Optimal Resource Allocation for Successful Task Transmission Probability Maximization in UAV Swarm Networks

Zhuojia Yang, Wei Su, Bin Yang, Yudong Ma, Yihua Peng, Qi Liu, Tarik Taleb

***Abstract*—Unmanned aerial vehicle (UAV) swarm networks (USNTs) are a crucial component of the emerging low-altitude intelligent networks. For supporting low-altitude economic activities, this paper explores successful task transmission probability (STP) maximization, while maintaining the performance fairness of each UAV swarm for USNTs with resource limitations (e.g., frequency, power). Towards this goal, this paper formulates STP maximization and its fairness as a nonlinear non-convex optimization problem. To solve the complex optimization problem, we first propose an adaptive frequency block sharing algorithm to determine whether different-sized UAV swarms use the same frequency blocks or not, providing an efficient initial solution for inter-swarm resource allocation. Then we develop a double deep Q-network-based algorithm for inter-swarm resource allocation to guarantee fairness among swarms. Based on the inter-swarm allocation results, we propose a genetic algorithm-based method for intra-swarm resource allocation to achieve the STP maximization, which ensures reliable transmission of high-priority tasks. Extensive simulation results are presented to validate the efficiency of our proposed algorithms, and also to illustrate the impact of system parameters on STP and fairness.**

***Index Terms*—UAV networks, resource allocation, successful task transmission probability, fairness, deep reinforcement learning.**

I. Introduction

Unmanned aerial vehicle (UAV) swarm networks (USNTs) are a key component of low-altitude intelligent networks (LAINs), which have many distinctive features like flexible deployment, high mobility and low cost [1]–[3]. In particular, UAV swarms can cooperate with each other to efficiently perform complex tasks for supporting various applications of Internet of Things (IoT) such as smart traffic, environmental monitoring, emergency rescue, etc. USNTs have received extensive attention from both industrial and academic communities [4]. For supporting various IoT applications, a fundamental issue is how to optimize resource (e.g., frequency, power) allocations to ensure reliability and fairness of task transmissions in USNTs with resource limitation.

Studies on resource allocation in UAV networks mainly focus on scenarios without UAV swarms [5]–[14] and UAV swarm scenarios [15]–[36]. For the scenarios without UAV swarms, UAVs usually serve as aerial users or base stations (BS). These works optimize the UAV deployment, UAV trajectory and allocation of spectrum and power resources to improve system performance in terms of throughput, energy consumption and efficiency, delay and cost using traditional optimization methods [5], [6], [8]–[12] and machine learning (ML) methods [7], [13], [14]. For the UAV swarm scenarios, the works focus on single-swarm [15]–[26] and multi-swarm scenarios [27]–[36], improving throughput [15]–[18], [36], energy efficiency [19]–[21], resilience [33], [34], fairness [31], system performance [25]–[28], [30], [32], and minimizing delay [22]–[24], [29] through heuristic methods, deep reinforcement learning (DRL) and multi-agent DRL (MADRL).

It is notable that the works on UAV swarms illustrate that system performance can be improved through proper resource allocations. We compare existing UAV swarm works on spectrum sharing, heterogeneity, fairness and priority, as shown in Table I. However, several critical challenges have not been well addressed so far. First, in the scenarios with limited resources and a large number of UAVs, the severe interference among UAVs will significantly degrade system performance. For example, in emergency scenarios, due to limited resources, spectrum sharing is required, which may lead to interference as multiple UAVs perform tasks such as search in the same area. Efficient resource allocation and sharing strategy mitigates interference, ensuring successful rescue task performance. Second, the inherent heterogeneity of UAV swarms further constrains the adaptability of conventional single-network resource allocation schemes within multi-swarm network environments. Swarms performing diverse rescue tasks (e.g. sensing, communication relays) exhibit variations in the number and distribution of UAVs, making it difficult to achieve optimal network performance applying the conventional resource allocation method across multiple heterogeneous UAV swarms. Third, the lack of explicit consideration of task transmission reliability and network fairness can lead to unstable service quality and degraded overall system performance. Transmission reliability ensures that critical data is delivered without loss or delay, while fairness guarantees all UAVs can perform their tasks effectively. Additionally, in emergency scenarios, tasks such as search and rescue, medical aid, and disaster relief operations hold varying levels of importance. Therefore, the task priority should be taken into account to ensure the efficient mission execution.

Therefore, this paper investigates a resource allocation method for USNTs that effectively addresses these challenges. In particular, UAV swarms are configured to cooperatively share spectrum resources while reducing mutual interference

Zhuojia Yang, Wei Su and Yihua Peng are with the School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing 100044, China (e-mail: 23111050@bjtu.edu.cn; wsu@bjtu.edu.cn; 22110024@bjtu.edu.cn).

Bin Yang is with the School of Computer and Information Engineering, Chuzhou University, Chuzhou 239000, Anhui, China (e-mail: yangbinchi@gmail.com).

Yudong Ma is with the Institute of Artificial Intelligence, Beihang University, Beijing 100191, China. (e-mail: yudongma1998@buaa.edu.cn).

Qi Liu is with the Technology Innovation Center, Smart City Research Institute of China Unicom, Beijing 100048, China (e-mail: liuqi49@chinaunicom.cn).

Tarik Taleb is with the Faculty of Electrical Engineering and Information Technology, Ruhr University Bochum, 44801 Bochum, Germany (e-mail: tarik.taleb@rub.de).

TABLE I
COMPARISON WITH EXISTING WORKS.

Work	Multi-swarm	Inter-swarm spectrum sharing	Hetero-geneity	Fairness	Priority	Optimization Objective	Solution Method
[15], [18], [24], [26]				✓	✓	Throughput, task loss rate, delay, energy consumption and fairness	MADRL and traditional methods
[16], [17], [22], [23], [25]					✓	Throughput, delay and long-term rewards	Reinforcement learning and traditional methods
[19]–[21]				✓		Energy consumption	MADRL and traditional methods
[27]	✓	✓	✓		✓	Age of information and consumption	MADRL
[28]	✓	✓				Performance and security of communication	Multi-object optimization
[29], [34]	✓					Task execution time [29] and resilience [34]	MADRL [29], heuristic method [34]
[30]	✓	✓		✓	✓	Data aggregation and offloading rate	Multi-agent reinforcement learning (MARL)
[31], [33], [35]	✓				✓	Trade-off for data packets [31], resilience [33] and cost [35]	Traditional method [31], heuristic method [33], MADRL [35]
[32]	✓			✓		Coverage rate and time	DRL
[36]	✓			✓	✓	Transmission success rate [36]	Traditional method [36]
This work	✓	✓	✓	✓	✓	Successful task transmission probability (STP)	DRL and heuristic methods

as much as possible, thereby mitigating the challenges of limited resources and severe interference. Then, inter-swarm and intra-swarm resources are allocated according to swarm size, structure, and task priority, taking heterogeneity and differentiated task requirements into account to enhance the adaptability of the resource allocation. Moreover, the successful task transmission probability (STP) and inter-swarm utility variance are introduced to evaluate task reliability and network fairness, respectively, and are jointly optimized to enhance overall system performance. The main contribution of this paper is summarized as follows.

- We consider a USNT consisting of multiple UAV swarms. We formulate the intra-swarm and inter-swarm resource allocations as two nonlinear nonconvex optimization problems. To solve these problems, we first propose an adaptive frequency block sharing (AFBS) algorithm, which provides a simple and effective mechanism to determine whether different-sized UAV swarms use the same frequency blocks (FBs) or not in the resource-limited USNT, and also serves as an initialization step for the subsequent optimization algorithms. Through mathematical analysis, we establish the necessary conditions to achieve interference-free operation for shared FBs. We further design an optimized graph coloring algorithm that accounts for the heterogeneity of UAV swarms to facilitate inter-swarm frequency sharing, thereby improving the STP under limited network resources.
- We then develop a frequency resource allocation method based on double deep Q-network (DDQN) for the inter-swarm resource allocation problem, aiming to achieve dynamic decision-making in inter-swarm resource allocation. Specifically, we design a reward mechanism based on the differences in swarm utility functions to guide the learning process, enabling the control center to perform centralized resource allocation based on the utility function results reported by swarm head nodes. For the discrete and highly combinatorial intra-swarm resource allocation problem, we also propose a power control and sub-channel selection method based on a genetic algorithm (GA), maximizing the STP of high-priority tasks by modeling a utility function that takes task priority into account.
- Finally, simulation experiments are conducted to evaluate the proposed algorithms. The experimental results demonstrate that the overall STP under the proposed algorithms outperforms that under the baseline algorithm, ensuring reliable transmission for high-priority tasks, and achieving superior fairness in inter-swarm resource allocation.

The rest of the paper is organized as follows. Section II provides an overview of the related work. In Section III, the system model is presented, followed by the problem formulation in Section IV. Section V details our proposed resource allocation method for USNTs. The numerical results are discussed in Section VI. Finally, Section VII concludes this paper, highlighting the key findings.

II. RELATED WORK

A. Scenarios without UAV Swarms

Several studies focus on improving system throughput [5]–[7], energy efficiency [8]–[10], minimizing energy consumption [11], [12], time delay [13] and cost [14] by optimizing UAV deployment [5], [11], [12], [14], power [6], trajectory, spectrum [7], [8], scheduling [9], [10] and offloading [13]. In

[5]–[7], the authors jointly optimized UAV deployment [5], its location, time, quantity, position [6], trajectory, power and spectrum allocation [7] to maximize system throughput, and the DRL was introduced in [7] to enhance the adaptability. To minimize energy consumption, [11] and [12] optimized UAV altitude, speed and other parameters using heuristic methods. Similarly, [8]–[10] optimized resource allocation, trajectory [8], scheduling [10], velocity [9] using traditional methods. Additionally, [12] minimized the time delay by optimizing deployment, resource allocation and offloading using DRL, and [14] minimized cost by optimizing computational cost and UAV deployment.

In summary, these studies can improve overall system performance and provide useful insights for UAV swarms. However, they are insufficient to cope with the challenges of cooperation, complex interference, and intensified resource competition in USNTs, and resource allocation in such networks requires further investigation.

B. Scenarios with UAV Swarms

The studies on UAV swarm scenarios can be classified into single-swarm and multi-swarm. For the single-swarm scenarios, several studies focus on improving system throughput [15]–[18], energy efficiency [19]–[21], reducing delay [22]–[24] and maximizing system total performance [25], [26] by optimizing spectrum, power, UAV deployment and related factors. In [15]–[18], the authors jointly optimized UAV deployment [15], [16], power-beam, spectrum [17] and antenna parameters [18] to improve system throughput, and the fairness was also considered in [18]. For the energy consumption, [20] and [21] minimized energy consumption by optimizing UAV trajectory [20] and spectrum and computing resources [21], and [24] minimized task delay and energy consumption through trajectory, association, and offloading optimization. In [22] and [23], parameters such as power, spectrum, UAV trajectory [23], offloading decisions, and computation resources [22] were jointly optimized to minimize delay. In addition, depending on the definition of the system utility function, spectrum, user association, power [25], and trajectory [26] were jointly optimized to maximize long-term utility [25], [37], and minimize the task loss rate [26].

In multi-swarm scenarios, current studies focus on resource allocation [27]–[31], deployment [32], resilience and recovery [33], [34], as well as traffic scheduling [35], [36]. For resource allocation, [27] and [28] aimed to minimize the age of information and power consumption [27], and improve the communication performance and security [28] by optimizing power and spectrum allocation, using multi-agent deep reinforcement learning (MADRL) and multi-object optimization, respectively. [29] optimized scheduling and communication, sensing and computing resource allocation to reduce execution time. [30] considered large-scale UAV swarm interference, optimized resource and trajectory scheduling, and integrated non-orthogonal multiple access and MARL to maximize network data offloading rate. [31] jointly optimized user association and spectrum allocation to balance the performance of different data packets. In [32], the authors focus on the deployment of UAV swarms to maximize coverage rate and minimize coverage time using DRL. Additionally, the resilience of the swarms is discussed in [33] and [34], where network topology optimization [33] and unfinished task allocation [34] methods are proposed to enhance resilience and minimize profit loss during attacks. [35] optimized the virtual swarm membership and traffic flow paths to reduce the communication overhead and cost, and [36] improved the success rate of data packet transmission by optimizing routing.

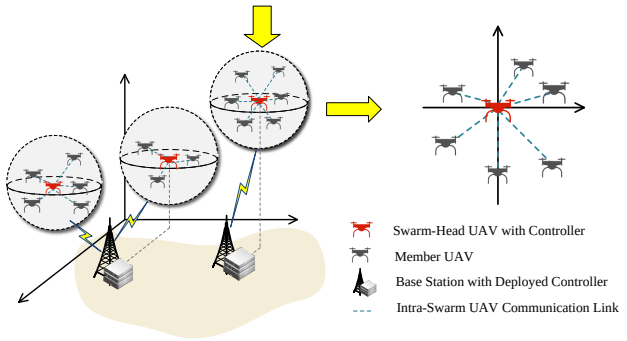


Fig. 1. A USNT system.

