

Joint Spectrum, Association, and Deployment Optimization for UAV Swarm-Assisted ISAC Networks

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

Abstract—Unmanned aerial vehicle (UAV) swarm-assisted integrated sensing and communication (ISAC) networks are a crucial technology for providing communication and sensing services in emergency rescue scenarios without base station support. However, the strong coupling between communication and sensing resources in such networks fundamentally limits the communication and sensing performance of ISAC systems. This paper jointly optimizes spectrum allocation, UAV association and deployment to maximize average system throughput while ensuring localization accuracy in such networks, where sensing is realized through localization. We begin by deriving an analytical expression for localization accuracy, which explicitly captures the joint effects of link quality and anchor geometry under shared communication-localization spectrum resources. We then formulate average system throughput maximization as a mixed-integer nonlinear and non-convex optimization problem with the constraints of localization accuracy, sub-channels, UAV association, UAV deployment and signal-to-interference-plus-noise ratio. We further develop an alternating iterative optimization method to solve this complex optimization problem. Within this method, a particle swarm optimization-based method is developed to jointly optimize spectrum allocation and UAV association, and a dueling double deep Q-network-based method is further employed for UAV deployment optimization. Finally, extensive simulation results are presented to validate the efficiency of our optimization method, and also to illustrate how key parameters influence average system throughput and localization accuracy.

Index Terms—UAV swarm networks, ISAC, localization, spectrum and association allocation, UAV deployment optimization, alternating iterative optimization.

I. INTRODUCTION

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), in part by the European Union's Horizon Europe research and innovation programme under the 6G-Path project (Grant No. 101139172). (Corresponding authors: Wei Su, Bin Yang)

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: yang-binchi@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).

INTEGRATED sensing and communication (ISAC) networks are widely regarded as a cornerstone technology for low-altitude intelligent networks and 6G systems, providing low-latency, highly reliable, and intelligent decision-making capabilities for critical applications such as emergency rescue, intelligent transportation, and industrial automation [1], [2]. Sensing includes target detection, environmental perception, parameter estimation, tracking, and localization. Among these, localization has attracted extensive attention for its ability to provide accurate and real-time user location information in emergency rescue scenarios. However, conventional global navigation satellite systems (GNSS) and cellular-based localization methods experience severe performance degradation in emergency scenarios characterized by non-line-of-sight (NLoS) propagation and infrastructure damage, limiting their ability to deliver reliable real-time localization and communication services [3], [4] and ultimately compromising rescue effectiveness. Unmanned aerial vehicle (UAV) swarm-assisted ISAC networks have emerged as a promising solution by leveraging UAVs' rapid deployment, high mobility, and line-of-sight (LoS)-enabled communication links [5]–[7]. To meet the stringent requirements of emergency rescue, it is therefore critical to enhance both sensing and communication performance in such networks.

Existing works on UAV-assisted networks can be broadly classified into three categories: communication performance optimization [8]–[16], localization performance optimization [17]–[25], and joint communication and localization performance optimization [26]–[34]. The joint optimization studies are most relevant to our work and are summarized in Table I. Although they provide valuable insights, they are still insufficient for base station (BS)-free emergency rescue ISAC networks. On the one hand, several existing UAV-assisted ISAC studies still rely on functional BS support [26]–[29], which limits their applicability in disaster scenarios where terrestrial infrastructure may be damaged or unavailable. On the other hand, the existing BS-free studies usually optimize only part of the decision variables, such as UAV anchor association [30], [32] or power control [32], [33], and some of them still rely on traditional optimization methods [30], [34]. As a result, they cannot fully capture how spectrum allocation, UAV association, and UAV deployment jointly shape interference, link quality and anchor geometry in dynamic environments, which together determine the communication and localization performance. In particular, communication and localization signals may share the same limited spectrum

TABLE I
COMPARISON WITH EXISTING WORKS.

Work	Base Station-Free	UAV Swarm	Optimization Parameters	Optimization Objective	Solution Method
[26]			3D UAV position, communication bandwidth and power allocation	Maximize communication rate	Traditional method
[27]		✓	UAV Deployment	Minimize UAV deployment numbers	Traditional method
[28]			Anchor selection	Optimize transmission power and localization accuracy	Traditional method
[29]			UAV trajectory and transmission scheduling	Minimize energy consumption	Traditional method
[30]	✓	✓	Anchor selection	Minimize sensing error	Traditional method
[31]			Waveform design	/	/
[32]	✓	✓	UAV trajectory, power and association	Maximize communication and localization performance	Reinforcement learning
[33]	✓	✓	Anchor selection and power allocation	Maximize energy efficiency	Multi-agent reinforcement learning
[34]	✓	✓	UAV trajectories and transmission scheduling	Minimize localization error	Traditional method
This work	✓	✓	Resource allocation, UAV association and UAV deployment	Maximize throughput and ensure localization accuracy	Deep reinforcement learning and heuristic method

resources, which introduces severe mutual interference and directly affects communication performance, while localization performance depends both on link quality and geometric configuration of multiple aerial anchors. Therefore, jointly optimizing spectrum allocation, association and UAV deployment to improve communication and localization performance in BS-free UAV swarm-assisted ISAC networks remains a challenging problem.

