACO module. Once the ACO module completes frequency band allocation, the DDQN responds promptly based on the frequency band scheme. It calculates the global feedback reward R_t, updates the state to s_1', and then feeds back to the ACO module's frequency allocation iteration. Throughout the network training process, experience samples (s_t, a_t, r_t, s_t') are stored in the experience replay buffer. Random sampling is adopted to update the ACO pheromones and the DDQN network parameters, respectively. The pseudocode of the proposed ACO-DDQN algorithm is presented in Algorithm 1. The ACO module re-optimizes the channel assignment once

at the beginning of each episode, and the assignment is then kept fixed within the episode. Apart from the switching cost discussed in Section 2.2.1, this also makes the power-learning subproblem easier to train, since the DDQN module always learns against a stable channel scheme. The channel assignment itself remains adaptive across episodes. The pheromones are updated by the per-slot feedback and are partially reset at the start of each episode. The past experience is retained while the colony can still follow changes in the jamming strategy.

Algorithm 1: ACO-DDQN Spectrum Decision Algorithm

Input: Env, agents, aco
Output: Reward
1. Initialize environmental parameters
2. for each episode do:
3. s ← env.reset()
4. τ ← ρ_elite τ + (1 − ρ_elite) · τ_init // Partially reset the ACO pheromones
5. F = (f_c,f_r) ← aco.sample and select best(env, M) // ACO first complete the frequency band assignment
6. for each t do:
7. s_freq ← AppendFreqOneHot(s,f_c,f_r) // Concatenate frequency bands
8. for each device do:
9. a_i ← ε_i-greedy(Q_i(s_freq)) // DDQN conducts power level action selection
10. end for
11. s′,R,{r_i},done ← env.step_freq_power(F, {a_i}) // Execute environmental interactions
12. s′_freq ← AppendFreqOneHot(s′,f_c,f_r)
13. for each device do
14. D_i ← {(s_freq, a_i, r_i, s′_freq, done)} // Store experience samples into the experience replay buffer
15. ℓ_i ← agent_i.update() // update DDQN network parameter
16. if ℓ_i > 0 then
17. ep_loss ← ep_loss + ℓ_i
18. update_count ← update_count + 1
19. end if
20. end for
21. {r_i_ACO} ← env.evaluate_f_assignment(F)
22. aco.update_pheromone(F,{r_i_ACO}) // Update ACO pheromones
23. s ← s′
24. if done then break
25. end for
26. end for

During the training phase, the ant colony gradually develops preferences for frequency band allocation under interference conditions, while the DDQN adapts the power selection for a given frequency band scheme. Through iterative interactions between the two modules, the overall system reward improves steadily. Once the ACO-DDQN network has converged, the algorithm can support two-timescale spectrum decision-making, with channel assignment renewed at the episode level and power adjusted at every time slot. The specific execution steps of the entire algorithm are outlined as follows:

Step 1: Initialize the environmental parameters and activate the ACO module. Each ant assigns the radar reconnaissance and command–communication frequencies for the UAV nodes.

Step 2: The DDQN module learns and trains based on the external environment and the frequency band allocation scheme obtained by the ACO module. Then it determines the optimal power action under the current scheme and generates the joint action space.

Step 3: The external environment returns rewards for the joint action and updates the state. The ACO module updates the pheromones, while the DDQN module updates the network parameters.

Step 4: Terminate network updating once training is completed, yielding the spectrum decision-making scheme in complex electromagnetic environments.

4. Experiments and Result Analysis

4.1. Simulation Scenarios and Parameters

This section presents simulation experiments to validate the feasibility of the proposed spectrum decision algorithm. In the constructed complex electromagnetic scenario, the proposed ACO-DDQN algorithm is compared with three baselines: the conventional DDQN algorithm, the ACO algorithm, and the MAPPO algorithm. The DDQN baseline represents learning-based methods with a joint action space, the ACO baseline represents pure heuristic search, and MAPPO is included as a representative multi-agent reinforcement learning method [18]. All experiments are carried out under different numbers of UAV nodes and diverse jamming strategies deployed by the red team. The horizontal spatial extent of the simulation mission area is bounded within Region A (120 km × 120 km), as illustrated in Figure 2. The parameter settings for the simulation experiments are listed in Table 2 [17].