In total, the studies on USNTs have primarily focused on optimizing power, spectrum [27]–[31], scheduling [29], trajectory [30], deployment [32], network recovery [33], [34] and routing [35], [36] to improve network performance in terms of communication [27]–[29], [31], [35], [36], offloading [30], coverage [32], and recovery [33], [34]. However, these studies do not jointly account for spectrum sharing, heterogeneity, fairness, priority, and task reliability, which are crucial for enhancing overall system performance and transmission efficiency. Therefore, it is essential to jointly optimize these aspects.

III. System Model

A. Network Model

As shown in Fig. 1, we consider a USNT consisting of C UAV swarms, where any swarm $c \in \mathcal{C} = \{c_i | i = 1, 2, \cdots, C\}$. Each swarm contains a swarm-head UAV equipped with multiple antennas and M_c member UAVs, each equipped with a single antenna, which are randomly distributed in the vicinity of swarm-head UAV. The swarm-head UAV of swarm c is denoted as CH_c, and the i-th member UAV in swarm c is denoted as CM_c^i, where $CM_c^i \in \mathcal{CM}_c = \{CM_c^1, CM_c^2, \cdots, CM_c^{M_c}\}$, $\mathcal{CM}_c \in \mathcal{CM} = \{\mathcal{CM}_1, \mathcal{CM}_2, \cdots, \mathcal{CM}_C\}$, and $CH_c \in \mathcal{CH} = \{CH_1, CH_2, \cdots, CH_C\}$. Each swarm has a star topology structure, where CM_c^i directly communicates with the CH_c. The main controller is deployed at each BS, with secondary controllers at the CH_c. To achieve load balancing and enhance network scalability, the main controller interacts exclusively with CH_c to manage external swarm status information and oversee overall network resources, while secondary controllers manage resource allocation between CM_c^i and the CH_c within each swarm.

TABLE II
LIST OF MAJOR NOTATIONS

Symbol	Description
$CM_c^i/C\mathcal{M}_c/C\mathcal{M}$	The i-th member UAV of swarm c / the set of the member UAVs in swarm c / the set of $C\mathcal{M}_c$
$CH_c/\mathcal{CH}$	The swarm head UAV of swarm c / the set of the head UAVs
$C/\mathcal{C}$	The number / set of swarms
M_c	The number of member UAVs in swarm c
$K/\mathcal{K}$	The number / set of sub-channels
$K_c/\mathcal{K}_c$	The number / set of sub-channels of swarm c
t/L_s	The index / length of the slot
d_c^i	The distance between CM_c^i and CH_c
G_c^i	The channel gain between CM_c^i and CH_c
$\mathbf{p}_c^i$ / $\mathbf{p}_c$	The coordinates of CM_c^i and CH_c
p_c^i	The transmit power of CM_c^i
ψ_{ci}^k/Ψ_c	The indicator variable for CM_c^i occupying the k-th sub-channel in swarm c / the set of ψ_{ci}^k
r_{ci}^k / r_c^i	The SINR from CM_c^i to CH_c on k-th sub-channel / the SINR from CM_c^i to CH_c
I_{ci}^k	The interference of other UAVs to the signal of CM_c^i on the k-th sub-channel
R_c^i	The transmission rate between CM_c^i and CH_c
$p_{ci}^j/\mathcal{P}_c^i/\mathcal{P}_c$	The CM_c^i transmits at power level P_j / the set of power level selection of CM_c^i / the set of $\mathcal{P}_c^i$
B	The bandwidth of a single sub-channel
$\mathcal{C}^q$	The set of swarms using FB $\mathcal{F}_q$
λ_c^m	The indicator variable of transmitting the m-th task successfully in swarm c
$b_c^m/\mathcal{B}_c$	The m-th task / the set of tasks in swarm c
$\pi_m/\sigma_m/\delta_m$	The priority / data size / delay requirement of m-th task in swarm c
τ_m	The actual transmission delay for the b_c^m
Y_c^π	The ratio of successfully transmitted tasks to total tasks with priority π in swarm c.
$\mathcal{F}_q/f_q/\mathcal{F}$	The q-th FB / the number of sub-channels of the q-th FB / the set of FBs
μ_c^q	The usage indicator variable of the q-th FB by swarm c

Assuming the total number of sub-channels in the network is K, each with bandwidth B, the sub-channels set is denoted by $\mathcal{K} = \{k_i | i = 1, 2, ..., K\}$, and the sub-channel set of swarm c is denoted by $\mathcal{K}_c \subset \mathcal{K}$. It is notable that the non-orthogonal access is employed for communication between the member UAVs and head UAV within a swarm, where each UAV is allowed to occupy at most one sub-channel, whereas each sub-channel may be simultaneously accessed by multiple UAVs. Time is divided into slots of length L_s, with $t \in \mathcal{T} = \{1, 2, ..., T\}$ as the time slot index. In our system model, the CM_c^i transmits the collected data to the CH_c via uplink, and then the CH_c forwards the aggregated tasks to the BS for processing. Such a system can be widely applied in various fields, including disaster rescue, battlefield operations, and scenarios with limited infrastructure. We focus on resource allocation among swarms and member UAVs during the uplink transmission phase. The main symbols used in this paper are listed in Table II.

Remark 1: The single-point failure is not considered in this paper. This is because in the scenarios discussed, such as short-term environmental monitoring, emergency rescue and small-scale inspection, UAV energy and operational duration are considered, allowing prompt updating and replacement of the swarm head before failure to maintain network reliability. Additionally, for the single point failure of the swarm head structure, several reselection methods can be employed to reselect a new head point upon its departure or failure to ensure reliable task transmission, such as a weighted swarm head selection method in [38].

B. Intra-Swarm Communication Model

As discussed in [39], the line of sight (LoS) model provides a good approximation for actual UAV communications. In the LoS model, the path loss between CM_c^i and CH_c relies on the distance and propagation conditions among UAVs. Specifically, under the LoS model, the channel gain between UAVs follows the free-space path loss model, and the path loss between the CM_c^i and CH_c in swarm c is

$$PL(\text{dB}) = 10 \log \frac{p_c^i(t)}{p_c(t)} = -10 \log \left[\frac{G_t G_r \lambda^2}{(4\pi)^2 {d_c^i}^2} \right], \quad (1)$$

where $p_c^i(t)$ and $p_c(t)$ represent the transmit power of CM_c^i and received power of CH_c at time t respectively. G_t and G_r represent the transmit antenna gain and receive antenna gain, respectively, while d_c^i denotes the distance between CM_c^i and CH_c. When the transmission frequency and antenna gains are fixed, the channel gain between CM_c^i and CH_c is primarily determined by the distance, expressed as

$$G_c^i = \beta_0 \cdot (d_c^i)^{-\alpha} = \frac{\beta_0}{(\|\mathbf{p}_c^i - \mathbf{p}_c\|_2)^\alpha}, \quad (2)$$

where β_0 represents the channel power gain at the reference distance $d_0 = 1$m, $\alpha \geq 2$ is the path loss exponent, $\mathbf{p}_c^i$ and $\mathbf{p}_c$ are the coordinates of CM_c^i and CH_c respectively, and $\|*\|_2$ is the L^2 norm of $*$.

In intra-swarm UAV communication, the interference for each CM_c^i-CH_c pair is caused by the operations of other UAV pairs sharing the same sub-channels. Let $\psi_{ci}^k(t)$ denote the indicator variable representing the occupancy of k-th sub-channel by CM_c^i in swarm c, where $\psi_{ci}^k(t) = 1$ if k-th sub-channel is occupied by CM_c^i at time slot t; $\psi_{ci}^k(t) = 0$, otherwise. We allow each UAV to occupy only one sub-channel, which is

$$\sum_{k \in \mathcal{K}_c} \psi_{ci}^k(t) \leq 1, \quad (3)$$

$$\Psi_c = \{\psi_{ci}^k(t)\}, k \in \mathcal{K}_c, c \in \mathcal{C}. \quad (4)$$

Accordingly, the signal-to-interference-plus-noise ratio (SINR) for the CM_c^i-CH_c pair on k-th sub-channel can be expressed as

$$r_{ci}^k(t) = \frac{G_c^i \cdot \psi_{ci}^k(t) \cdot p_c^i(t)}{I_{ci}^k(t) + \sigma^2}, \quad (5)$$

where $\sigma^2 = N_0 B$ is the noise power. $I_{ci}^k(t)$ is the interference power caused by other UAVs on k-th sub-channel to the signal of CM_c^i, which is given by

$$I_{ci}^k(t) = \sum_{l=1, l \neq i}^{M_c} G_c^l \cdot \psi_{cl}^k(t) \cdot p_c^l(t). \quad (6)$$

Finally, the total SINR from CM_c^i to CH_c can be calculated by

$$r_c^i(t) = \sum_{k=1}^{K_c} \frac{G_c^i \cdot \psi_{ci}^k(t) \cdot p_c^i(t)}{I_{ci}^k(t) + \sigma^2}. \tag{7}$$

Then we have that the transmission rate between CM_c^i and CH_c is

$$R_c^i(t) = B \log_2(1 + r_c^i(t)). \tag{8}$$

To ensure that the signal can be correctly demodulated at the receiver, the received SINR should be no less than a threshold SINR_{th}, i.e.,

$$r_c^i(t) \geq \text{SINR}_{th}. \tag{9}$$

In this paper, we adopt discrete transmit power control. The power level is denoted as $\{P_1, P_2, \cdots, P_J\}$. For each CM_c^i, we define a binary variable $p_{ci}^j(t)$ to represent the power level, where $j \in \mathcal{J} = \{1, 2, ..., J\}$. If CM_c^i selects power level P_j for transmission at time slot t, then $p_{ci}^j(t) = 1$; otherwise, $p_{ci}^j(t) = 0$. Each UAV selects one power level at time t, i.e.,

$$\sum_{j \in \mathcal{J}} p_{ci}^j(t) \leq 1, \forall i \in \{1, 2, \cdots, M_c\}. \tag{10}$$

For the power selection of all member UAVs within swarm c, we obtain

$$\mathcal{P}_c = \{\mathcal{P}_c^1, \mathcal{P}_c^2, \cdots, \mathcal{P}_c^{M_c}\}, \tag{11}$$

where $\mathcal{P}_c^i = \{p_{ci}^1(t), p_{ci}^2(t), \cdots, p_{ci}^J(t)\} \in \mathcal{P}_c$.

As a result, the transmit power of CM_c^i can be expressed as

$$p_c^i(t) = \sum_{j \in \mathcal{J}} \left[p_{ci}^j(t) \cdot P_j \right]. \tag{12}$$

C. Task Model

The transmission tasks within swarm c at time slot t are represented by

$$\mathcal{B}_c = \{b_c^1, b_c^2, ..., b_c^{M_c}\}, c \in \mathcal{C}, \tag{13}$$

where $b_c^m = \{\pi_m, \sigma_m, \delta_m\}$, $m \in \{1, 2, \cdots, M_c\}$ denotes the m-th task in swarm c with priority π_m, data size σ_m and delay requirement δ_m, and $\Pi = \{\pi_1, \pi_2, \cdots, \pi_{|\Pi|}\}$.

Based on the task size and transmission rate, the actual transmission delay for the b_c^m is calculated as

$$\tau_c^m = \sigma_c^m / R_c^m. \tag{14}$$

R_c^m is the transmission rate at current time slot. Let λ_c^m indicate whether a task is successfully transmitted, which is expressed as

$$\lambda_c^m = \begin{cases} 1, & \tau_c^m \leq \delta_m \\ 0, & \tau_c^m > \delta_m. \end{cases} \tag{15}$$

Then, Y_c^π is defined as the ratio of the number of successfully transmitted tasks with priority π to the total number of tasks with priority π in swarm c.

$$Y_c^\pi = \begin{cases} \frac{\sum_{\pi_m} \lambda_c^m}{|\{b_c^m\}_{\pi_m = \pi|}}, & |\{b_c^m\}_{\pi_m = \pi}| \neq 0 \\ 0, & \text{others.} \end{cases} \tag{16}$$

We adopt the following scheduling strategy for UAV transmissions: each UAV is allowed to transmit only one task per time slot, with the transmission starting at the beginning of the slot. The task must be completed within the same time slot, after which the inter-UAV resources are reallocated according to the proposed method. This scheduling method provides the foundation for designing an online learning algorithm aimed at optimizing the long-term STP of the multi-UAV network. By leveraging real-time feedback, the algorithm can adapt to dynamic network conditions and improve overall performance across multiple time slots.

IV. Problem Formulation

In this section, we first present an intra-swarm optimization problem and an inter-swarm optimization problem, respectively. Then, we formulate the intra-swarm and inter-swarm resource allocations as an optimization problem in USNT.