Motivated by these limitations, this paper investigates a BS-free UAV swarm-assisted ISAC network and jointly optimizes spectrum allocation, UAV association, and deployment to maximize the average system throughput while ensuring accurate localization. This problem is particularly challenging because these variables are deeply coupled, and any local adjustment may simultaneously affect communication and localization links, interference patterns, and anchor geometry. Moreover, communication and localization compete for limited resources, leading to an intrinsic performance tradeoff between throughput enhancement and localization accuracy assurance. To address these challenges, we derive an analytical expression for localization accuracy, formulate a unified mixed-integer nonlinear and non-convex optimization problem (MINNLP), and develop a collaborative alternating iterative optimization method for BS-free UAV swarm-assisted ISAC networks. The main contributions of this paper are summarized as follows:

- Considering the complex effects of multiple factors on communication and localization performance, we derive an analytical expression for localization accuracy that explicitly captures the joint effects of link quality and anchor geometry in BS-free UAV swarm-assisted ISAC networks with shared communication-localization spectrum resources. Based on this, we formulate the joint spectrum allocation, UAV association, and UAV deployment problem as a MINNLP to maximize the average system throughput, subject to localization accuracy, sub-channel allocation, UAV association, UAV deployment, and signal-to-interference-plus-noise ratio (SINR) con-

straints.

- To handle the strong coupling among resource allocation, UAV association, and deployment involving both discrete and continuous decision variables, we propose a collaborative alternating iterative optimization method consisting of two complementary methods that leverage global search capability and effective optimization in complex decision spaces. In particular, a particle swarm optimization (PSO)-based method is developed to solve the discrete spectrum-allocation and UAV-association subproblem by exploiting its strong global search capability in combinatorial solution spaces, providing initial inputs for UAV deployment. Then a dueling double deep Q-network (DQN)-based method is designed for UAV deployment, exploiting the adaptability of DQN and its effectiveness in complex decision spaces, jointly maximizing average system throughput and ensuring localization accuracy. The integration of these two methods enables effective coordination among spectrum resource allocation, UAV association and UAV deployment strategies.
- Finally, extensive simulations demonstrate that our proposed method significantly improves average system throughput while maintaining high localization accuracy, outperforming baseline methods. These results validate that the proposed method can provide highly efficient communication and reliable localization services in disaster-affected environments.

The rest of this paper is organized as follows. Section II provides an overview of the related work. In Section III, we present the system model. Section IV gives the problem formulation. Section V develops an alternating iterative optimization method to solve the optimization problem. The simulation results are provided in Section VI. Finally, Section VII concludes this paper, highlighting the key findings.

II. RELATED WORK

Existing works on UAV-assisted networks can be categorized into communication-oriented, localization-oriented, and

joint communication-localization works.

A. UAV-assisted Communication or Localization

Regarding the communication service, UAVs usually serve as aerial BSs or relay nodes to ensure emergency communications of ground users in disaster-affected and infrastructure-damaged areas. Cheng et al. [8] and He et al. [9] jointly optimized user association and UAV deployment in UAV-assisted networks to improve fairness, system throughput [8], and quality of experience [9]. Geng et al. [10] further incorporated resource allocation into the optimization framework, aiming to minimize the system cost. Ge et al. [11] jointly optimized the bandwidth allocation, UAV-user association, and UAV trajectories to enhance the overall service stability using a traditional mathematical method. In [12], UAV positions, bandwidth allocation, and transmit power are jointly optimized to maximize the average utility in UAV-relay networks. Sun et al. [13] jointly optimized UAV deployment, association and resource allocation to ensure system fairness. Wang et al. [14] proposed a convolutional neural network and Q-learning-based framework for joint power allocation and service zone selection, followed by a deterministic user scheduling policy in UAV-assisted relay systems. Chen et al. [15] further proposed an approach combining non-cooperative theoretical game and multi-agent deep reinforcement learning (MADRL) to maximize system throughput and energy efficiency by a joint optimization of user association and UAV trajectory. MADRL was also used in [16] to optimize UAV deployment, association and resource allocation.

As for localization service, most existing studies focus on optimizing UAV parameters to enhance localization accuracy. Qu et al. [17] and Fatima et al. [18] respectively optimized energy beamforming and bandwidth allocation [17], and receiver placement [18] to improve localization or target estimation accuracy. Famili et al. [19], [20] enhanced localization accuracy under unknown transmitter locations by exploiting time difference of arrival measurements and Doppler shifts using a DRL-based method. Graff et al. [21] optimized pilot placement and power allocation to minimize time-of-arrival (ToA) estimation errors. Saikko et al. [22] exploited Fisher information and prior tracking information to select informative range, angle, and received signal strength measurements, thereby improving localization accuracy with fewer measurements. Jesús-Azabal et al. [23] investigated energy-aware anchor optimization by adjusting active anchors, balancing localization accuracy and energy consumption. Ebrahimi et al. [24] employed reinforcement learning to optimize UAV trajectories for improving localization accuracy in global positioning system (GPS)-denied environments. Kato et al. [25] developed a radio-propagation-aware UAV path planning method that optimizes observation strategies and location selection to reduce localization errors caused by signal strength fluctuations.

B. UAV-assisted ISAC

The UAV-assisted joint communication and localization service can be categorized into BS-supported and BS-free scenarios. For BS-supported scenarios, Bi et al. [26] considered

a UAV as both an airborne BS and anchor node, and jointly optimized UAV deployment, power allocation and bandwidth to maximize system throughput under localization accuracy constraints. Yang et al. [27] jointly optimized the number of UAVs, their spatial positions, and UAV-user associations to satisfy both ground communication and localization requirements. Yang et al. [28] used simulated annealing for UAV deployment and integrated heuristics to jointly optimize association, power, and sub-channel allocation, improving coverage and communication efficiency while reducing energy consumption. Wang et al. [29] proposed a joint optimization method for UAV trajectory and device transmission scheduling, achieving reduced energy consumption and improved 3D localization performance. Zhu et al. [30] proposed a Doppler-based UAV-swarm-assisted vehicular network framework for joint sensing and communication, and employed a differential evolution algorithm to achieve a trade-off of communication and sensing performance. However, in disaster scenarios where terrestrial infrastructure is damaged or unavailable, traditional localization methods become ineffective. Therefore, for BS-free scenarios, Tao et al. [31] proposed a novel ISAC scheme, where integrated waveform design, resource reuse and receiver processing enable joint data demodulation and sensing parameter estimation. This work primarily focuses on physical-layer waveform and transceiver design, while higher-layer resource scheduling is beyond its main scope. More specifically, Zhu et al. [33] and Xie et al. [32] mainly focus on localization performance in sensing, and investigate higher-layer resource scheduling and network-level optimization. Zhu et al. adopted the time-difference-of-arrival technique, and proposed a UAV swarm-based rural localization method to jointly optimize user association and power level to maximize communication and localization performance. Xie et al. jointly optimized UAV trajectory, transmit power and user association to balance communication throughput and localization performance. Similarly, Zhu et al. [34] achieved the objective of minimizing the average localization error of sensors by optimizing UAV trajectories and sensor transmission scheduling strategies, while ensuring successful data uploading.