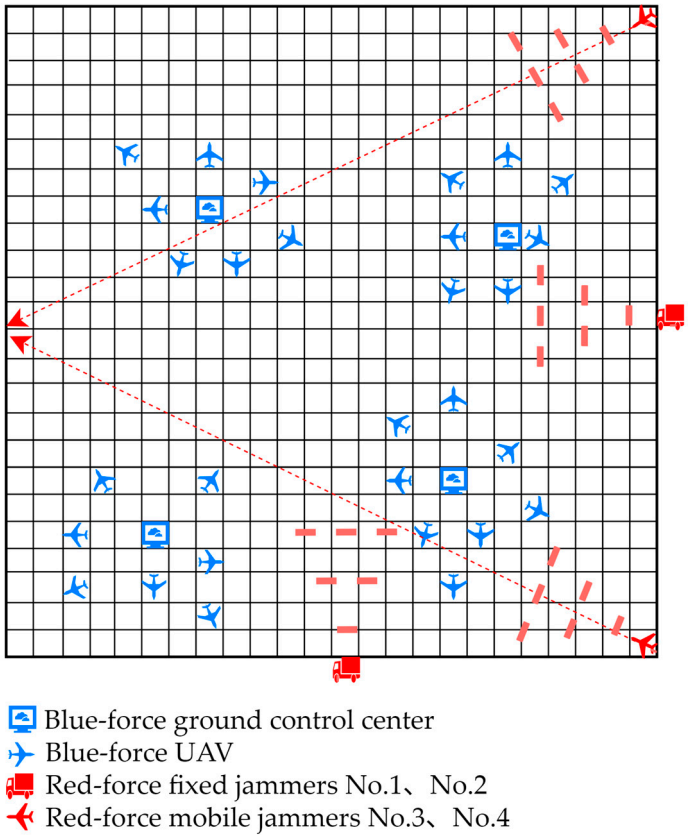


Figure 2. Red–blue equipment distribution.

4.2. Analysis of Algorithm Effectiveness

Figure 3 illustrates the training process of the proposed ACO-DDQN algorithm. In the decoupled strategy, the ACO module performs frequency band allocation, and the

DDQN module handles power level selection. Evidently, the hybrid algorithm can escape poor local solutions during training and converge to a stable solution. Once converged, the algorithm enables the power decisions of each UAV node to adapt slot by slot to time-varying interference and evolving jamming strategies, while the channel assignment is renewed at the beginning of each episode.

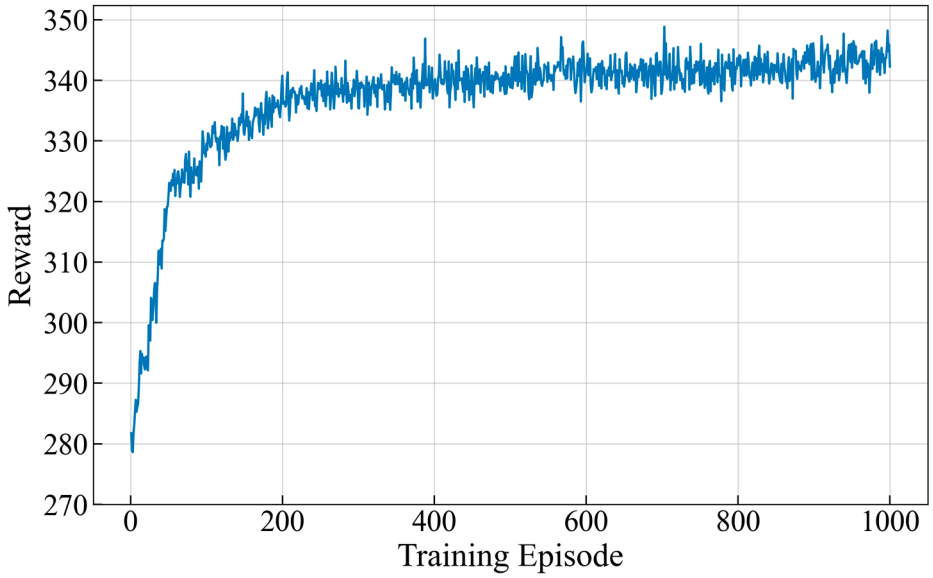
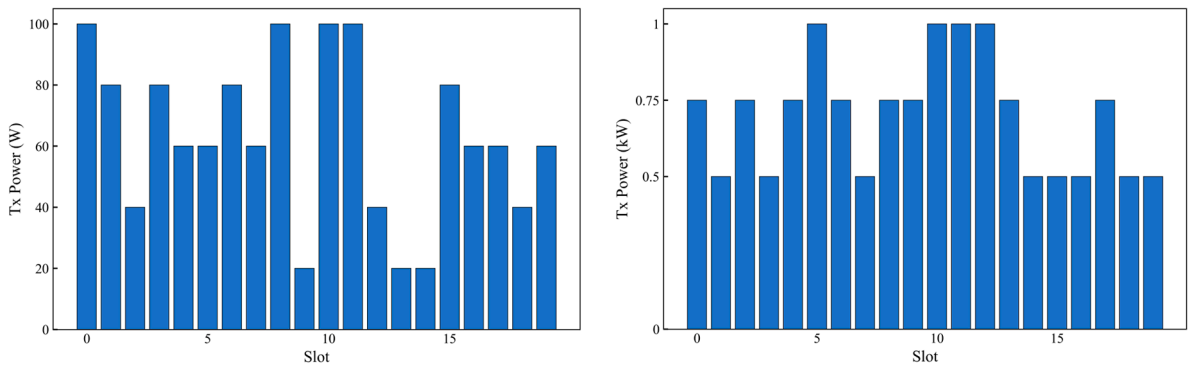