A. Intra-Swarm Resource Allocation

Within the swarm, the utility function serves to provide optimization guidance for the resource allocation and decision-making of each UAV. The goal of intra-swarm optimization is to maximize the STP of high-priority tasks, ensuring network reliability. Therefore, the intra-swarm utility function, which is positively correlated with STP, is expressed as:

$$U_c = \sum_{m=1}^{|\Pi|} w_m \cdot Y_c^{\pi_m}. \tag{17}$$

Here, the constant w_m represents the weight of the task with priority π_m, and $W = [w_1, w_2, \cdots, w_{|\Pi|}]$. Accounting for the task priorities, the weights satisfy $w_1 > w_2 > \cdots > w_{|\Pi|}$ and $w_m \in (0, 1)$, with $\sum w_m = 1$.

We define the intra-swarm optimization problem as $\boldsymbol{P1}$. The objective is to maximize the STP of the swarm under a given number of sub-channels by selecting the optimal transmit power and sub-channels configuration for member UAVs within the swarm. The problem $\boldsymbol{P1}$ is formulated as follows.

$$\boldsymbol{P1}: \max_{\mathcal{P}_c, \Psi_c} \quad U_c \tag{18}$$

$$\text{s.t.} \quad \sum_{k \in \mathcal{K}_c} \psi_{ci}^k(t) \leq 1, \tag{18a}$$

$$r_c^i(t) \geq \text{SINR}_{th}, \tag{18b}$$

$$0 \leq p_c^i(t) \leq P_{max}, \tag{18c}$$

$$\sum_{j \in \mathcal{J}} p_{ci}^j(t) \leq 1, \tag{18d}$$

$$\psi_{ci}^k(t),\ p_{ci}^j(t),\ \lambda_c^m \in \{0, 1\}, \tag{18e}$$

$$i, m \in \mathcal{M}_c,\ c \in \mathcal{C},\ k \in \mathcal{K}_c,\ j \in \mathcal{J}$$

where (18a) ensures that each UAV occupies at most one sub-channel. (18b) specifies the minimum SINR required for the swarm head UAV to decode received signals indiscriminately.

(18c) defines the power limitation of UAVs. (18d) ensures that each UAV selects only one power level, while (18e) imposes binary restrictions on the decision variables.

B. Inter-Swarm Resource Allocation

In the inter-swarm scenario, all swarms compete for the total network resources. As the scale of UAVs expands, the average resources available to each UAV decrease. To enhance resource efficiency, this paper introduces a cooperative and competitive framework among swarms, with a primary focus on optimizing the allocation of effective spectrum resources to each swarm.

We first divide all spectrum resources of the USNT into non-overlapping FBs, with each FB serving as the spectrum unit for swarm usage. A FB is defined as a set of continuous sub-channels, and $\mathcal{F}_q \subset \mathcal{K}$ represents the q-th FB in the network, which can be expressed by

$$\mathcal{F}_q = \{k_a, k_{a+1}, \cdots, k_{a+f_q-1}\}, 0 \le a \le K, \tag{19}$$

where f_q represents the total number of sub-channels in $\mathcal{F}_q$.

It is notable that FBs are non-overlapping and collectively constitute the spectrum resources of network, which is

$$\mathcal{F}_p \cap \mathcal{F}_q = \varnothing, p \ne q, \forall p, q \in \{1, 2, ..., Q\}, \tag{20}$$

$$\mathcal{F}_1 \cup \mathcal{F}_2 \cup \cdots \cup \mathcal{F}_Q = \mathcal{K}. \tag{21}$$

Here, Q represents the total number of FBs. With this in mind, we use $\mathcal{C}^q$ to denote the set of swarms sharing $\mathcal{F}_q$, where $\mathcal{C}^q \in \{\mathcal{C}^1, \mathcal{C}^2, \cdots, \mathcal{C}^Q\}$. Apparently, for swarms sharing the same FB, their sub-channels satisfy:

$$K_c = f_q, \forall c \in \mathcal{C}^q. \tag{22}$$

Then, we use μ_c^q to indicate the usage of $\mathcal{F}_q$ by swarm c, where $\mu_c^q = 1$ if swarm c utilizes $\mathcal{F}_q$, and $\mu_c^q = 0$ otherwise. This satisfies:

$$\sum_{q=1}^{Q} \mu_c^q = 1, \forall c \in \mathcal{C}, \tag{23}$$

$$\prod_{q=1}^{Q} \max_{c \in \mathcal{C}} \{\mu_c^q\} = 1. \tag{24}$$

Accordingly, we define $\mathcal{U} = \{\mathcal{U}_1, \mathcal{U}_2, \cdots, \mathcal{U}_C\}$ as the set of utility functions of all swarms, and $\mathcal{U}^{f_q}$ as the utility function set for swarms sharing $\mathcal{F}_q$, where $\mathcal{U}^{f_q} \subset \mathcal{U}$.

Furthermore, it is important to note that when different FBs are shared among swarms, a certain condition has to be satisfied to avoid mutual interference. The detailed derivation and explanation will be presented in V, and here we directly adopt the conclusion, as follows:

$$\prod_{c_i, c_l \in \mathcal{C}^q, c_i \ne c_l} d_{c_i c_l} \ge 10^{\frac{\rho + \alpha(n)}{20}} \cdot \sqrt{\frac{(P_J)^{n-1}}{P_1}} \cdot d_{c_i}, \forall q \in \{1, 2, \cdots Q\}, \tag{25}$$

where n denotes the number of swarms sharing the same FB, $\alpha(n) = 20(n-2)\lg(\frac{\lambda}{4\pi})$, $d_{c_i c_l}$ denotes the shortest distance between CH_{c_i} and member UAVs in swarm c_l, while d_{c_i} represents the farthest distance between member UAVs and the head UAV in swarm c_i. This expression implies that interference can be avoided if the distance between UAVs exceeds a certain threshold.

As a result, we formulate the inter-swarm resource allocation problem as $\boldsymbol{P2}$ as follows:

$$\begin{aligned}
\boldsymbol{P2}: \min_{\mathcal{P}_c, \Psi_c, \mathcal{K}_c} \quad & \mathbb{E}\left[\left(\frac{1}{|\mathcal{U}^{f_q}|}\sum_{U \in \mathcal{U}^{f_q}} U - \mathbb{E}\left[\frac{1}{|\mathcal{U}^{f_q}|}\sum_{U \in \mathcal{U}^{f_q}} U\right]\right)^2\right], \\
& q \in \{1, 2, \cdots Q\} && (26)\\
\text{s.t.} \quad & U_c = \sum_{m=1}^{|\Pi|} w_m \cdot Y_c^{\pi_m}, && (26a)\\
& \sum_{q=1}^{Q} |\mathcal{K}_q| = |\mathcal{K}|, && (26b)\\
& \mathcal{F}_p \cap \mathcal{F}_q = \varnothing, p \ne q, \forall p, q \in \{1, 2, ..., Q\}, && (26c)\\
& \mathcal{F}_1 \cup \mathcal{F}_2 \cup \cdots \cup \mathcal{F}_Q = \mathcal{K}, && (26d)\\
& K_c = f_q, \quad \forall c \in \mathcal{C}^q,\ q \in \{1, 2, \cdots, Q\} && (26e)\\
& \sum_{q=1}^{Q} \mu_c^q = 1, c \in \mathcal{C}, && (26f)\\
& \prod_{q=1}^{Q} \max_{c \in \mathcal{C}} \{\mu_c^q\} = 1, \mu_c^q \in \{0, 1\}, && (26g)\\
& \prod_{\substack{c_i, c_l \in \mathcal{C}^q, \\ c_i \ne c_l}} d_{c_i c_l} \ge 10^{\frac{\rho + \alpha(n)}{20}} \cdot \sqrt{\frac{(P_J)^{n-1}}{P_1}} \cdot d_{c_i}, \\
& \qquad \forall q \in \{1, 2, \cdots Q\}, && (26h)\\
& Q \le C, && (26i)\\
& (18a) - (18e).
\end{aligned}$$

The optimization objective is to minimize the variance of the sum of utility functions for swarms sharing the same FB. The constraint in (26a) represents the utility function of the swarm. (26b) guarantees that the total number of sub-channels allocated across all swarms is equal to the overall available sub-channels. (26c) and (26d) represent FBs that do not overlap with each other and collectively form the network spectrum resources. (26e) ensures that swarms sharing the same FB have an equal number of sub-channels, and each swarm is restricted to occupying at most one FB, as stated in (26f). (26g) indicates that each FB is utilized by at least one swarm. (26h) specifies the condition for multiple swarms to share the same FB without causing interference. (26i) is the constraint of FB number. Finally, (18a)-(18e) are the constraints of intra-swarm resource allocation.

Given the aforementioned, we define the USNT resource allocation problem as $\boldsymbol{P}$, comprising two related subproblems. The optimization objective is to maximize the STP across the network while ensuring inter-swarm fairness.

$$\boldsymbol{P}: \max_{\mathcal{P}_c,\Psi_c,\mathcal{K}_c} \quad \sum_{c=1}^{C} U_c \tag{27}$$

$$\text{s.t.} \min_{\mathcal{P}_c,\Psi_c,\mathcal{K}_c} \quad \mathbb{E}\left[\left(\frac{1}{|\mathcal{U}^{f_q}|}\sum_{U\in\mathcal{U}^{f_q}} U - \mathbb{E}\left[\frac{1}{|\mathcal{U}^{f_q}|}\sum_{U\in\mathcal{U}^{f_q}} U\right]\right)^2\right],$$

$$q \in \{1,2,\cdots Q\}, \tag{27a}$$

$$(18a) - (18e),$$

$$(26a) - (26i),$$

where (27a) represents the inter-swarm optimization constraints, while (18a)-(18e) define the intra-swarm constraints, and (26a)-(26i) specify the inter-swarm constraints.

V. Resource allocation method for the UAV swarm Network

In this section, we first derive the conditions for inter-swarm FB sharing and propose the AFBS algorithm, which abstracts the network model and utilizes an optimized graph coloring method to achieve FB sharing among different swarms. Subsequently, we present a DDQN-based dynamic frequency block competition algorithm (DDQN-FBCA), which uses DRL to allocate a specific number of sub-channels for each FB, enabling dynamic adaptation of sub-channels resources among different swarms, while ensuring fairness among the swarms. We further introduce a GA-based multi-prioritization dynamic resource allocation algorithm (GAP-DRA) for power and sub-channel allocation within each swarm, with the optimization goal of maximizing STP. At the same time, the utility function is calculated to provide reward guidance for DDQN. Finally, through the aforementioned steps, we achieve resource sharing under limited network resources, maximizing the STP for a USNT while ensuring fairness among the swarms. The overall framework of the USNT resource allocation method is shown in Fig. 2.

A. Adaptive Frequency Block Sharing Algorithm (AFBS)

In this section, we provide a detailed mathematical derivation of the conditions for FB sharing. Based on this, we propose the AFBS algorithm to achieve FB sharing among swarms, addressing the issue of limited network resources in the USNT. In **Algorithm 1**, since the swarm interference can be modeled as edges in an undirected graph, and FB sharing among swarms can be represented as assigning the same color to adjacent nodes, the graph coloring method is introduced to generate an initial color (or FB) allocation solution for the swarms. In **Algorithm 2**, we initially use the graph coloring method to generate an initial color (or FB) allocation solution for the swarms. Then, we evaluate whether the swarms contained within each color in each solution satisfy condition (32). If not, we add the swarms contained within that color to the set until condition (32) is satisfied, and reallocate colors among the swarms in the set to form a new solution. With the aim of alleviating the impact of inter-swarm heterogeneity, we further select the one with the minimum

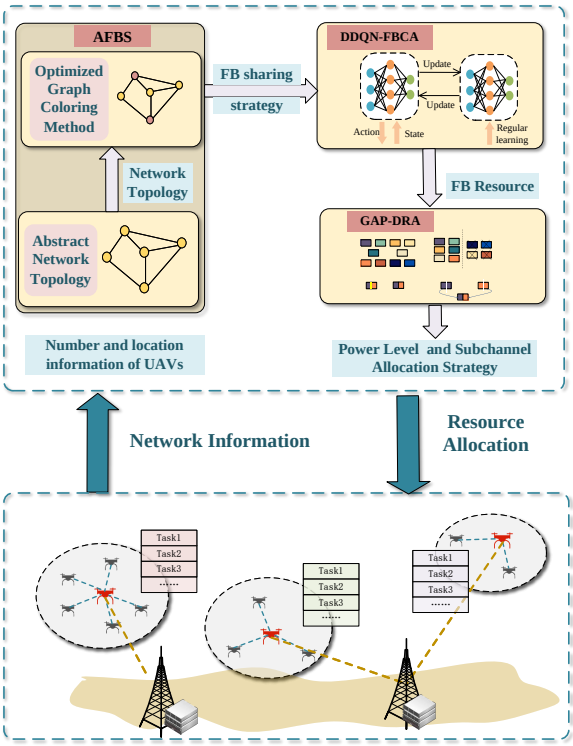


Fig. 2. Framework for resource allocation in the USNT.

variance of the number of UAVs in each swarm as the optimal solution among all feasible solutions, and finally obtain the FB allocation scheme for each swarm. The pseudocode of AFBS is shown in **Algorithm 1**.