III. SYSTEM MODEL

A. Network Model

As shown in Fig. 1, we consider a UAV swarm-assisted ISAC network for disaster rescue with M UAVs and N ground users. One UAV serves as the cluster head (CH) for decision-making, while the others act as cluster members (CMs) interacting with ground users. GPS is unreliable due to mountainous terrain and forest coverage, and damaged terrestrial infrastructure degrades communication, so the UAV swarm enables the network to provide communication and localization services. Ground users transmit signals to UAVs, which extract relevant information to determine UAV association, sub-channel allocation, and localization strategies. It is worth noting that communication and localization signals may occupy the same spectrum resources, as described in [35], [36]. Furthermore, the network has K spectrum resource blocks, each corresponding to a sub-channel of bandwidth

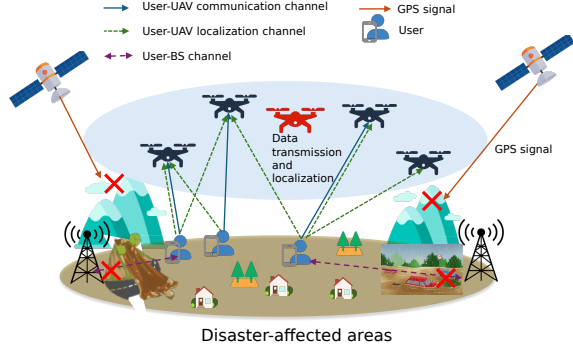


Fig. 1. The system model

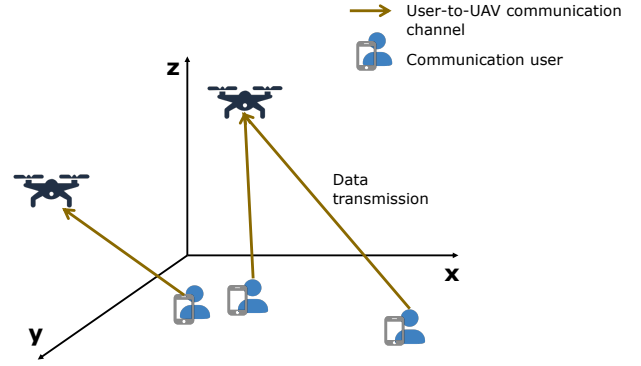


Fig. 2. Communication model

TABLE II
LIST OF MAJOR NOTATIONS

Symbol	Description
$\mathcal{U} / U_m$	The set of UAVs / the m -th UAV
$\mathcal{C} / C_n$	The set of ground users / the n -th user
M / N	The number of UAVs / ground users
$U_n^{\text{com}} / U_n^{\text{loc}}$	The set of UAVs communicating with C_n / The set of UAVs providing localization service for C_n
$\mathbf{p}_m / \mathbf{P}$	The 2-D coordinates of U_m / the set of $\mathbf{p}_m$
H	The altitude of the UAVs
$\mathbf{q}_n$	The 2-D coordinates of C_n
$\mathcal{K} / k_i / K$	The set of sub-channels / the i -th sub-channel / the number of sub-channels
g_n^m	The channel gain between U_m and C_n
$p^{\text{com}} / p^{\text{loc}}$	The transmit power of the communication signal / the localization signal
$a_{nm}^{\text{com}} / \mathcal{A}_n^{\text{com}}$	The indicator variable whether C_n selects the U_m for communication / the set of a_{nm}^{com}
$b_{nm}^{\text{loc}} / \mathcal{B}_n^{\text{loc}}$	The indicator variable whether C_n selects the U_m for localization
$\psi_{nk}^{\text{com}} / \Psi_n^{\text{com}}$	The indicator variable for C_n selects the k -th sub-channel for communication / the set of ψ_{nk}^{com}
$\psi_{nk}^{\text{loc}} / \Psi_n^{\text{loc}}$	The indicator variable for C_n selects the k -th sub-channel for localization / the set of ψ_{nk}^{loc}
$r_{nk}^{\text{com}} / r_n^{\text{com}}$	The communication SINR of U_n^{com} on k -th sub-channel / the total communication SINR at U_n^{com}
$I_{nk}^{\text{com}} / I_{nmk}^{\text{loc}}$	The interference power of other users to the communication / localization signal of C_n on the k -th sub-channel
N_{loc}	The number of UAVs used for localization
$(\sigma_\rho^{nm})^2$	The pseudorange measurement variance between U_m and C_n
ϵ_n / δ_n	The localization RMSE / accuracy of C_n
$r_{nm}^{\text{loc}} / U_n^{\text{loc}}$	The SINR of the localization signal at the $U_m \in U_n^{\text{loc}}$

B , denoted by $\mathcal{K} = \{k_i \mid i = 1, \dots, K\}$. Communication and localization sub-channels may overlap, and each user's communication or localization signal occupies at most one sub-channel. A CH serves as a system coordination center to manage sub-channel sharing between communication and localization signals across all UAVs and users.