Figure 3. Convergence of the ACO-DDQN algorithm.

Taking UAV Node 1 as a specific example, as the electromagnetic environment changes dynamically, Figure 4 depicts the per-time-slot power level selections for radar reconnaissance and command–communication modules, respectively. Additionally, for a single representative episode, Figure 5 illustrates the frequency band selection of each UAV node for a swarm of 21 UAVs.

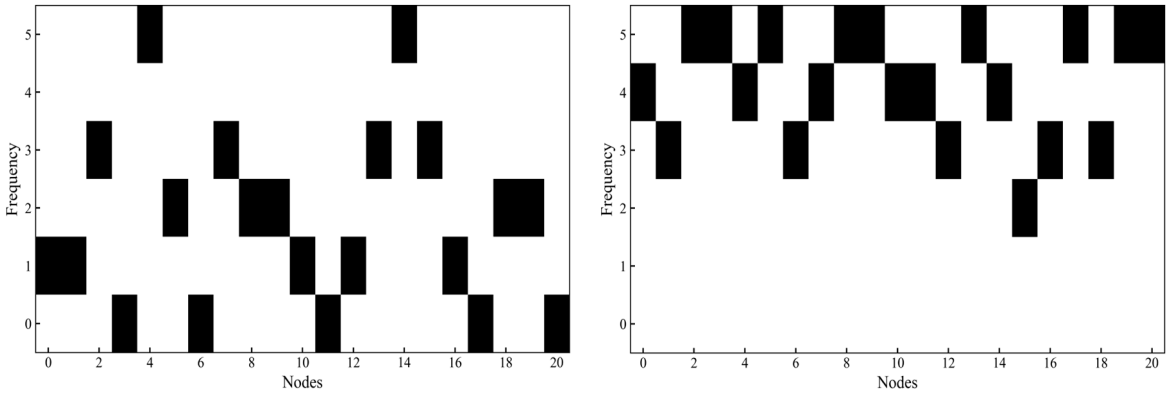


(**a**) (**b**)

Figure 4. (**a**) Per-time-slot command–communication power decision; (**b**) per-time-slot radar power decision.

These results demonstrate that the spectrum decision-making process of the ACO-DDQN algorithm can adapt to variations in the electromagnetic environment on two timescales. Within each episode, it adjusts the power levels of every UAV node slot by slot, and across episodes it renews the channel assignment based on accumulated pheromone feedback. And it can improve spectrum resource utilization efficiency. Overall, the proposed ACO-DDQN algorithm achieves reliable convergence when addressing spectrum

decision-making problems in complex electromagnetic environments, and delivers effective joint optimization of frequency band allocation and power level selection.



(**a**) (**b**)

Figure 5. (**a**) Per-node command–communication frequency decision; (**b**) per-node radar frequency decision.

4.3. Comparison of Algorithm Decision-Making Performance

In the simulation scenario, the number of UAV swarm nodes ranges from 6 to 30. Comparative experiments are conducted across different algorithms. As the number of UAV nodes increases, self-interference and mutual interference among the blue team become increasingly severe, and competition for electromagnetic spectrum resources intensifies. At the same time, the mobile jammers perform dynamic jamming following predefined jamming trajectories. The decision-making performance of the ACO-DDQN algorithm is compared with the DDQN algorithm, the ACO algorithm and the MAPPO algorithm in the aforementioned complex electromagnetic environment. The results are presented in Figure 6 and Table 3. To demonstrate the statistical consistency of the comparison, each configuration is evaluated with three independent runs using different random seeds. The mean values are shown in Figure 6, and the standard deviations and confidence intervals are listed in Table 3.