In AFBS, it is necessary to determine whether the swarms can use the same FB. Therefore, we will provide a detailed derivation of the conditions for FB sharing. To illustrate the conditions for FB sharing, we provide an example in Fig. 3, which involves two swarms, c_1 and c_2. The condition for FB sharing between c_1 and c_2 is that they do not interfere with each other. For c_1, the signals of c_2 are viewed as interference (green dotted lines in the figure) and the signal of c_1 is a useful signal (red dotted line in the figure). Due to the orthogonal partitioning of the FB, if both c_1 and c_2 occupy $\mathcal{F}_q$, there exists at least one UAV pair, one from each swarm, sharing the same sub-channel $k \in \mathcal{F}_q$.

To avoid interference among UAVs in distinct swarms and to ensure efficient spectrum utilization, the frequency reuse methods employed in urban cellular networks are adopted. Each swarm can be viewed as a microcell in the USNT, where the swarm head UAV functions as a BS to receive and transmit signals from member UAVs. We consider a more general case: Assume that swarms $c_1, c_2, \ldots, c_n$ use the same FB. For swarm head CH_{c_1}, the signals within the swarm are considered useful signals, while the signals from other swarms are considered interference. For convenience, we denote the distance from the farthest UAV within swarm c_1 to its swarm head CH_{c_1} as d_{c_1}, and the distance from the nearest UAV within other swarms to the swarm head CH_{c_1} as $d_{c_1 c_n}$. To ensure correct demodulation at c_1, the received

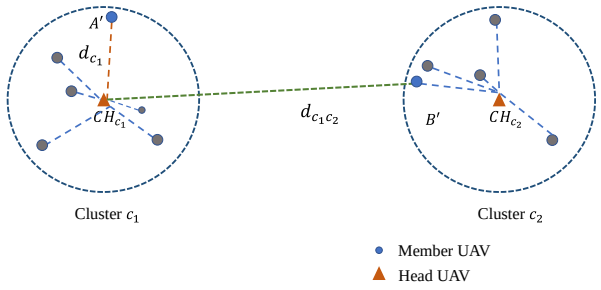


Fig. 3. Two swarms and their UAVs.

power of the useful signal and the interference signal must satisfy the following condition (the antenna gain is assumed to be $G_t = G_r = 1$):

$$\begin{aligned} &P'_{c_1} - \left(P'_{c_2} + P'_{c_3} + \cdots + P'_{c_n}\right) \\ &= 10\lg\left(\frac{P_{c_1}\lambda^2}{(4\pi)^2 d_{c_1}^2}\right) - \sum_{i=2}^{n} 10\lg\left(\frac{P_{c_i}\lambda^2}{(4\pi)^2 d_{c_1c_i}^2}\right) \\ &= 10\lg\left(\frac{P_{c_1}}{\prod_{i=2}^{n} P_{c_i}}\right) + 20\lg\left(\frac{\prod_{i=2}^{n} d_{c_1c_i}}{d_{c_1}}\right) \\ &+ 20(n-2)\lg(\frac{4\pi}{\lambda}) \geq \rho, \end{aligned} \tag{28}$$

where ρ is the interference-free threshold. Since we employ discrete power control, then

$$10\lg\frac{P_1}{P_J^{n-1}} \leq 10\lg\frac{P_{c_1}}{\prod_{i=2}^{n} P_{c_i}} \leq 10\lg\frac{P_J}{P_1^{n-1}}. \tag{29}$$

Let $\alpha(n) = 20(n-2)\lg(\frac{\lambda}{4\pi})$, and the (28) can be rewritten as

$$20\lg\frac{\prod_{i=2}^{n} d_{c_1c_i}}{d_{c_1}} \geq \rho + \alpha(n) + 10\lg\frac{P_J^{n-1}}{P_1}. \tag{30}$$

Simplifying the above equation, we obtain

$$\prod_{i=2}^{n} d_{c_1c_i} \geq 10^{\frac{\rho+\alpha(n)}{20}} \cdot \sqrt{\frac{(P_J)^{n-1}}{P_1}} \cdot d_{c_1}. \tag{31}$$

Equation (31) represents the condition for no interference at c_1 when $c_1, c_2, \cdots, c_n$ use the same FB. At the same time, $c_2, \ldots, c_n$ must also satisfy similar conditions to ensure that the FB is shared by $c_1, c_2, \ldots, c_n$ without interference. Based on the above analysis, we express the condition for n swarms to share the FB without interference as follows:

$$\prod_{l\neq i, l=1,2,\cdots n}^{n} d_{c_ic_l} \geq 10^{\frac{\rho+\alpha(n)}{20}} \cdot \sqrt{\frac{(P_J)^{n-1}}{P_1}} \cdot d_{c_i}, \forall i \in \{1, 2, \cdots n\}. \tag{32}$$

Algorithm 1 AFBS.

INPUT: Number and location of swarm head UAV and member UAVs, discrete power P_1 and P_j
OUTPUT: FB sharing solution

Generate network $G = \langle V, E\rangle$, $V = \varnothing$, $E = \varnothing$;
For c in $\mathcal{C}$, **do**:
  For c' in $\mathcal{C}$, **do**:
    if c and c' not satisfy (32) and (c, c') not in E and $c \neq c'$, **then**
      add c and c' to V;
      add (c, c') to E;
Execute **Algorithm 2**, and obtain the FB sharing solution.

Algorithm 2 Optimized Graph Coloring.

INPUT: $G = \langle V, E\rangle$, number of colors Q, $\mathcal{TM} = \varnothing$
OUTPUT: Optimal graph coloring solution

Calculate all available coloring solutions $\mathbb{O} = \{O_1, O_2, \ldots\}$, where $O_i = \{o_1^i, o_2^i, \ldots, o_Q^i\}$, and $o_q^i = \{c_k\}$ is a set of c with color q in solution O_i [40];
For O_i in $\mathbb{O}$, **do**:
  For o_q^i in O_i, **do**:
    While $|o_q^i| \geq 3$ and (32) is not satisfied, **do**:
      place one c from o_q into $\mathcal{TM}$;
    For o_q^i in $\mathcal{TM}$:
      Place the colors sequentially and save the allocation once (32) is satisfied. Repeat the process until all o_q^i are allocated, and update O_i accordingly;
For O_i in $\mathbb{O}$, **do**:
  Calculate $\text{VAR}_{value} = \frac{1}{Q}\sum_{q=1}^{Q} \text{Var}(|\mathcal{CM}_k| + 1)$, where $c_k \in o_q^i$;
Return the solution O_i which has the minimum VAR_{value}.

Equation (32) describes the conditions under which n swarms can share the same FB without interference. Building upon the previous discussion, we propose the AFBS algorithm, which first abstracts the network and then allocates FBs using an optimized graph coloring method that accounts for inter-swarm differences. The pseudocode for AFBS is presented in **Algorithm 1** and **Algorithm 2**.

B. DDQN-based Dynamic Frequency Block Competition Algorithm (DDQN-FBCA)

In this section, we propose a DDQN-FBCA for sub-channel resource allocation across FBs. DDQN [41] enables adaptive decision-making in dynamic environments by incorporating a target network and a double update mechanism, effectively reducing the overestimation of Q-values. This method enhances algorithmic stability and accuracy, making it highly suitable for addressing complex resource allocation challenges. The principle of DDQN is shown in Fig. 4.

State ($\mathcal{S}$): The state space is defined as the number of sub-channels available in different FBs.

$$\mathcal{S} = [f_1, f_2, \cdots, f_Q]. \tag{33}$$

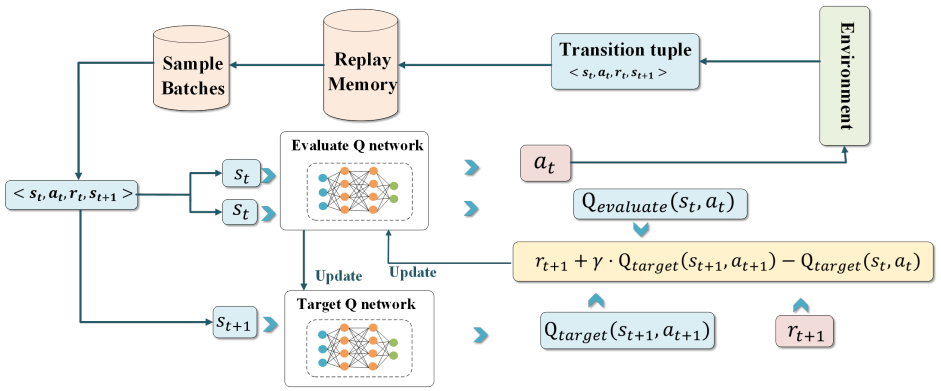


Fig. 4. DDQN algorithm framework.

Action ($\mathcal{A}$): The action represents the operation performed by the agent at each time step. In this paper, the action space is defined as the adjustments made by the agent to the number of sub-channels allocated to FBs for swarms at each step, expressed as

$$\mathcal{A} = \{A_i\}, \tag{34}$$

where $A_i = [a_{i1}, a_{i2}, \cdots, a_{iQ}]$, $a_{iq} \in \{-1, 0, 1\}$, $q \in \{1, 2, \cdots, Q\}$.

1) In each A_i, only one $a_{iq} = 1$;
2) The sum of elements in each A_i satisfies $\sum a_{iq} = 0$;
3) $|\mathcal{A}| = A_Q^2$, where A_Q^2 denotes the permutation of $(Q, 2)$;
4) For any two actions A_i and A_j $(i \neq j)$, $A_i \neq A_j$, ensuring all actions are distinct.

Reward (R): The reward represents the feedback received by the agent upon executing a specific action in the current environment.

Typically, the reward is closely associated with the utility function. To effectively translate the optimization objective into a reward, the following considerations are incorporated:

1) The primary objective is to maximize the STP. Thus, the reward function should integrate the intra-swarm utility function as presented in equation (17);
2) Additionally, fairness in inter-swarm resource allocation is deemed a critical consideration to enhance the utility value generated per unit of resource. Consequently, the reward function is designed to reflect disparities in the utility functions across different swarms.
3) To address the challenge of reward sparsity and enhance the stability of DRL, a stepwise structure is introduced into the reward function.

Therefore, we design the reward of the agent as:

$$R = g\left\{-\log\left(\mathbb{E}\left[\frac{1}{|\mathcal{U}^{f_q}|}\sum_{U\in\mathcal{U}^{f_q}} U - \mathbb{E}\left[\frac{1}{|\mathcal{U}^{f_q}|}\sum_{U\in\mathcal{U}^{f_q}} U\right]\right]\right)^2\right\}. \tag{35}$$

where $g(x)$ is a step function designed to ensure that the reward value remains within a certain range, thereby stabilizing the performance of DDQN. At each time slot, the agent obtains the current network state, denoted as S_t, from the network. Based on the agent parameters ω, the agent identifies the current action corresponding to the maximum Q-value in the current Q-network, expressed as:

$$A_t = \arg\max_{A'} Q(S_t, A', \omega). \tag{36}$$

Then A_t is applied to the UAV network, resulting in R_t and S_{t+1}, and the tuple (S_t, A_t, R_t, S_{t+1}) is subsequently stored in the experience replay buffer $\mathcal{D}$.

Once $\mathcal{D}$ accumulates sufficient experiences, the agent randomly selects samples to compute the target Q-value:

$$y_t = R_t + \gamma \cdot Q'(S_{t+1}, \arg\max_{A'} Q(S_{t+1}, A', \omega), \omega'). \tag{37}$$

The loss function of DDQN is defined as:

$$Loss = E\left[(y_t - Q(S_t, A_t, \omega))^2\right]. \tag{38}$$

As the neural network parameters ω are randomly initialized, it is crucial to encourage exploration of different actions during the early training stages. We employ the ϵ-greedy method for action selection and dynamically update the exploration rate in each episode to balance exploration and exploitation:

$$\xi = 1 - \frac{2}{\pi}\arctan(episode/40), \tag{39}$$

where $episode$ denotes the training episodes. To accelerate the training process and improve its effectiveness, we adopt a pre-training method, applying the resulting model to a network with dynamic task arrivals.

The overall flow of DDQN-FBCA is as follows: First, we execute AFBS (**Algorithm 1**) to obtain the used FBs in swarms, using the number of sub-channels contained in each FB as the state and initializing the DRL network parameters. During each $episode$, the DDQN-FBCA initializes the state space, selects an action, and calculates the reward through GAP-DRA (see Subsection V-C) to obtain the new state, then stores the data in the experience replay buffer. Subsequently, a data batch is randomly sampled, and the Q-network parameters are updated. After a certain number of steps, the target Q-network parameters are updated. Finally, the Q-network parameters are saved and the pre-training is completed. When tasks dynamically arrive, actions are performed using the current network parameters, followed by reward calculation.

If the reward is below a set threshold, the state is reset. The pseudocode for DDQN-FBCA is presented in **Algorithm 3**.