B. UAV-User Channel Model

We first denote the set of UAVs in the network as $\mathcal{U} = \{U_1, U_2, \dots, U_M\}$, with the two-dimensional Cartesian coordinates of the m -th UAV U_m represented by $\mathbf{p}_m = (x_m^p, y_m^p) \in \mathbf{P}$, where $m \in \{1, 2, \dots, M\}$, $\mathbf{P}$ represents the coordinates set of the M UAVs, and the flight altitude

of the UAVs is H . The set of ground users is denoted as $\mathcal{C} = \{C_1, C_2, \dots, C_N\}$, and the two-dimensional Cartesian coordinates of each C_n are given by $\mathbf{q}_n = (x_n^q, y_n^q)$, where $n \in \{1, 2, \dots, N\}$.

As indicated in [37], the communication link between the UAVs and the ground users can be modeled using a probabilistic path loss model, which accounts for the occurrence probabilities of both LoS and NLoS conditions. The probability that the channel between U_m and C_n is a LoS channel is given by

$$P_{\text{LoS}} = \frac{1}{1 + \mathbf{a} \cdot \exp\left(-\mathbf{b} \cdot \left(\sin^{-1}\left(\frac{H}{d_n^m}\right) - \mathbf{a}\right)\right)}, \quad (1)$$

where $\mathbf{a}$ and $\mathbf{b}$ are environment-dependent constants, and $d_n^m = \sqrt{\|\mathbf{p}_m - \mathbf{q}_n\|^2 + H^2}$ represents the distance between U_m and C_n . Moreover, the probability of the NLoS channel is given by

$$P_{\text{NLoS}} = 1 - P_{\text{LoS}}. \quad (2)$$

Then, the average path-loss encompassing both LoS and NLoS components for the link between C_n and U_m can be determined as follows [38]:

$$PL_{nm} = (P_{\text{LoS}} \cdot \mu_{\text{LoS}} + P_{\text{NLoS}} \cdot \mu_{\text{NLoS}}) \left(\frac{4\pi f d_n^m}{c} \right)^2, \quad (3)$$

where f is the signal transmission frequency, c is the speed of light, and μ_{LoS} and μ_{NLoS} denote the attenuation factors for LoS and NLoS links, respectively. Then, the average channel gain between C_n and U_m can be expressed as $g_n^m = (1/PL_{nm})$ [38].

C. UAV-User Communication Model

Fig. 2 shows that a UAV swarm provides communication services for ground users. We use a_{nm}^{com} to indicate whether U_m provides communication service to C_n , and $\mathcal{A}_n^{\text{com}} = [a_{n1}^{\text{com}} \ a_{n2}^{\text{com}} \ \dots \ a_{nM}^{\text{com}}]$. If U_m provides communication service, $a_{nm}^{\text{com}} = 1$; otherwise, $a_{nm}^{\text{com}} = 0$. It satisfies

$$\sum_{m=1}^M a_{nm}^{\text{com}} \in \{0, 1\}, n \in \{1, 2, \dots, N\}. \quad (4)$$

The above equation ensures that at most one UAV can provide communication service to a given ground user. Let

$U_n^{\text{com}} \subset \mathcal{U}$ denote the set of UAVs assigned to provide communication service to C_n .

We then use $\psi_{nk}^{\text{com}} \in \Psi_n^{\text{com}}$ as the indicator variables, which indicate whether the communication signal of C_n occupies the k -th sub-channel, and $\Psi_n^{\text{com}} = [\psi_{n1}^{\text{com}} \ \psi_{n2}^{\text{com}} \ \dots \ \psi_{nK}^{\text{com}}]$. If C_n occupies the k -th sub-channel for communication, $\psi_{nk}^{\text{com}} = 1$; otherwise, $\psi_{nk}^{\text{com}} = 0$, which is constrained as follows

$$\sum_{k=1}^K \psi_{nk}^{\text{com}} \in \{0, 1\}. \quad (5)$$

The above constraint indicates that the communication signal can occupy at most one sub-channel.

Accordingly, the communication SINR of U_n^{com} on k -th sub-channel can be expressed as

$$r_{nk}^{\text{com}} = \sum_{m=1}^M \frac{g_n^m \cdot a_{nm}^{\text{com}} \cdot \psi_{nk}^{\text{com}} \cdot p^{\text{com}}}{I_{nk}^{\text{com}} + \sigma^2}, \quad (6)$$

where $\sigma^2 = N_0 \cdot B$ is the power of noise, and p^{com} is the transmit power of communication signal. I_{nk}^{com} represents the interference received by U_n^{com} from the communication signals of other users and all localization signals on k -th sub-channel, which can be expressed as

$$I_{nk}^{\text{com}} = \sum_{l=1, l \neq n}^N g_l^m \cdot p^{\text{com}} \cdot \psi_{lk}^{\text{com}} + \sum_{l=1}^N g_l^m \cdot p^{\text{loc}} \cdot \psi_{lk}^{\text{loc}}, U_m \in U_n^{\text{com}}. \quad (7)$$

Here, the first and second terms in (7) represent the interference caused by other communication signals and localization signals, respectively. p^{loc} denotes the transmit power of the localization signal, and ψ_{lk}^{loc} denotes the indicator variable indicating whether the localization signal of C_l occupies the k -th sub-channel, as discussed in the later section.