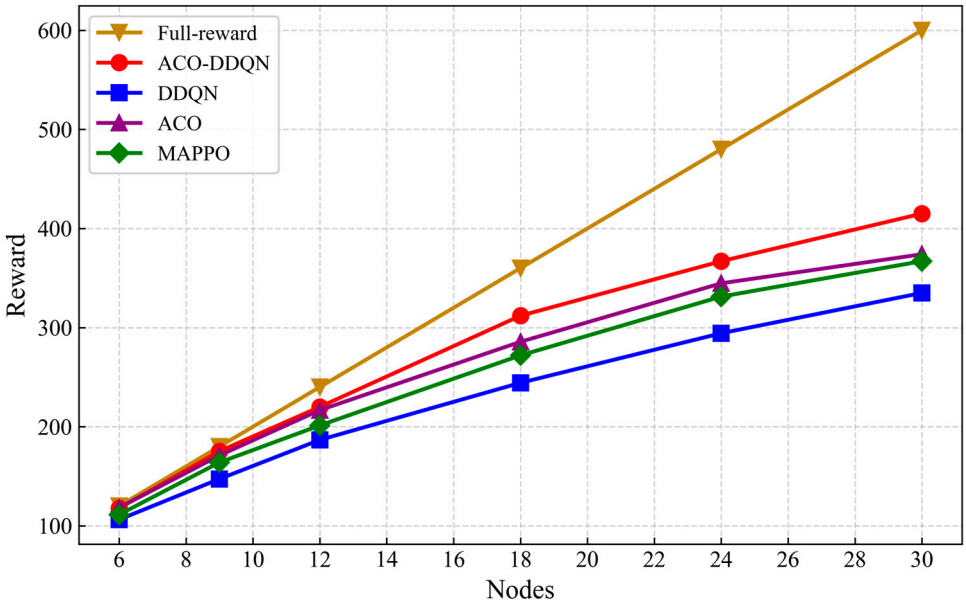


Figure 6. Performance comparison under different quantities of UAV nodes.

Table 3. Rewards of four algorithms at different UAV nodes.

Nodes	ACO-DDQN	DDQN	ACO	MAPPO
6	117.7 ± 1.5 [114.0, 121.4]	106.3 ± 1.5 [102.6, 110.0]	118.3 ± 1.5 [114.6, 122.0]	111.3 ± 0.6 [109.8, 112.8]
9	175.3 ± 3.8 [165.9, 184.7]	147.3 ± 1.5 [143.6, 151.0]	171.0 ± 2.0 [166.0, 176.0]	164.0 ± 1.0 [161.5, 166.5]
12	220.0 ± 3.0 [212.5, 227.5]	186.7 ± 2.1 [181.5, 191.9]	216.7 ± 2.1 [211.5, 221.9]	201.3 ± 1.5 [197.6, 205.0]
18	312.0 ± 3.0 [304.5, 319.5]	244.3 ± 1.5 [240.6, 248.0]	285.7 ± 1.5 [282.0, 289.4]	272.0 ± 1.0 [269.5, 274.5]
24	367.0 ± 3.0 [359.5, 374.5]	294.3 ± 5.0 [281.9, 306.7]	344.7 ± 1.5 [341.0, 348.4]	331.3 ± 6.4 [315.4, 347.2]
30	415.0 ± 3.6 [406.1, 423.9]	335.0 ± 3.6 [326.1, 343.9]	374.0 ± 4.0 [364.1, 383.9]	367.0 ± 4.4 [356.1, 377.9]

Note: Values are mean ± std over three independent runs with different random seeds. Values in brackets are 95% confidence intervals (t-distribution, df = 2). Given the small number of runs, the intervals and tests should be read as indicative.

Table 3 reports the episode rewards of the four algorithms at six swarm scales, with the mean and standard deviation over three independent runs and the corresponding 95% confidence intervals. Two trends are visible directly from the absolute values. When the swarm is small, the proposed ACO-DDQN and the standalone ACO perform almost identically, with a gap below 1% (117.7 ± 1.5 vs. 118.3 ± 1.5 at 6 nodes), and both are clearly ahead of DDQN and MAPPO. Figure 6 shows that, as the number of UAV nodes increases, spectrum competition among the nodes becomes fierce, and the advantage of ACO-DDQN becomes increasingly visible. At 30 nodes it outperforms DDQN by 23.9%, MAPPO by 13.1%, and ACO by 11.0%.