Algorithm 3 DDQN-FBCA.

INPUT: EP, replay buffer $\mathcal{D}$ and its capacity D_{max}, training batch size N_b, target Q-network update frequency N_u, task set $\mathcal{B}$, environment parameters and UAVs parameters.

Execute **Algorithm 1** to obtain the set of swarms sharing the FBs;
Initialize the parameters of the Q-network and the target Q-network as ω and ω' respectively, with $\omega \rightarrow \omega'$, and set global_step $= 0$;
For $episode = \{1, 2, \cdots, EP\}$, **do**:
    STEP $= 0$
    While STEP $\leq$ MAX STEP, **do**:
        Randomly initialize state space S_t;
        Obtain Q-values for all actions from the Q-network. Use ϵ-greedy method to select A_t;
        Get the current FBs resource status, execute the intra-swarm **Algorithm 4** (see Subsection V-C) with the action A_t to determine the new state S_{t+1} and the reward R_t.
        Store $\langle S_t, A_t, R_t, S_{t+1} \rangle$ in $\mathcal{D}$ for training;
        Randomly sample a minibatch of N_b tuples $\sim$ Unif($\mathcal{D}$) and update the Q network parameters ω;
        If global_step$\%N_u = 0$:
            $\omega \rightarrow \omega'$;
        Update state count;
        $S_t = S_{t+1}$;
Save the Q-network parameters.
For $t = \{1, 2, \cdots\}$, **do**:
    Obtain the current state S_t and select action A for tasks arriving at the current time, then calculate S_{t+1} and R_t.
    If $R_t < R_{th}$:
        Reset State

C. GA-Based Multi-Prioritization Dynamic Resource Allocation (GAP-DRA)

In this section, a GA-based intra-swarm power and sub-channels allocation strategy is proposed. This strategy leverages the efficiency of GA in optimizing complex and nonlinear resource distributions, enabling the system to adapt to dynamic conditions and achieve near-optimal solutions. The GA mainly involves encoding, fitness function design, and the processes of selection, crossover, and mutation. The detailed design and description of each of these components are provided as follows.

1) Encoding: The transmit power and sub-channels selection of UAVs are both represented using binary encoding. The encoding for power is defined as

$$Power_c = \{\varphi_{11}, \cdots, \varphi_{1n}; \cdots; \varphi_{M_c1}, \cdots, \varphi_{M_cn}\}, \quad (40)$$

where $\forall \varphi \in Power_c$, $\varphi \in \{0, 1\}$, $n = \lceil \log_2 J \rceil$, and $\lceil \cdot \rceil$ denotes the ceiling function. The actual value of the encoding satisfies

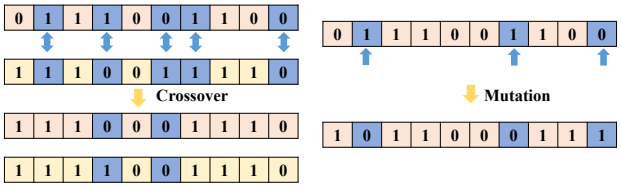


Fig. 5. Crossover and mutation.

$$f_{bin}(\varphi_{i1}, \cdots, \varphi_{in}) = p_c^i, \quad (41)$$

where $f_{bin}(\cdot)$ represents the binary-to-decimal mapping. Similarly, the encoding for sub-channels selection is defined as follows:

$$Sub_c = \{\nu_{11}, \cdots, \nu_{1f_q}; \cdots; \nu_{M_c1}, \cdots, \nu_{M_cf_q}\}, \quad (42)$$

where $|Sub_c| = M_c \times f_q$, ν_{ij} denotes the occupancy of the j-th sub-channel by the CM_c^i, and $\nu_{ij} \in \{0, 1\}$.

Therefore, a chromosome can then be expressed as

$$\zeta = \{Power_c, Sub_c\} = \{\varphi_{11}, \cdots, \varphi_{1n}; \cdots; \varphi_{M_c1}, \cdots, \varphi_{M_cn}; \nu_{11}, \cdots, \nu_{1f_q}; \cdots; \nu_{M_c1}, \cdots, \nu_{M_cf_q}\}. \quad (43)$$

2) Fitness Function: We define the fitness function of swarm c as:

$$F_c(Power_c, Sub_c) = U_c + \Gamma(Power_c), \quad (44)$$

where Γ is the penalty function. To avoid the UAV power from exceeding its maximum value during chromosome crossover and mutation in each generation, the penalty function is formulated as:

$$\Gamma(Power_c) = \begin{cases} -\left[f_{bin}(Power_c) - P_{max}\right], & \text{if } f_{bin}(Power_c) \geq P_{max} \\ 0, & \text{otherwise.} \end{cases} \quad (45)$$

3) Selection, Crossover and Mutation: The GA employs a survival-of-the-fittest selection method, where the probability of an individual being selected to advance to the next generation is given by:

$$p_i = \frac{F_{c,i}}{\sum F_{c,i}}, i \in \{1, 2, \cdots, N_{pop}\}. \quad (46)$$

Here, $F_{c,i}$ represents the fitness function of the i-th chromosome, and N_{pop} denotes the population size. To enhance the global search capability of the algorithm, GA performs crossover between chromosomes with a probability p_c during each iteration. To prevent the algorithm from converging to a local optimum, mutations are introduced with a certain probability, as depicted in Fig. 5.

Based on the above analysis and design, we present GAP-DRA to optimize the power and sub-channel allocation within the swarm. First, the population is initialized, and the fitness is calculated. Then, through selection, crossover, and mutation operations, the population is iteratively updated. Ultimately, the optimal power and sub-channel allocation results are obtained to maximize the STP of high-priority tasks. The pseudocode for GAP-DRA is shown in **Algorithm 4**.

Algorithm 4 GAP-DRA.

INPUT: Population size N_{pop}, number of member UAVs M_c, number of sub-channels f_q, power encoding length n, number of iterations N_{it}.

OUTPUT: Power levels and sub-channels allocation solution.

1: Initialize the population $pop = \{\zeta_1, \zeta_2, \cdots, \zeta_{N_{pop}}\}$, $|\zeta| = M_c \times (n + f_q)$;
2: For $iter = \{1, 2, \cdots, N_{it}\}$, **do**:
3: For $\zeta_i \in pop$, calculate the fitness function values $\{F_{c,i}\}$, where $i = 1, 2, \cdots, N_{pop}$;
4: Select N_{pop} chromosomes based on the probability p_i in (46), and form a new population $pop1$;
5: Perform crossover on $pop1$ with the probability p_c to generate $pop2$;
6: For $\zeta_i \in pop2$, apply mutation with a mutation probability p_m, and form population $pop3$, where $i = 1, 2, \cdots, N_{pop}$;
7: Calculate the fitness function values for both the pop and $pop3$, denoted as $F_{c,i}^{Parent}$ and $F_{c,i}^{Child}$, respectively, where $i = 1, 2, \cdots, N_{pop}$;
8: Compare the fitness values between $F_{c,i}^{Parent}$ and $F_{c,i}^{Child}$, and retain the chromosome corresponding to the higher fitness value to update the pop.

D. Complexity Analysis

1) Time Complexity: For the DDQN-FBCA, the time complexity is $O(EP \times STEP \times N_b \times F(\theta))$, where θ represents all learnable parameters of the DDQN, and $F(\theta)$ denotes the floating point operations required for a single forward pass of the DDQN. For the GAP-DRA, the time complexity is $O(N_{it} \times N_{pop} \times M_c \times (n + f_q))$.

2) Space Complexity: For the DDQN-FBCA, the space complexity is $O(\theta) + O(D_{max} \times (|\mathcal{S}| + 1 + 1)) = O(\theta) + O(D_{max} \times (Q+2)) \approx O(\theta + D_{max} \times Q)$. For the GAP-DRA, the space complexity is $O(N_{pop} \times M_c \times (n + f_q))$.

Remark 2: The proposed algorithm demonstrates strong potential for practical deployment in emergency rescue and post-disaster communication scenarios. By enabling UAV swarms to provide efficient and reliable communication services under constrained network resources, the algorithm ensures the prioritized and reliable transmission of critical data, such as rescue command and decision information. Additionally, our proposed algorithm is designed with low computational overhead, where both the DDQN-FBCA and GAP-DRA algorithms exhibit polynomial time and space complexity, ensuring computational efficiency and scalability. Moreover, the adoption of a hierarchical management architecture facilitates both global and local resource optimization, while possible adaptive swarm head reselection can further enhance system robustness against single-point failures. Therefore, the proposed algorithm exhibits high scalability, low coordination overhead, and strong resilience, making it well-suited for large-scale, resource-limited, and dynamic operational environments.

VI. Simulation Results

A. Simulation Settings

Simulations for USNT resource allocation were conducted using Python on a computer running Debian, equipped with a 3.20 GHz Intel i9 CPU, 64 GB of RAM, and an NVIDIA GeForce RTX 4090D GPU with 24 GB of VRAM. The simulation parameters are detailed in Table III. Three different network scales are considered, corresponding to swarm-UAV combinations of $[6, 50]$, $[8, 62]$, and $[10, 82]$. In order to clearly illustrate the distribution of different UAVs and their resource usage, Fig. 6 depicts the two-dimensional distribution of UAVs, comprising six swarms, with the number of UAVs and their positions within each swarm randomly distributed. Dashed lines between swarm head UAVs indicate that these swarms cannot share the same FB. Task transmissions begin at each time slot, with decisions made and effective for the entire slot duration. Each UAV is limited to transmitting one task per time slot.

B. Performance Analysis

We define STP as the ratio of the number of successfully transmitted tasks to the total number of tasks in swarm c over a given period. It represents the performance of network transmission tasks, which is expressed as follows:

$$\text{STP} = \frac{|\mathcal{B}_c^{succ}|}{|\mathcal{B}_c|}, c \in \mathcal{C}, \tag{47}$$

where $\mathcal{B}_c^{succ}$ represents the set of successfully transmitted tasks in swarm c, and $\mathcal{B}_c$ denotes the set of total tasks in swarm c. As tasks arrive dynamically, $|\mathcal{B}_c^{succ}|$ and $|\mathcal{B}_c|$ mean the total number of tasks over a period of time.

Then we use the variance of STP across swarms, denoted as STPV, to represent the fairness of inter-swarm resource allocation:

$$\text{STPV} = \frac{1}{C} \sum_{i=1}^{C} \left(\text{STP}_i - \overline{\text{STP}}\right)^2. \tag{48}$$

where $\overline{\text{STP}}$ is the average STP of all swarms, and STP_i is the STP of the i-th swarm.

For intra-swarm resource allocation, we adopt particle swarm optimization (PSO) [42] and simulated annealing (SA) [43] as baseline algorithms to validate the performance of GAP-DRA. In addition, multi-agent DQN (MADQN) and pre-alliance concept from [44] are also introduced into intra-swarm resource allocation framework (denoted as MADQN and PAPS, respectively). For the former, each UAV in the swarm is considered as an individual agent, selecting sub-channel and power allocation based on the environmental state and utility function. For the latter, PAPS evaluates the attraction value of UAVs under various power and sub-channel combinations to determine the optimal intra-swarm resource allocation method, and performs joint power and sub-channel optimization through utility function assessment. To evaluate the performance of DDQN-FBCA, for inter-swarm resource allocation, the baseline algorithm employs an equal distribution method, named the equal-partition-based inter-swarm

FB allocation method (EP-FBA). And the sub-channels are uniformly allocated across FBs, i.e., $f_1 = f_2 = \cdots = f_Q$. All simulation results are collected and analyzed after the network reaches stability.

TABLE III
THE VALUE OF THE SIMULATION PARAMETERS

Parameter	Value
UAV altitude range	$[100\ \text{m}, 200\ \text{m}]$
The range of UAVs	$5\ \text{km} \times 5\ \text{km}$
The number of UAVs	50, 62, 82
The number of swarms	6, 8, 10
K	$20, 25, 30$
f	2 GHz
N_0	-174 dBm/Hz
β_0	-60 dB
P_{max}	0.3 W, 0.5 W, 0.7 W
SINR_{th}	3 [25]
α	2 [45]
τ	1 s
B	75 kHz
ρ	9 [46]
W	$[0.5, 0.3, 0.2]$
N_{it}	50
p_c	0.7
p_m	0.3
N_{pop}	50
D_{max}	800
N_b	128
Learning rate	0.001
γ	0.9
Priority number	3
Task size	$[500\ \text{kbit}, 1\ \text{Mbit}]$

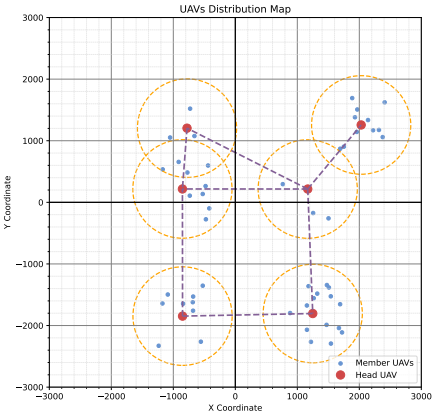


Fig. 6. A USNT with six swarms and corresponding UAV distribution.