Accordingly, the SINR at U_n^{com} for communication can be expressed as

$$r_n^{\text{com}} = \sum_{k=1}^K r_{nk}^{\text{com}}. \quad (8)$$

Then, the throughput from C_n to U_n^{com} is given by

$$R_n = B \log_2(1 + r_n^{\text{com}}). \quad (9)$$

The average system throughput is given by

$$\bar{R} = \frac{1}{N} \sum_{n=1}^N R_n. \quad (10)$$

IV. LOCALIZATION MODELLING AND PROBLEM FORMULATION

A. Localization Modelling

As illustrated in Fig. 3, we develop a localization model tailored for disaster-affected environments where terrestrial communication infrastructure is unavailable, and communication and localization signals can share the same sub-channel. In practice, sub-channel reuse is common due to limited spectrum resources, where transmissions from multiple nodes often share the same sub-channels, incurring mutual interference [8],

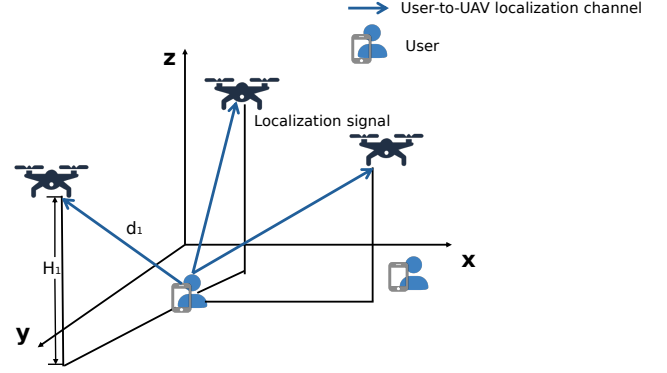


Fig. 3. Localization model

[11], [14]. Prior work mainly allocates separate sub-channels for communication and localization [29], [33], which are more applicable in scenarios with abundant spectrum resources. This localization model is designed for interference-prone emergency rescue scenarios, with the objective of enabling reliable communication and localization services for ground users within the ISAC network.

In the localization model, ground users transmit localization signals to UAVs, which subsequently receive and process these signals to determine the optimized sub-channel allocation, UAV association, and the UAV deployment. Then, the system estimates the users' positions and evaluates the corresponding localization accuracy.

Equipped with high-precision navigation instruments, such as map-referenced multi-sensor navigation systems, UAVs can accurately determine their locations, thereby serving as reliable aerial anchor nodes for localization and localization applications [39], [40]. Therefore, ToA, which is a time-based localization method, is employed to estimate the user positions¹. It offers simplicity, efficiency, and wide applicability, making it suitable for resource-limited scenarios [41]. Specifically, each user's position is determined based on ToA measurements collected from N_{loc} associated UAVs. The corresponding estimation model can be expressed as follows

$$\hat{d}_n^m = c \cdot \tau_{nm} = \sqrt{\|\mathbf{p}_m - \mathbf{q}_n\|^2 + H^2}, \quad (11)$$

where $\hat{d}_n^m$ is the pseudorange between C_n and U_m , and τ_{nm} denotes the time duration for the localization signal transmitted from C_n to the associated $U_m \in U_n^{\text{loc}} \subset \mathcal{U}$. $|U_n^{\text{loc}}| = N_{\text{loc}} \geq 3$ is the set of associated UAVs of C_n , indicating that each user is associated with at least three UAVs for localization [41].

Following the method in [35], [36], the receiver uses a delay-locked loop (DLL) and an early-late discriminator. Under the above conditions, and based on the pseudorange measurement variance given in [42], we derive the pseudorange measurement variance adopted in this paper as (see Appendix for details)

¹It is worth noting that ISAC sensing generally encompasses target detection, environmental perception, parameter estimation, and target tracking. This paper focuses on a ToA-based ground-user localization method.

$$(\sigma_\rho^{nm})^2 = \frac{a_{\text{DLL}} T_0}{2g_n^m p^{\text{loc}} B_{fe}^2} (N_0 B_{fe} + X + Y), \quad (12)$$

where $X = \sum_{n'=1}^N \sum_{k=1}^K g_{n'}^m \psi_{n'k}^{\text{com}} \psi_{nk}^{\text{loc}} p^{\text{com}}$, $Y = \sum_{n'=1, n \neq n'}^N \sum_{k=1}^K g_{n'}^m \psi_{n'k}^{\text{loc}} \psi_{nk}^{\text{loc}} p^{\text{loc}}$. T_0 denotes the duration of the communication signal and the localization signal. a_{DLL} is determined by the relevant parameters of the receiver's loop, and B_{fe} is the double-sided front-end bandwidth. In equation (12), the first item represents the contribution from noise, the second item accounts for interference from communication signals occupying the same sub-channels as the localization signals, and the third item reflects the impact of co-sub-channel transmission from other localization signals overlapping with the current localization signal.

Then the pseudorange measurement standard deviations of C_n are represented by $\sigma_\rho^n = [\sigma_\rho^{n1}, \sigma_\rho^{n2}, \dots, \sigma_\rho^{nmN_{\text{loc}}}]^T$. Accordingly, the localization root mean square error (RMSE) ϵ_n of C_n is given by

$$\epsilon_n = \sqrt{\text{Tr}(H_n D_n (H_n)^T)}, \quad (13)$$

where

$$H_n = [(G_n)^T G_n]^{-1} (G_n)^T, \quad (14)$$

and

$$D_n = \text{diag}((\sigma_\rho^{nm1})^2, (\sigma_\rho^{nm2})^2, \dots, (\sigma_\rho^{nmN_{\text{loc}}})^2). \quad (15)$$

Here, G_n is a matrix determined by the geometric configuration of the localization anchors, and D_n is a diagonal matrix with the variances of pseudorange measurements between C_n and its anchors.

$$G_n = \begin{bmatrix} u_x^{nm1} & u_y^{nm1} \\ u_x^{nm2} & u_y^{nm2} \\ \vdots & \vdots \\ u_x^{nmN_{\text{loc}}} & u_y^{nmN_{\text{loc}}} \end{bmatrix}, \quad (16)$$

and

$$\begin{cases} u_x^{nm_i} = \frac{x_n^q - x_{m_i}^p}{d_n^{m_i}} \\ u_y^{nm_i} = \frac{y_n^q - y_{m_i}^p}{d_n^{m_i}} \end{cases}, \quad (17)$$

where $u_x^{nm_i}$ and $u_y^{nm_i}$ respectively denote the unit vectors of C_n and U_{m_i} along the x and y axes. At a 95% confidence level, the localization accuracy of C_n can be expressed as:

$$\delta_n = \sqrt{\frac{k_{0.95}}{2}} \cdot \epsilon_n \approx 1.731 \epsilon_n, \quad (18)$$

where $k_{0.95}$ denotes the 95% quantile of the chi-square distribution with two degrees of freedom.