Since the theoretical maximum reward scales linearly with the swarm size, the absolute rewards cannot be compared directly across scales. Table 4 normalizes each result by the theoretical maximum. All four methods degrade as spectrum competition intensifies, but the proposed method degrades the slowest: its normalized reward drops from 98.1% at six nodes to 69.2% at 30 nodes, a loss of 28.9 percentage points, while DDQN, ACO and MAPPO lose 32.8, 36.3 and 31.6 points, respectively. The near-tie with ACO at six nodes (98.1% vs. 98.6%) and the widening gap afterward tell a consistent story: the decoupled architecture pays off exactly where the scenario is hardest, which makes it the more suitable choice for large-scale UAV swarms. This scale dependence has a simple explanation. With six sub-bands and only a handful of UAVs, a good channel assignment alone nearly saturates the achievable reward. Once the swarm outgrows the available channels, co-channel reuse becomes unavoidable, and power control turns into the deciding factor, which is exactly where the DDQN module earns its gain.

Table 4. Normalized rewards.

Nodes	ACO-DDQN	DDQN	ACO	MAPPO
6	98.1%	88.6%	98.6%	92.8%
9	97.4%	81.8%	95.0%	91.1%
12	91.7%	77.8%	90.3%	83.9%
18	86.7%	67.9%	79.4%	75.6%
24	76.5%	61.3%	71.8%	69.0%
30	69.2%	55.8%	62.3%	61.2%

Note: Values are the mean rewards from Table 3 normalized by the theoretical maximum (20 per UAV per episode), expressed as a percentage.

Welch's t-tests on the three independent runs support this reading. The improvements over DDQN and MAPPO are statistically significant at every scale ($p < 0.05$). The improvement over ACO is not significant at 6 to 12 nodes ($p > 0.05$), which matches the near-tie seen in Tables 3 and 4, and becomes significant from 18 nodes onward ($p < 0.01$). We note that with only three runs, these tests have limited statistical power, and we therefore report them as supporting evidence. Our conclusions rest primarily on the large effect sizes (Cohen's d above 10 at 30 nodes) and on the consistent ordering of the four algorithms from 18 nodes onward.

The three baselines were chosen to represent three different paradigms: heuristic search (ACO), deep reinforcement learning over the joint action space (DDQN), and multi-agent reinforcement learning (MAPPO). Since ACO and DDQN are the parent components of the proposed architecture, they also serve an ablation role. Each of them extends one paradigm to cover both decision dimensions, so the performance gap directly reflects the contribution of the replaced module.

To examine whether the proposed method sacrifices one task for the other, Table 5 decomposes the reward into its communication and radar components. At six nodes both efficiencies are close to one (0.9999 and 0.9747), which indicates near-perfect decisions when spectrum resources are abundant. As the UAV nodes grow, the two efficiencies degrade in a balanced manner, and neither component collapses. At 30 nodes the two components converge to almost the same value (0.695 vs. 0.686), which shows that the decoupled architecture keeps the two mission objectives balanced even under heavy spectrum congestion.

Table 5. Command–communication and radar performance.

Nodes	Command–Communication	Radar
6	0.9999 ± 0.0011	0.9747 ± 0.0089
9	0.9952 ± 0.0016	0.9618 ± 0.0050
12	0.8994 ± 0.0022	0.9097 ± 0.0050
18	0.8548 ± 0.0020	0.8636 ± 0.0040
24	0.7281 ± 0.0131	0.8019 ± 0.0227
30	0.6947 ± 0.0011	0.6861 ± 0.0013

Note: Values are per-UAV per-slot averages.

The jammer-power study is conducted at 21 nodes, a representative mid-to-large scale within the evaluated range, where spectrum competition is already significant while the computational cost of the sweep remains moderate. Figure 7 shows the episode rewards of the four algorithms at 21 nodes as the jammer power increases from 0 to 200 W. All methods degrade gracefully, and the proposed ACO-DDQN retains the highest reward at every jammer power, dropping by only 4.3% (from 345 to 330) over the full range. At 200 W, it still outperforms ACO by 9.3%, DDQN by 27.9%, and MAPPO by 13.0%, confirming stable decision performance under intensifying jamming.