We first evaluate the convergence of the overall algorithm GA-DDQN-FBCA, which adopts DDQN-FBCA for inter-swarm operations and GAP-DRA for intra-swarm operations. The total number of training steps is set to approximately 120k. Fig. 7 illustrates the variations in the Q-network loss and rewards of DDQN under the combinations of $P_{max} = 0.3$ W, $P_{max} = 0.5$ W and $P_{max} = 0.7$ W with $K = 30$. During the initial 10k steps, both the loss and reward curves exhibit rapid changes. Around 20k steps, the loss curve stabilizes, and the reward curve approaches a steady state. After 20k steps, the reward curve demonstrates minor fluctuations locally, which can be attributed to the variations caused by state resets in the DDQN. Based on this analysis, we conclude that GA-DDQN-FBCA achieves successful convergence during training and demonstrates the capability to make high-reward decisions under the given system states. This also indicates that the proposed algorithm efficiently allocates resources in dynamic environments.

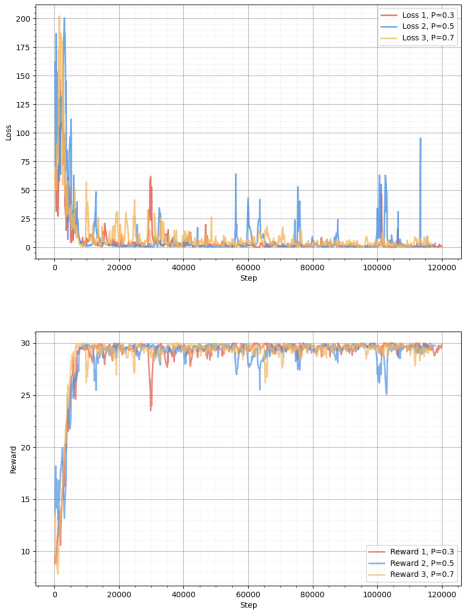


Fig. 7. Loss and reward curve.

Table IV presents the training time and inference time of GA-DDQN-FBCA under different network scenarios and sub-channel conditions. The training time is the duration from the algorithm's start to convergence, while the inference time is the algorithm's execution time over 500 steps. The results indicate that the overall training time and inference time of the algorithm are reasonable, validating the feasibility of the algorithm in practical applications. Specifically, Scenario 1 shows the lowest training time and inference time, while Scenario 3 exhibits the highest values for both training time and inference time across all sub-channel conditions. Similarly, the training time and inference time also increase as the sub-channel size increases, with Scenario 3 having the highest training time and inference time in all cases, which is consistent with the analysis in Section V-D. This means that the GA-DDQN-FBCA method demonstrates strong practical deployment capability and suitability over different conditions.

Fig. 8 presents the performance of AFBS. It can be observed that, under all parameter settings, AFBS consistently enhances the algorithm's performance on STP. This indicates that AFBS enhances the utilization of network resources through FB

TABLE IV
THE TRAINING TIME AND INFERENCE TIME OF GA-DDQN-FBCA ($P_{max} = 0.3, 0.5, 0.7$ W).

Network Scenarios	Training time			Inference time		
	$K = 20$	$K = 25$	$K = 30$	$K = 20$	$K = 25$	$K = 30$
Scenario 1	83.269 s	93.011 s	104.112 s	39.789 s	40.760 s	43.190 s
Scenario 2	89.516 s	98.493 s	127.893 s	53.312 s	57.307 s	58.522 s
Scenario 3	122.493 s	134.104 s	137.205 s	62.340 s	65.423 s	67.073 s

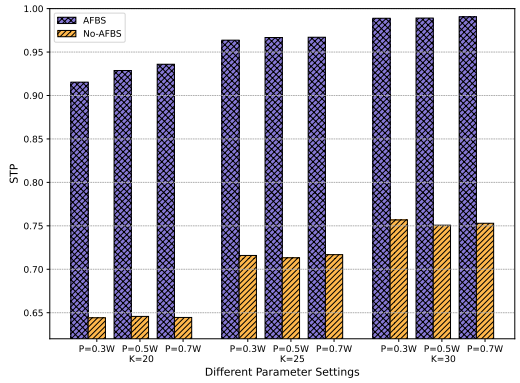
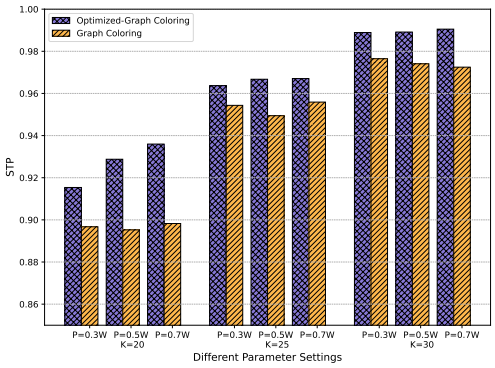
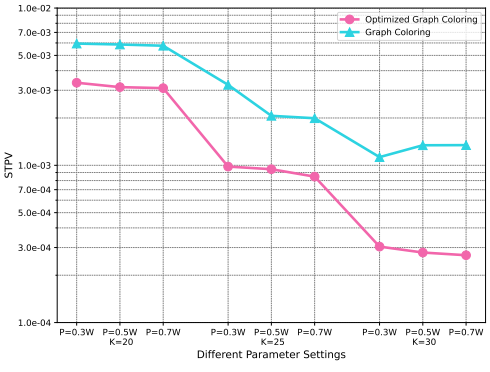


Fig. 8. The STP with different parameters.



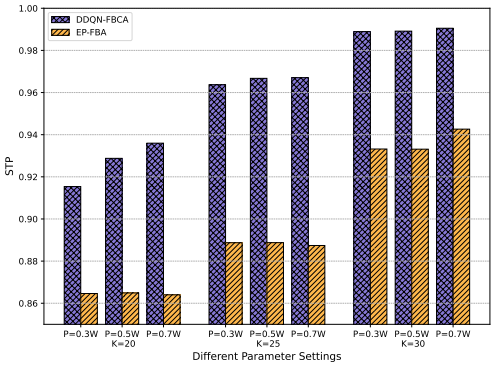
(a)



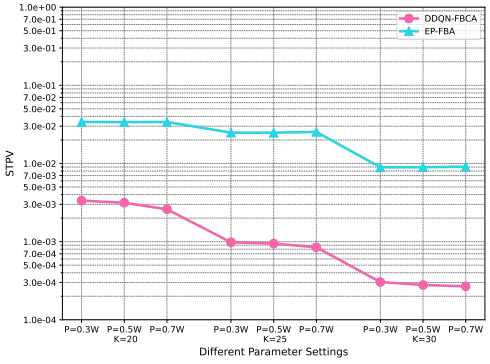
(b)

Fig. 9. Optimized graph coloring vs. graph coloring comparison (a) STP. (b) STPV.

sharing, thereby ensuring the transmission of more tasks by efficiently allocating resources. Overall, AFBS achieves a



(a)



(b)

Fig. 10. DDQN-FBCA vs. EP-FBA under all parameter settings. (a) STP. (b) STPV.

25.61% improvement in STP.

To demonstrate the effectiveness of the optimized graph coloring method, Fig. 9 presents a comparison of the average STP and STPV between the conventional graph coloring method and the optimized graph coloring method. The results indicate that the optimized method consistently outperforms the conventional method across all parameter settings. This improvement can be attributed to the consideration of inter-swarm variations in the optimized method, which enables more efficient FB allocation, leading to higher STP and improved inter-swarm fairness. Overall, compared with the conventional method, the optimized graph coloring method improves the STP by 2.07% and reduces the STPV by 62.45%.

To demonstrate the effectiveness of the inter-swarm algorithm, we present the performance of DDQN-FBCA and EP-FBA in terms of average STP and STPV as shown in Fig. 10, with intra-swarm operations employing the GAP-DRA

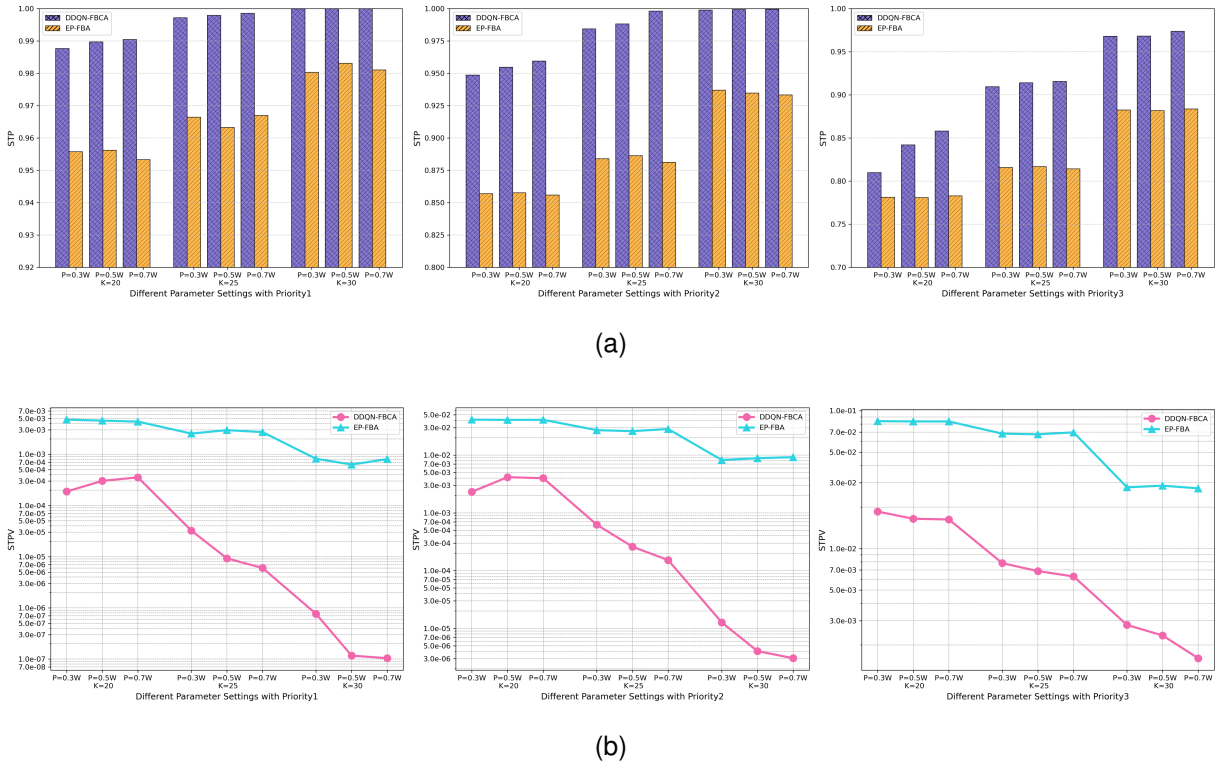


Fig. 11. DDQN-FBCA vs. EP-FBA at different priorities (a) STP. (b) STPV.

TABLE V
PERFORMANCE IMPROVEMENT OF STP UNDER DIFFERENT CONDITIONS.

Parameter	$P = 0.3$ W	$P = 0.5$ W	$P = 0.7$ W
$K = 20$	5.87%	7.39%	8.33%
$K = 25$	8.44%	8.77%	8.97%
$K = 30$	5.79%	6.00%	5.08%

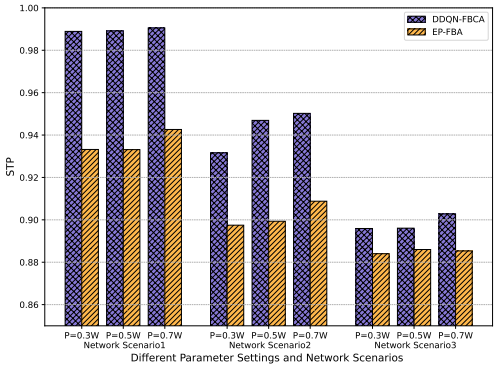
TABLE VI
PERFORMANCE IMPROVEMENT OF STPV UNDER DIFFERENT CONDITIONS.