Notably, the pseudorange variance model in (12) explicitly couples sub-channel allocation, UAV association and deployment with localization performance. Sub-channel allocation affects the pseudorange variance through the interference terms X and Y , while UAV association and deployment influence the geometry matrix and channel gain. Through (13)-(18), these effects are propagated into localization accuracy,

thereby coupling the optimization variables with localization performance.

We use b_{nm}^{loc} to represent whether U_m is selected to provide localization for C_n , and $\mathcal{B}_n^{\text{loc}} = [b_{n1}^{\text{loc}} \ b_{n2}^{\text{loc}} \ \dots \ b_{nM}^{\text{loc}}]$. If so, $b_{nm}^{\text{loc}} = 1$; otherwise, $b_{nm}^{\text{loc}} = 0$. These variables satisfy the following constraint:

$$\sum_{m=1}^M b_{nm}^{\text{loc}} = N_{\text{loc}}, n \in \{1, 2, \dots, N\}. \quad (19)$$

$\psi_{nk}^{\text{loc}} \in \Psi_n^{\text{loc}}$ are the indicator variables, which indicate whether the localization signal of C_n occupies the k -th sub-channel, and $\Psi_n^{\text{loc}} = [\psi_{n1}^{\text{loc}} \ \psi_{n2}^{\text{loc}} \ \dots \ \psi_{nK}^{\text{loc}}]$. If C_n occupies the k -th sub-channel for localization, the corresponding binary indicator variables $\psi_{nk}^{\text{loc}} = 1$; otherwise, $\psi_{nk}^{\text{loc}} = 0$. These variables are constrained as follows:

$$\sum_{k=1}^K \psi_{nk}^{\text{loc}} \in \{0, 1\}, \quad (20)$$

which indicates that the localization signals can occupy at most one sub-channel.

Then, the SINR of the localization signal at the $U_m \in U_n^{\text{loc}}$ associated with C_n can be expressed as:

$$\gamma_{nm}^{\text{loc}} = \sum_{k=1}^K \frac{g_n^m \cdot b_{nm}^{\text{loc}} \cdot \psi_{nk}^{\text{loc}} \cdot p^{\text{loc}}}{I_{nmk}^{\text{loc}} + \sigma^2}. \quad (21)$$

I_{nmk}^{loc} denotes the interference from the communication signals and other localization signals to the considered localization signal between C_n and U_m . Similarly, I_{nmk}^{loc} also consists of two components, expressed as:

$$I_{nmk}^{\text{loc}} = \sum_{l=1}^N g_l^m \cdot p^{\text{com}} \cdot \psi_{lk}^{\text{com}} + \sum_{l=1, l \neq n}^N g_l^m \cdot p^{\text{loc}} \cdot \psi_{lk}^{\text{loc}}. \quad (22)$$

Here, the former term represents the interference caused by communication signals, while the latter term corresponds to the interference induced by the other localization signals.

Building upon the analyses in this section and Section III-C, the mutual interference caused by sub-channel allocation impacts pseudorange measurements variance, while UAV association for anchor selection and UAV deployment influence the channel gain and geometry matrix. These factors jointly determine the system communication and localization performance. Therefore, joint optimization of sub-channel allocation, UAV association, and UAV deployment is necessary to provide efficient communication and reliable localization services.

B. Problem Formulation

In this paper, we aim to maximize the average system throughput by jointly optimizing the UAV association and sub-channel allocation for communication and localization, and UAV deployment, while ensuring the localization accuracy. The optimization problem is formulated as follows

$$\mathcal{P} : \max_{\mathcal{A}_n^{\text{com}}, \mathcal{B}_n^{\text{loc}}, \Psi_n^{\text{com}}, \Psi_n^{\text{loc}}, \mathbf{P}} \bar{R}, \quad (23)$$

- s.t. **C1:** $\delta_n \leq \delta_r$,
- C2:** $\sum_{m=1}^M a_{nm}^{\text{com}} \in \{0, 1\}$,
- C3:** $\sum_{m=1}^M b_{nm}^{\text{loc}} = N_{\text{loc}}$,
- C4:** $\sum_{k=1}^K \psi_{nk}^{\text{com}} \in \{0, 1\}$,
- C5:** $\sum_{k=1}^K \psi_{nk}^{\text{loc}} \in \{0, 1\}$,
- C6:** $r_{nm}^{\text{loc}} \geq r_{\text{th}}^{\text{loc}}$,
- C7:** $r_n^{\text{com}} \geq r_{\text{th}}^{\text{com}}$,
- C8:** $N_{\text{loc}} \geq 3$,
- C9:** $\mathbf{P} \in \mathbb{R}^{M \times 2}, x_{\min} \leq x_m^p \leq x_{\max}, y_{\min} \leq y_m^p \leq y_{\max}$,
- C10:** $a_{nm}^{\text{com}}, b_{nm}^{\text{loc}}, \psi_{nk}^{\text{com}}, \psi_{nk}^{\text{loc}} \in \{0, 1\}$,
 $\forall n \in \{1, 2, \dots, N\}$.