4.4. Computational Cost and Communication Overhead

The single reward evaluation costs $O(N^2)$, since the $SINR$ and radar range of each UAV sum over all co-channel interferers. The proposed method pays this price 50 times per episode for the ACO channel assignment, while the per-slot DDQN power selection involves only one forward pass per agent on a compact network. By contrast, the standalone ACO re-searches the joint channel-power space at every slot, i.e., 400 $O(N^2)$ evaluations per episode. We conducted the tests on a laptop CPU (Intel i7-10750H, PyTorch 2.4.1). At 30 nodes, the proposed ACO-DDQN method decides in 82.9 ms per slot, about five times faster than the standalone ACO (409.9 ms), and far below any realistic slot duration. Training is performed offline; it takes 5.1 s per episode at 30 nodes (about 1.4 h for 1000 episodes), comparable to DDQN (1.3 h) and well below ACO (2.4 h). MAPPO has the lowest training

cost (0.75 h), owing to its parameter-shared actor. But it converges to markedly lower rewards (Section 4.3). The proposed method remains the better choice when performance and cost are weighed together. As for communication overhead, a coordination node handles the global observation, gathering it from all UAVs and broadcasting it back. The raw observation has $10N + 18$ scalars; after one-hot channel encoding, each agent receives 678 values per slot. At 30 nodes, this works out to roughly 1.6 MB per episode, and it scales as $O(N^2)$. The ACO channel assignment, by contrast, stays fixed for an entire episode; only $2N$ integers need to be broadcast at the start. Agents do not negotiate with each other during online operation.

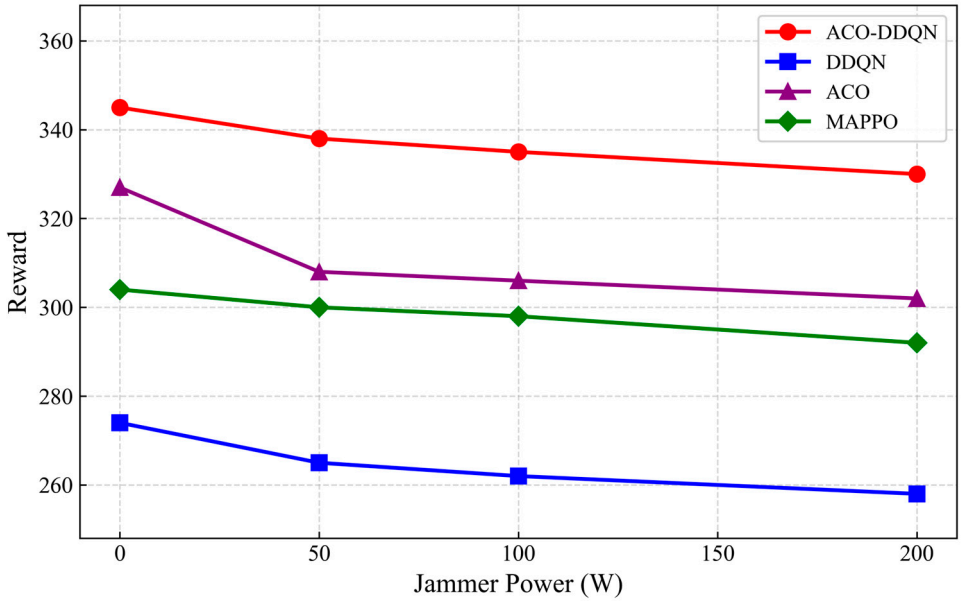


Figure 7. Performance comparison under different jamming power.

The main practical limitation of the current design is the global observation, which assumes a reliable broadcast channel from a coordination node. Its per-episode exchange grows as $O(N^2)$: about 1.6 MB at 30 nodes, and already roughly 18 MB at 100 nodes. This is easily affordable at the scales tested here, but for much larger swarms it becomes the binding constraint, in terms of both bandwidth and the single point of failure introduced by the coordination node. Moving to local observations with communication-efficient aggregation is therefore the natural direction forward. We also keep all UAVs hovering at fixed positions, a deliberate choice that isolates the spectrum decision problem from mobility effects. Adding mobility would introduce time-varying topology and path loss, which we leave to future work. On the evaluation side, the current experiments run from 6 to 30 nodes, and training stays cheap throughout this range (0.74 to 5.1 s per episode). In a reduced three-node scenario, exhaustive search shows that the proposed method attains 97.64% of the exact optimum. The real obstacle to going larger is not computation but validation, since exact optimal references are only computable in reduced scenarios. Scaling to 50 or more UAVs with distributed observations and mobility remains open for future work.