Parameter	$P = 0.3$ W	$P = 0.5$ W	$P = 0.7$ W
$K = 20$	90.14%	90.71%	92.36%
$K = 25$	96.04%	96.19%	96.66%
$K = 30$	96.60%	96.88%	97.05%

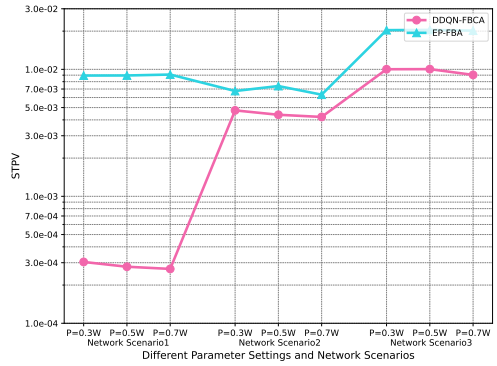
algorithm. It can be observed that DDQN-FBCA consistently outperforms EP-FBA in both average STP and STPV, demonstrating its ability to allocate resources effectively by taking inter-swarm differences into account. Under the same power level, the average STP of DDQN-FBCA increases with the number of sub-channels, while under the same sub-channel conditions, the improvement is less significant, indicating that sub-channels significantly impact the algorithm by enhancing network capacity. In Fig. 10b, it can also be observed that with a larger number of sub-channels, the two algorithms perform better in ensuring fairness, indicating that, under conditions of sufficient network resources, the algorithm is more capable of effectively scheduling network resources. In general, this method ensures fairness among swarms and contributes to an improvement in the overall STP, thereby ensuring the reliability of task transmission. Specifically, DDQN-FBCA improves STP by 7.20% and reduces STPV by 94.74% compared to EP-FBA.

Table V and Table VI summarize the performance improvements in STP and STPV achieved by DDQN-FBCA and EP-FBA. As shown in Table V, it can be observed that under the same power conditions, the impact of sub-channels on performance improvement is significant. In addition, under the same power conditions, the best performance improvement occurs at $K = 25$, while the minimum performance improvement is observed at $K = 30$. This can be attributed to the fact that when resources are limited (e.g., $K = 20$, $K = 25$), DDQN-FBCA has more flexibility to allocate appropriate resources for task transmission based on inter-swarm differences. However, when resources are more abundant (e.g., $K = 30$), the EP-FBA can meet the transmission requirements of most tasks within the network. From Table VI, it can be observed that under conditions of more abundant resources, DDQN-FBCA demonstrates better performance in enhancing fairness. Overall, DDQN-FBCA achieves an average STP improvement of 7.33% and reduces the STPV by 94.57%.

Fig. 11 compares the average STP and STPV of DDQN-FBCA and EP-FBA for priority 1, 2, and 3 under different power levels and numbers of sub-channels. The results illustrate that, under priority levels 1, 2, and 3, DDQN-FBCA consistently outperforms the EP-FBA in both STP and STPV. Moreover, priority 1 tasks achieve better STP and STPV than priority 2 and priority 3 tasks, with the STP for priority 1 tasks approaching 1. This indicates that the proposed algorithm effectively enhances the STP for high-priority tasks while

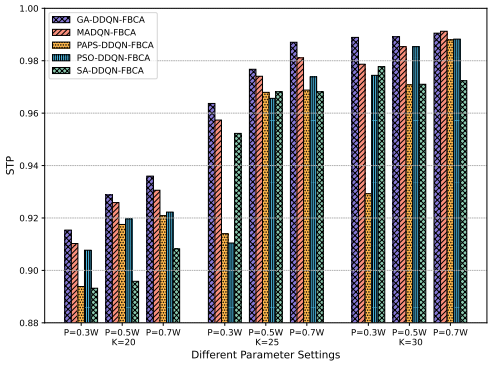


(a)

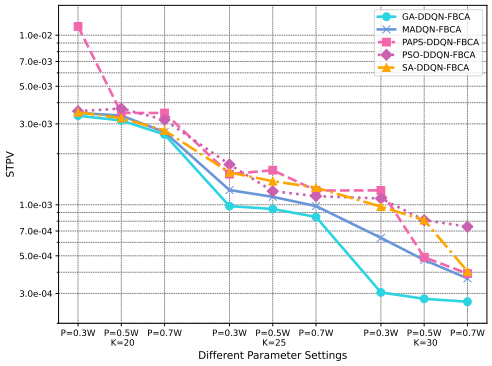


(b)

Fig. 12. DDQN-FBCA vs. EP-FBA under different network scenarios, $K = 30$. (a) STP. (b) STPV.



(a)



(b)

Fig. 13. STP and STPV of GA, MADQN, PAPS, PSO and SA under all parameter settings. (a) STP. (b) STPV.

ensuring fairness.

To validate the effectiveness of DDQN-FBCA under different network scenarios, Fig. 12 presents the variations of STP and STPV for DDQN-FBCA and EP-FBA across three network scenarios. The results show that DDQN-FBCA consistently achieves significantly higher STP than EP-FBA under all parameter settings and network scenarios, demonstrating its advantage in ensuring successful task transmissions. Furthermore, Fig. 12b indicates that although the STPV of DDQN-FBCA tends to increase sharply as the network scale expands, its average performance still surpasses that of EP-FBA.

For each of the three groups of parameter settings, Fig. 13 illustrates the variations in the average STP and STPV under different intra-swarm algorithms (GA, MADQN, PAPS, PSO and SA) and an inter-swarm algorithm (DDQN-FBCA) across three priority levels, namely GA-DDQN-FBCA, MADQN-FBCA, PAPS-DDQN-FBCA, PSO-DDQN-FBCA and SA-DDQN-FBCA. We can observe that GA consistently outperforms the other algorithms in both STP and STPV, confirming its superior efficiency in enhancing system performance and task reliability. Under favorable sub-channel conditions, STP growth slows as power increases due to limited potential for further improvement. As shown in Fig. 13b, GA also achieves the best load-balancing performance. Additionally, we observe that the proposed algorithm outperforms the MADQN algorithm in both STP and STPV. This is because MADQN faces dimensionality explosion when handling discrete problems, making it difficult to achieve optimal solutions within a short time, and its non-stationarity leads to significant performance fluctuations. The proposed algorithm effectively combines the advantages of heuristic methods for handling discrete problems and DRL for addressing dynamic tasks, thereby ensuring fairness while achieving more reliable task transmission. Overall, GA improves average STP by 0.62%, 2.45%, 1.53%, and 2.01%, and reduces STPV by 23.2%, 40.16%, 37.21%, and 31.38% compared with MADQN, PAPS, PSO, and SA, respectively.

Fig. 14 illustrates the variations of STP and STPV for different intra-swarm algorithms under various network scenarios and transmission powers with the setting of $K = 30$. The results show that GA-DDQN-FBCA consistently achieves the highest STP, indicating that it maintains optimal performance across different network scenarios and power conditions. In addition, the STPV results reveal that GA-DDQN-FBCA exhibits the lowest variability, demonstrating better stability and robustness under varying network scenarios, whereas the baselines experience significantly increased fluctuations as the network scale expands. Furthermore, we also observe that the proposed algorithm outperforms the MADQN algorithm under different network scenarios. These results demonstrate that the combined GA-DDQN framework effectively integrates the advantages of DRL and GA, ensuring more reliable task transmission and swarm fairness.

Fig. 15 shows the STP performance of GA, MADQN,

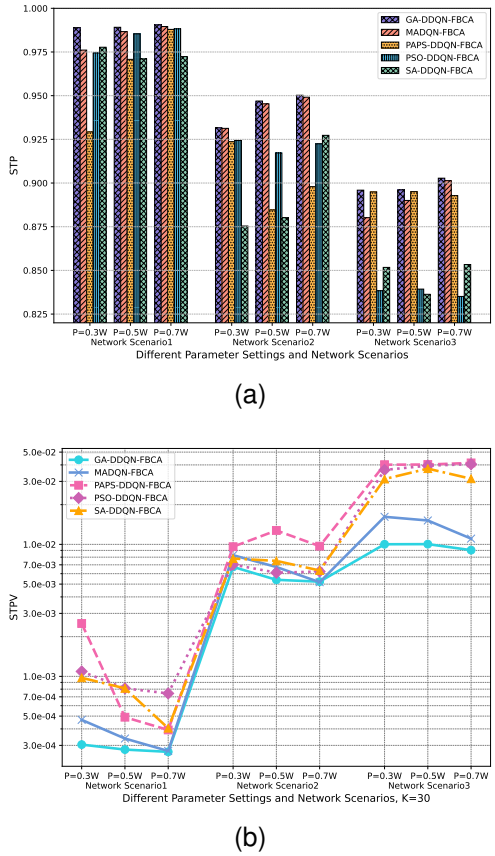


Fig. 14. STP and STPV of GA, MADQN, PAPS, PSO and SA under different parameter settings and different network scenarios, $K = 30$. (a) STP. (b) STPV.

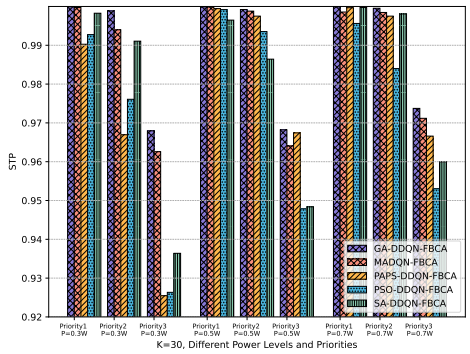


Fig. 15. The variation of STP with different parameters under $K = 30$.

PAPS, PSO, and SA under different power levels and priorities for $K = 30$. Across all power levels, GA achieves the highest STP for all tasks, with priority 1 tasks consistently outperforming lower priorities. Notably, the STP improvement for priority 3 tasks is more pronounced than for priority 1 and 2 tasks under abundant network resources. This is because, when network resources are abundant, additional resources can be allocated to low-priority tasks while still guaranteeing the transmission requirements of high-priority tasks.

VII. Conclusion

In this paper, we investigate resource allocation algorithms for USNTs, aiming to maximize STP while ensuring inter-swarm fairness. Motivated by the limited wireless resources and inter-swarm differences, we propose AFBS, which integrates FB sharing with an optimized graph coloring method. We then decompose the resource allocation problem into two sub-problems, presenting optimization expressions for both sub-problems and the global problem. Based on this, we propose DDQN-FBCA and GAP-DRA for specific sub-channel and power allocation. In DDQN-FBCA, we design an efficient reward function with the goal of inter-swarm fairness to guide the learning process of the DDQN agent, effectively allocating sub-channels for FBs among different swarms. In GAP-DRA, we implement intra-swarm power level and sub-channel allocation based on GA by designing intra-swarm utility and fitness functions. Finally, simulation experiments were conducted to validate the proposed algorithms, demonstrating the effectiveness of the optimized graph coloring, inter-swarm resource allocation method, and intra-swarm resource allocation method. Specifically, the AFBS increases the average STP by 25.61% compared with no-AFBS, while DDQN-FBCA enhances the average STP by 7.33% and reduces STPV by 94.57% compared with EP-FBA. Additionally, GAP-DRA effectively ensures the reliable transmission of high-priority tasks. Thus, the proposed algorithms significantly improve STP, ensure inter-swarm fairness, and provide a valuable reference for resource allocation in USNTs in the future.

In the future, an important and challenging research direction is the swarm head failure issue, which can lead to a significant degradation or collapse of network performance. Therefore, we will focus on investigating protection, re-election, and recovery mechanisms for the swarm head to enhance the adaptability and robustness of USNTs.

Acknowledgment

This work is supported in part by the Science and Technology Innovation Program of Xiongan New Area (Grant No.2025XAGG0055), in part by the National Natural Science Foundation of China (Grant No. 62372076), in part by the Education Department Research Foundation of Anhui Province (Grant No. DTR2023051), in part by the Federal Ministry of Research, Technology, and Space (BMFTR), Germany, through the Project 6GEM+ (Grant 16KIS2411), and in part by the European Union's Horizon Europe research and innovation programme under the 6G-Path project (Grant No. 101139172).

References

[1] N. Mohamed, J. Al-Jaroodi, I. Jawhar, A. Idries, and F. Mohammed, "Unmanned aerial vehicles applications in future smart cities," *Technological forecasting and social change*, vol. 153, p. 119293, 2020.

[2] N. Zhao, W. Lu, M. Sheng, Y. Chen, J. Tang, F. R. Yu, and K.-K. Wong, "UAV-assisted emergency networks in disasters," *IEEE Wireless Communications*, vol. 26, no. 1, pp. 45–51, 2019.

[3] N. Cheng, S. Wu, X. Wang, Z. Yin, C. Li, W. Chen, and F. Chen, "AI for UAV-assisted IoT applications: A comprehensive review," *IEEE Internet of Things Journal*, vol. 10, no. 16, pp. 14 438–14 461, 2023.

[4] D. Mourtzis, J. Angelopoulos, and N. Panopoulos, "UAVs for industrial applications: Identifying challenges and opportunities from the implementation point of view," *Procedia Manufacturing*, vol. 55, pp. 183–190, 2021.

[5] N. Gupta, S. Agarwal, and D. Mishra, "Trajectory design for throughput maximization in UAV-assisted communication system," *IEEE Transactions on Green Communications and Networking*, vol. 5, no. 3, pp. 1319–1332, 2021.

[6] S. K. Nobar, M. H. Ahmed, Y. Morgan, and S. A. Mahmoud, "Resource allocation in cognitive radio-enabled UAV communication," *IEEE Transactions on Cognitive Communications and Networking*, vol. 8, no. 1, pp. 296–310, 2021.

[7] A. Benfaid, N. Adem, and B. Khalfi, "Adaptsky: A DRL based resource allocation framework in NOMA-UAV networks," in *2021 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2021, pp. 01–07.