In problem $\mathcal{P}$, constraint **C1** ensures that the localization accuracy, jointly determined by sub-channel allocation, UAV association and deployment, remains below the predefined threshold, thereby guaranteeing reliable localization service. **C2** and **C3** specify the communication and localization association constraints, respectively, i.e., each user is associated with at most one UAV for communication and exactly N_{loc} UAVs for localization. **C4** and **C5** indicate that both communication signal and localization signal can occupy only one sub-channel for transmission. Constraints **C6** and **C7** respectively ensure that the SINR of localization signals and communication signals received at the UAVs must meet or exceed a predefined threshold, thereby guaranteeing reliable demodulation and successful decoding. **C8** guarantees that each user is associated with at least three UAV anchors for localization to satisfy the fundamental geometric requirement. The constraint **C9** specifies that the optimized UAV coordinates are defined in the two-dimensional Cartesian space and constrained within the feasible deployment region bounded by $x_{\min}$, $x_{\max}$, $y_{\min}$, and $y_{\max}$. Finally, **C10** defines the binary restrictions of the decision variables.

Due to the strong coupling among sub-channel allocation, UAV association and deployment in both the objective and constraints, as well as the binary variables a_{nm}^{com} , b_{nm}^{loc} , ψ_{nk}^{com} , and ψ_{nk}^{loc} , problem $\mathcal{P}$ is a MINNLP. These variables jointly affect the throughput objective and constraints **C1**–**C10** on localization accuracy, association, sub-channel allocation, and communication and localization SINR, making the problem difficult to solve by conventional mathematical methods. To obtain a satisfactory high quality solution, we propose an alternating iterative optimization method and decompose the original problem into two subproblems: UAV association and sub-channel allocation (subproblem 1), and UAV deployment optimization (subproblem 2). The detailed method is provided in Section V.

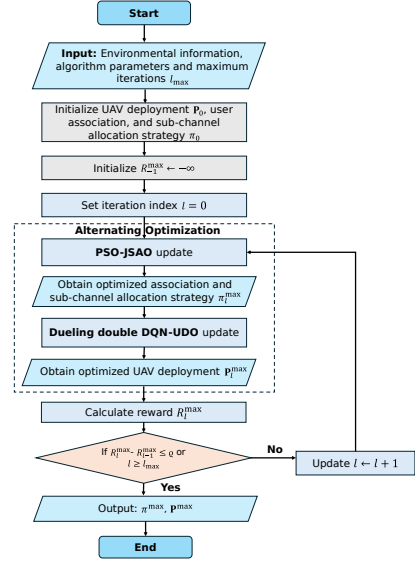


Fig. 4. The framework of the CAIO method

V. OPTIMIZATION METHOD

In this section, we propose a collaborative alternating iterative optimization (CAIO) method that integrates PSO and duelling double DQN to jointly optimize sub-channel allocation, UAV association, and deployment, thereby both enhancing average system throughput and ensuring localization accuracy. The CAIO method operates via an iterative mechanism comprising two complementary methods: a particle swarm optimization-based joint spectrum and UAV association (PSO-JSAO) method, and a duelling double DQN-based UAV deployment optimization (duelling double DQN-UDO) method. The PSO-JSAO method addresses the discrete optimization of sub-channel allocation and UAV association by constructing a utility function with penalty factors, maximizing average system throughput while ensuring localization accuracy under fixed UAV deployment. Building upon this, the duelling double DQN-UDO method uses the PSO-JSAO output as its initial state and leverages a difference-based reward function to guide UAV deployment optimization. Through alternating iterative updates, these two methods collaboratively optimize sub-channel allocation, UAV association, and UAV deployment, ensuring efficient communication and reliable localization. The overall framework of the proposed CAIO method is illustrated in Fig. 4.

A. CAIO Method

In the CAIO method, each outer iteration performs one PSO-JSAO update and one duelling double DQN-UDO update, which enables timely coordination, avoids over-optimization of outdated intermediate variables, and balances solution quality with computational complexity. At the beginning, the UAV association, sub-channel allocation for both communication and localization and UAV positions $\mathbf{P}_0$ are initialized. During each outer iteration, given the current UAV positions $\mathbf{P}_l$, the PSO-JSAO method is executed to obtain the optimized UAV association and sub-channel allocation solution π_l^{max} ,

Algorithm 1 CAIO Method

Input: Maximum number of outer iterations $l_{\max}$, parameters of PSO-JSAO and dueling double DQN-UDO.

Output: Optimized sub-channel allocation, association solution $\pi^{\max}$ and UAV positions $\mathbf{P}^{\max}$.

- 1: Randomly initialize UAV positions $\mathbf{P}_0$.
 - 2: Initialize $R_{-1}^{\max} \leftarrow -\infty$.
 - 3: **for** $l = 0, 1, \dots, l_{\max} - 1$ **do**
 - 4: Obtain $\pi_l^{\max}$ by **Algorithm 2** under $\mathbf{P}_l$ and calculate R_l^0 .
 - 5: Obtain $\mathbf{P}_l^{\max}$ by **Algorithm 3** under $\pi_l^{\max}$ and $\mathbf{P}_l$.
 - 6: Calculate $R_l^{\max}$ from (41) and (42) under $\pi_l^{\max}$ and $\mathbf{P}_l^{\max}$.
 - 7: Update $\pi_l^{\max} \leftarrow \pi_l^{\max}$ and $\mathbf{P}^{\max} \leftarrow \mathbf{P}_l^{\max}$.
 - 8: $\mathbf{P}_{l+1} \leftarrow \mathbf{P}_l^{\max}$.
 - 9: **if** $l > 0$ and $|R_l^{\max} - R_{l-1}^{\max}| \leq \varrho$ **then**
 - 10: **break**
 - 11: **return** $\pi^{\max}$ and $\mathbf{P}^{\max}$.
-

and the corresponding initial reward R_l^0 is then calculated. Subsequently, dueling double DQN-UDO is employed to optimize UAV deployment under $\pi_l^{\max}$, thereby obtaining the optimized UAV positions $\mathbf{P}_l^{\max}$. This alternating iterative process continues until the convergence criterion is satisfied, at which point the final sub-channel, UAV association and deployment solution is obtained.