5. Conclusions

This paper proposes a hybrid ACO-DDQN spectrum decision-making algorithm to address the spectrum conflicts and resource competition for UAV swarms in complex electromagnetic environments. The proposed algorithm mainly integrates the heuristic ACO algorithm with the DDQN reinforcement learning algorithm. The ACO component allocates spectrum channels, while the DDQN component implements adaptive power se-

lection. This cooperative framework effectively resolves spectrum conflicts, enables precise resource scheduling, and enhances spectrum decision-making efficiency. Additionally, comparative simulation experiments demonstrate that the proposed ACO-DDQN algorithm achieves stable decision performance and better anti-jamming capability than the baseline algorithms within the evaluated scenarios. This advantage is scale-dependent. At small swarm sizes, the proposed method performs comparably to standalone ACO. At medium and large scales, where spectrum competition is at its most intense, the gain over baseline algorithms becomes substantial. These conclusions are drawn from simulation experiments where UAVs hover at fixed positions and swarm sizes range from 6 to 30 nodes. In our current evaluation, we adopted a simplified channel setup (six sub-bands). Future research will extend the evaluation to more diverse spectrum and traffic settings, and will include a systematic sensitivity analysis of algorithm parameters, including the network size and the power discretization levels. Future work will also explore local observation schemes with communication-efficient aggregation, and extend evaluations to swarms of 50 or more UAVs with distributed observation capabilities and mobility.

Author Contributions: Conceptualization, Y.C.; methodology, Y.C. and Y.W.; software, Y.C. and Y.W.; validation, Y.C., G.W. and Y.W.; formal analysis, Y.C., G.W. and J.Z.; investigation, W.W.; resources, W.W.; data curation, Y.C. and W.W.; writing—original draft preparation, Y.C. and W.W.; writing—review and editing, Y.C., G.W., J.Z. and W.W.; visualization, Y.W. and J.Z.; supervision, G.W.; project administration, W.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data that support the findings of this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hou, J.; Zhang, Q.; Zhang, P.; Lu, W.; Liu, X.; Fang, G. Deinterleaving Mixed Radar Pulse Sequences in Complex Electromagnetic Environments. *IEEE Trans. Aerosp. Electron. Syst.* **2025**, *61*, 15966–15979. [CrossRef]
2. Chen, T.; Sun, G.; Xiao, Y.; Xin, Q. Open Set Target Recognition via ALGGAT with Incomplete Electromagnetic Spectrum Data. *IEEE Sens. J.* **2025**, *25*, 36554–36568. [CrossRef]
3. Posherstnik, Y.; Rocksvald, E.R.; Lussier, B.; Makowski, M. Framework for Interoperable Command & Control of Joint Electromagnetic Spectrum Operations. In *IEEE Military Communications Conference*; IEEE: New York, NY, USA, 2019.
4. Brown, C.; Rutagemwa, H. Revolutionizing Spectrum Management: Data-Driven Strategies. *IEEE Wirel. Commun.* **2025**, *32*, 191–199. [CrossRef]
5. Hou, J.; Hou, Z.; Chen, X.; Wang, H.; Tao, S.; Hu, N. Blockchain and Edge Computing Empowered Battlefield Spectrum Management Architecture. *IEEE Trans. Veh. Technol.* **2025**, *74*, 16335–16348. [CrossRef]
6. Zhang, H.; Zhou, F.; Wu, Q.; Yuen, C. Spectrum Cognition: Semantic Situation for Next-Generation Spectrum Management. *IEEE Netw.* **2026**, *40*, 190–198. [CrossRef]
7. Zhu, X.; Huang, Y.; Wu, Q.; Zhou, F.; Ge, X.; Liu, Y. Dynamic Channel Selection and Transmission Scheduling for Cognitive Radio Networks. *IEEE Internet Things J.* **2022**, *9*, 24429–24443. [CrossRef]
8. Fiengo, R.; Adams, A. Test and evaluation of radar systems operating in the modern electromagnetic spectrum. *IET Radar Sonar Navig.* **2024**, *18*, 2189–2198. [CrossRef]
9. Jia, Z.; You, J.; Dong, C.; Wu, Q.; Zhou, F.; Niyato, D.; Han, Z. Cooperative Cognitive Dynamic System in UAV Swarms: Reconfigurable Mechanism and Framework. *IEEE Veh. Technol. Mag.* **2024**, *19*, 90–101. [CrossRef]
10. Surekha, S.; Rahman, M.Z.U. Spectrum Sensing and Allocation Strategy for IoT Devices Using Continuous-Time Markov Chain-Based Game Theory Model. *IEEE Sens. Lett.* **2022**, *6*, 1–4. [CrossRef]
11. Qi, N.; Huang, Z.; Zhou, F.; Shi, Q.; Wu, Q.; Xiao, M. A Task-driven Sequential Overlapping Coalition Formation Game for Resource Allocation in Heterogeneous UAV Networks. *IEEE Trans. Mob. Comput.* **2023**, *22*, 4439–4455. [CrossRef]
12. Zhou, H.; Jia, L.; Chu, F.; Qi, N.; Zhang, L. Joint Optimization of Task and Spectrum in Autonomous Swarm: A Dual-Rationality-Guided Partially Overlapping Coalition Formation Game Approach. *IEEE Internet Things J.* **2026**, *13*, 14160–14174. [CrossRef]