[8] Y. Li, H. Zhang, K. Long, C. Jiang, and M. Guizani, "Joint resource allocation and trajectory optimization with QoS in UAV-based NOMA wireless networks," *IEEE Transactions on Wireless Communications*, vol. 20, no. 10, pp. 6343–6355, 2021.

[9] Y. Cai, Z. Wei, R. Li, D. W. K. Ng, and J. Yuan, "Energy-efficient resource allocation for secure UAV communication systems," in *2019 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2019, pp. 1–8.

[10] F. Zeng, Z. Hu, Z. Xiao, H. Jiang, S. Zhou, W. Liu, and D. Liu, "Resource allocation and trajectory optimization for QoE provisioning in energy-efficient UAV-enabled wireless networks," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 7, pp. 7634–7647, 2020.

[11] J. Huang, F. Shan, J. Luo, R. Xiong, and W. Wu, "ASSUME: An optimal algorithm to minimize uav energy by altitude and speed scheduling," *IEEE Transactions on Mobile Computing*, 2025.

[12] Y. Wang, J. Huang, F. Shan, Y. Gao, R. Xiong, and J. Luo, "Optimizing joint speed and altitude schedule for UAV data collection in low-altitude airspace," *IEEE Transactions on Mobile Computing*, 2025.

[13] H. Wang, H. Zhang, X. Liu, K. Long, and A. Nallanathan, "Joint UAV placement optimization, resource allocation, and computation offloading for thz band: A DRL approach," *IEEE Transactions on Wireless Communications*, vol. 22, no. 7, pp. 4890–4900, 2022.

[14] T. Khurshid, W. Ahmed, M. Rehan, R. Ahmad, M. M. Alam, and A. Radwan, "A DRL strategy for optimal resource allocation along with 3D trajectory dynamics in UAV-MEC network," *IEEE Access*, vol. 11, pp. 54 664–54 678, 2023.

[15] I. Valiulahi and C. Masouros, "Multi-UAV deployment for throughput maximization in the presence of co-channel interference," *IEEE Internet of Things Journal*, vol. 8, no. 5, pp. 3605–3618, 2020.

[16] A. Shamsoshoara, F. Afghah, A. Razi, S. Mousavi, J. Ashdown, and K. Turk, "An autonomous spectrum management scheme for unmanned aerial vehicle networks in disaster relief operations," *IEEE access*, vol. 8, pp. 58 064–58 079, 2020.

[17] Y. Lin, M. Wang, X. Zhou, G. Ding, and S. Mao, "Dynamic spectrum interaction of UAV flight formation communication with priority: A deep reinforcement learning approach," *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 3, pp. 892–903, 2020.

[18] B. Yang, T. Taleb, Y. Shen, X. Jiang, and W. Yang, "Performance, fairness, and tradeoff in UAV swarm underlaid mmwave cellular networks with directional antennas," *IEEE Transactions on Wireless Communications*, vol. 20, no. 4, pp. 2383–2397, 2020.

[19] H. Gao, J. Feng, Y. Xiao, B. Zhang, and W. Wang, "A UAV-assisted multi-task allocation method for mobile crowd sensing," *IEEE Transactions on Mobile Computing*, vol. 22, no. 7, pp. 3790–3804, 2022.

[20] Y. He, Y. Gan, H. Cui, and M. Guizani, "Fairness-based 3-D multi-UAV trajectory optimization in multi-UAV-assisted MEC system," *IEEE Internet of Things Journal*, vol. 10, no. 13, pp. 11 383–11 395, 2023.

[21] Y. Shen, Y. Qu, C. Dong, F. Zhou, and Q. Wu, "Joint training and resource allocation optimization for federated learning in UAV swarm," *IEEE Internet of Things Journal*, vol. 10, no. 3, pp. 2272–2284, 2023.

[22] H. Hao, C. Xu, W. Zhang, S. Yang, and G.-M. Muntean, "Joint task offloading, resource allocation, and trajectory design for multi-UAV cooperative edge computing with task priority," *IEEE Transactions on Mobile Computing*, vol. 23, no. 9, pp. 8649–8663, 2024.

[23] S. Guo and X. Zhao, "Multi-agent deep reinforcement learning based transmission latency minimization for delay-sensitive cognitive satellite-UAV networks," *IEEE Transactions on Communications*, vol. 71, no. 1, pp. 131–144, 2022.

[24] M. M. Alam and S. Moh, "Joint optimization of trajectory control, task offloading, and resource allocation in air–ground integrated networks," *IEEE Internet of Things Journal*, 2024.

[25] J. Cui, Y. Liu, and A. Nallanathan, "Multi-agent reinforcement learning-based resource allocation for UAV networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 2, pp. 729–743, 2019.

[26] C. Zhu, G. Zhang, and K. Yang, "Fairness-aware task loss rate minimization for multi-UAV enabled mobile edge computing," *IEEE Wireless Communications Letters*, vol. 12, no. 1, pp. 94–98, 2022.

[27] M. Li, A. Dong, G. Wang, X. Tian, S. Li, and J. Yu, "Wireless resource optimization for UAV swarm cooperative sensing via multi-agent multi-task deep reinforcement learning," *IEEE Internet of Things Journal*, 2026.

[28] L. Kuixian, L. Jinjie, G. Xin, Y. Yandie, C. Haipeng, W. Liangtian, L. Yun *et al.*, "Dynamic decision-making of UAV swarm based on constrained multi-objective optimization under incomplete interference information," *Chinese Journal of Aeronautics*, p. 103846, 2025.

[29] T. Li, S. Leng, Z. Wang, K. Zhang, and L. Zhou, "Intelligent resource allocation schemes for UAV-swarm-based cooperative sensing," *IEEE Internet of Things Journal*, vol. 9, no. 21, pp. 21 570–21 582, 2022.

[30] X. Dai, Z. Lu, X. Chen, X. Xu, and F. Tang, "Multiagent RL-based joint trajectory scheduling and resource allocation in NOMA-assisted UAV swarm network," *IEEE Internet of Things Journal*, vol. 11, no. 8, pp. 14 153–14 167, 2024.

[31] C. Fan, C. She, H. Zhang, B. Li, C. Zhao, and D. Niyato, "Learning to optimize user association and spectrum allocation with partial observation in mmwave-enabled UAV networks," *IEEE Transactions on Wireless Communications*, vol. 21, no. 8, pp. 5873–5888, 2022.

[32] Z. Mou, Y. Zhang, F. Gao, H. Wang, T. Zhang, and Z. Han, "Deep reinforcement learning based three-dimensional area coverage with UAV swarm," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 10, pp. 3160–3176, 2021.

[33] K. Wang, T. Zhao, Y. Yuan, J. Chu, Z. Chen, and H. Dui, "Resilience evaluation and resource allocation in UAV-enabled IoT via multi-swarm logistics support," *IEEE Internet of Things Journal*, 2025.

[34] H. Li, Q. Sun, Y. Zhong, Z. Huang, and Y. Zhang, "A soft resource optimization method for improving the resilience of UAV swarms under continuous attack," *Reliability Engineering & System Safety*, vol. 237, p. 109368, 2023.

[35] T. Li, S. Leng, X. Liao, and Y. Zhang, "Digital twin-based task-driven resource management in intelligent UAV swarms," *IEEE Transactions on Intelligent Transportation Systems*, 2025.

[36] Q. Li, Z. Wang, H. Yao, T. Mai, Z. Li, and M. Guizani, "Dynamic routing mechanism for load distribution in UAV swarm networks with edge caching," *IEEE Transactions on Mobile Computing*, 2025.

[37] Z. Hu, Z. Zheng, L. Song, T. Wang, and X. Li, "UAV offloading: Spectrum trading contract design for UAV-assisted cellular networks," *IEEE Transactions on Wireless Communications*, vol. 17, no. 9, pp. 6093–6107, 2018.

[38] X. Yang, T. Yu, Z. Chen, J. Yang, J. Hu, and Y. Wu, "An improved weighted and location-based clustering scheme for flying ad hoc networks," *Sensors*, vol. 22, no. 9, p. 3236, 2022.

[39] Y. Zeng, R. Zhang, and T. J. Lim, "Wireless communications with unmanned aerial vehicles: Opportunities and challenges," *IEEE Communications Magazine*, vol. 54, no. 5, pp. 36–42, 2016.

[40] D. W. Matula, G. Marble, and J. D. Isaacson, "Graph coloring algorithms," in *Graph theory and computing*. Elsevier, 1972, pp. 109–122.

[41] H. Van Hasselt, A. Guez, and D. Silver, "Deep reinforcement learning with double Q-learning," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 30, no. 1, 2016.

[42] R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *MHS'95. Proceedings of the sixth international symposium on micro machine and human science*. IEEE, 1995, pp. 39–43.

[43] S. Kirkpatrick, C. D. Gelatt Jr, and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671–680, 1983.

[44] Q. Peng, H. Wu, N. Li, and F. Wang, "A dynamic task allocation method for unmanned aerial vehicle swarm based on wolf pack labor division model," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 8, no. 6, pp. 4075–4089, 2024.

[45] Q. Wu, Y. Zeng, and R. Zhang, "Joint trajectory and communication design for multi-UAV enabled wireless networks," *IEEE Transactions on Wireless Communications*, vol. 17, no. 3, pp. 2109–2121, 2018.

[46] H. Binjie, D. Xinyan, and Z. Jianbin, "GSM principles and network optimization," 2009.

Zhuojia Yang is currently pursuing a Ph.D. degree at the School of Electronics and Information, Beijing Jiaotong University. She received her Master's degree in Information and Communication Engineering in 2022 and Bachelor's degree in Communication Engineering in 2019 from the School of Electronic and Information Engineering, Beijing Jiaotong University. Her research interests include network resource allocation and UAV network communications.

Wei Su was born in October 1978. He got the Ph.D. degrees in Communication and Information Systems from Beijing Jiaotong University in January 2008. Now he is a professor in the School of Electronic and Information Engineering, Beijing Jiaotong University. Dr. Su Wei is mainly engaged in researching key theories and technologies for the next generation Internet and has taken part in many national projects such as National Basic Research Program (also called 973 Program), the Projects of Development Plan of the State High Technology Research, the National Natural Science Foundation of China. He currently presides over the research project Fundamental Research on Cognitive Services and Routing of Future Internet, a project funded by the National Natural Science Foundation of China.

Bin Yang received his Ph.D. degree in systems information science from Future University Hakodate, Japan in 2015. He was a research fellow with the School of Electrical Engineering, Aalto University, Finland, from Nov. 2019 to Nov. 2021. He is currently a professor with the School of Computer and Information Engineering, Chuzhou University, China. His research interests include unmanned aerial vehicle networks, cyber security, edge computing, and Internet of Things.

Yudong Ma received the B.E. degree in Electronic and Information Engineering from the School of Electronic and Information Engineering, Tiangong University, Tianjin, China, in 2020 and the M.S. degree in Information and Communication Engineering from the School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing, China, in 2024. He is currently pursuing his Ph.D. in Artificial Intelligence at the Institute of Artificial Intelligence, Beihang University, Beijing, China. His research focuses on navigation and surveillance technology based on L-band digital aeronautical communication systems.

Yihua Peng is currently pursuing the Ph.D. degree with School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing, China. His primary research interests include satellite network and in-band network telemetry.

Qi Liu received the B.S. degree in information and communication engineering and Ph.D. degree in communication and information system from Beijing Jiaotong University, Beijing, China, in 2003 and 2009, respectively. Then she works as a post-doctor in electronic engineering department in Tsinghua University from 2009 to 2011. She is currently a professorate senior engineer in Smart City Research Institute of China Unicom. Her research interests focus on 5G, cooperation of heterogeneous networks, Internet of Vehicles and High-Precision Positioning.

Tarik Taleb Prof. received the B.E. degree with distinction in Information Engineering and the M.Sc. and Ph.D. degrees in Information Sciences from Tohoku University, Sendai, Japan, in 2001, 2003, and 2005, respectively. He is currently a Full Professor at Ruhr University Bochum (RUB), Germany, where he leads research activities on next-generation mobile and distributed systems. Prior to joining RUB, he was a Professor at the University of Oulu, Finland (2018–2023), and an Associate Professor at Aalto University, Finland (2014–2021). Earlier in his career, he served as a Senior Researcher and 3GPP Standards Expert at NEC Europe Ltd., Heidelberg, Germany, where he contributed to the evolution of mobile network architectures and standardization activities. Before joining NEC, he was an Assistant Professor at the Graduate School of Information Sciences, Tohoku University, Japan, working in a research laboratory fully funded by KDDI. He also held a Research Fellowship at the Intelligent Cosmos Research Institute, Japan, from 2005 to 2006. Prof. Taleb is widely recognized for his pioneering contributions to mobile network softwarization, network slicing, cloud-edge continuum management, and autonomous networking. His current research interests include autonomous network and service management, edge-cloud continuum systems, network softwarization and slicing, software-defined security, and AI-native communication networks.