13. Badran, E.F.; Bashir, A.A.; Kheirallah, H.N.; Farag, H.H. Dynamic Spectrum Co-Access in Multicarrier-Based Cognitive Radio Using Graph Theory Through Practical Channel. *Appl. Sci.* **2024**, *14*, 10868. [CrossRef]
14. Si, J.; Huang, R.; Li, Z.; Hu, H.; Jin, Y.; Cheng, J.; Al-Dhahir, N. When Spectrum Sharing in Cognitive Networks Meets Deep Reinforcement Learning: Architecture, Fundamentals and Challenges. *IEEE Netw.* **2024**, *38*, 187–195. [CrossRef]
15. Liao, X.; Wang, Y.; Han, Y.; Li, Y.; Lin, C.; Zhu, X. Heterogeneous Multi-Agent Deep Reinforcement Learning for Cluster-Based Spectrum Sharing in UAV Swarms. *Drones* **2025**, *9*, 377. [CrossRef]
16. Zhao, H.; Duan, G.; Si, J.; Huang, R.; Shi, J. Research on intelligent spectrum allocation techniques for incomplete electromagnetic information. *J. Electron. Inf. Technol.* **2024**, *46*, 2694–2702.
17. Yin, K.; Shi, J.; Duan, G.; Li, L.; Si, J. Greedy-PPO intelligent spectrum sharing decision for complex electromagnetic interference environments. *Acta Aeronaut. Astronaut. Sin.* **2024**, *45*, 233–246.
18. Yuan, X.; Wen, J. A Curriculum-Learning-Assisted MAPPO-Based Algorithm for Dynamic Spectrum Access and Anti-Jamming in UAV Swarms. *Sensors* **2026**, *26*, 2912. [CrossRef] [PubMed]
19. Du, J.; Cui, L.; Xu, Y.; Feng, Z.; Zhang, Y.; Li, Y.; Liang, K.; Xu, Y. Optimizing UAV Trajectory and Spectrum Access for Anti-Jamming Communications: A Heterogeneous MADRL Approach. *IEEE Commun. Lett.* **2026**, *30*, 1642–1646. [CrossRef]
20. Lou, D.; Wan, B.; Zhang, Y.; Du, Y.; Wan, F.; Chen, Y. Orthogonal Gradient Multi-Agent Deep Reinforcement Learning-based Algorithm for Spectrum Resource Optimization in UAV Swarm. *Phys. Commun.* **2025**, *72*, 102770. [CrossRef]
21. Sheng, H.; Zhou, W.; Zheng, J.; Zhao, Y.; Ma, W. Transfer Reinforcement Learning for Dynamic Spectrum Environment. *IEEE Trans. Wirel. Commun.* **2024**, *23*, 1447–1458. [CrossRef]
22. Ghaderibaneh, M.; Zhan, C.; Gupta, H. DeepAlloc: Deep Learning Approach to Spectrum Allocation in Shared Spectrum Systems. *IEEE Access* **2024**, *12*, 8432–8448. [CrossRef]
23. Zhang, Z.; Liu, Y.; Wang, C.; Chang, H.; Bian, J.; Zhang, J. Machine Learning Based Clustering and Modeling for 6G UAV-to-Ground Communication Channels. *IEEE T. Veh. Technol.* **2024**, *73*, 14113–14126. [CrossRef]
24. Zhou, S.; Niu, X.; Lu, S.; Xiang, Y.; Pu, C. Self-organized anti-jamming reinforcement learning for resource allocation in UAV-assisted networks. *Signal Process.* **2026**, *240*, 110346. [CrossRef]
25. Wei, W.; Wang, J.; Fang, Z.; Chen, J.; Ren, Y.; Dong, Y. 3U: Joint design of UAV-USV-UUV networks for cooperative target hunting. *IEEE Trans. Veh. Technol.* **2023**, *72*, 4085–4090. [CrossRef]
26. Rahman, M.T.; Shah, A.F.M.S.; Karabulut, M.A.; Ilhan, H. FANET-enabled cluster-based emergency communication with 3D mobility in 5G and beyond. *Veh. Commun.* **2025**, *56*, 100971. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.