B. PSO-JSAO Method

In this section, we propose the PSO-JSAO method for joint sub-channel allocation and UAV association in both the communication and localization domains. By leveraging PSO, the method efficiently explores the solution space and achieves a balance between convergence speed and solution quality, thereby maximizing the average system throughput while ensuring localization accuracy.

PSO is a population-based optimization method inspired by the foraging behavior of bird flocks and fish schools. In PSO, each candidate solution is represented as a particle, and all particles collectively form a swarm.

We define the position of the i -th particle at PSO iteration ι as:

$$\mathbf{X}_{i,\iota} = \{X_{i,\iota}^{(\text{ca})}, X_{i,\iota}^{(\text{pa})}, X_{i,\iota}^{(\text{cs})}, X_{i,\iota}^{(\text{ps})}\}, \quad i \in \{1, 2, \dots, N_{\text{particles}}\}, \quad (24)$$

where $X_{i,\iota}^{(\text{ca})}$, $X_{i,\iota}^{(\text{pa})}$, $X_{i,\iota}^{(\text{cs})}$ and $X_{i,\iota}^{(\text{ps})}$ represent the communication UAV association, localization UAV association, communication sub-channel allocation and localization sub-channel allocation, respectively². $N_{\text{particles}}$ is the number of particles. Specifically,

$$X_{i,\iota}^{(\text{ca})} = \{x_{1,\iota}^{(\text{ca})}, x_{2,\iota}^{(\text{ca})}, \dots, x_{N,\iota}^{(\text{ca})}\}, x_{n,\iota}^{(\text{ca})} \in \{1, 2, \dots, M\}. \quad (25)$$

$$X_{i,\iota}^{(\text{pa})} = \{x_{1,\iota}^{(\text{pa})}, x_{2,\iota}^{(\text{pa})}, \dots, x_{N,\iota}^{(\text{pa})}\},$$

²For notational consistency, the superscripts (ca), (pa), (cs), and (ps) denote the variable categories, while the subscripts i and ι denote the particle index and the PSO iteration index, respectively.

$$\begin{aligned} x_{n,\iota}^{(\text{pa})} &= \{x_{n1,\iota}^{(\text{pa})}, x_{n2,\iota}^{(\text{pa})}, \dots, x_{nN_{\text{loc}},\iota}^{(\text{pa})}\}, x_{ni,\iota}^{(\text{pa})} \in \{1, 2, \dots, M\}. \\ x_{nj,\iota}^{(\text{pa})} &\neq x_{nj',\iota}^{(\text{pa})}, j, j' \in \{1, \dots, N_{\text{loc}}\}, j \neq j' \end{aligned} \quad (26)$$

$$X_{i,\iota}^{(\text{cs})} = \{x_{1,\iota}^{(\text{cs})}, x_{2,\iota}^{(\text{cs})}, \dots, x_{N,\iota}^{(\text{cs})}\}, x_{n,\iota}^{(\text{cs})} \in \{1, 2, \dots, K\}. \quad (27)$$

$$X_{i,\iota}^{(\text{ps})} = \{x_{1,\iota}^{(\text{ps})}, x_{2,\iota}^{(\text{ps})}, \dots, x_{N,\iota}^{(\text{ps})}\}, x_{n,\iota}^{(\text{ps})} \in \{1, 2, \dots, K\}. \quad (28)$$

Accordingly, the velocity of the i -th particle at iteration ι is defined as:

$$\mathbf{V}_{i,\iota} = \{V_{i,\iota}^{(\text{ca})}, V_{i,\iota}^{(\text{pa})}, V_{i,\iota}^{(\text{cs})}, V_{i,\iota}^{(\text{ps})}\}. \quad (29)$$

Based on the current position and velocity, each particle in the swarm updates its position at the next iteration. The velocity update formula is given by:

$$\mathbf{V}_{i,\iota+1} = \omega \mathbf{V}_{i,\iota} + c_1 \xi_1 (\mathbf{X}_{i,\text{pbest},\iota} - \mathbf{X}_{i,\iota}) + c_2 \xi_2 (\mathbf{X}_{\text{gbest},\iota} - \mathbf{X}_{i,\iota}), \quad (30)$$

where $\mathbf{X}_{i,\text{pbest},\iota}$ and $\mathbf{X}_{\text{gbest},\iota}$ denote the personal-best position of the i -th particle and the global-best position of the swarm at iteration ι , respectively. ω denotes the inertia weight, c_1 and c_2 represent the cognitive and social learning factors, respectively, which control the influence of the particle's own best experience and the swarm's global best position. $\xi_1, \xi_2 \sim U(0, 1)$ are uniformly distributed random variables, introduced to enhance the stochasticity of the algorithm and help avoid premature convergence to local optima.

According to the updated velocity, we get the position of the particles at the next iteration:

$$\mathbf{X}_{i,\iota+1} = \text{clip}(\lfloor \mathbf{X}_{i,\iota} + \mathbf{V}_{i,\iota} \rfloor). \quad (31)$$

where $\lfloor \cdot \rfloor$ is the floor function, ensuring that each element in the particle is discrete. And the $\text{clip}(\cdot)$ is the clipping function, preventing the particle from exceeding the bounds.

According to problem $\mathcal{P}$, the optimization objective of this paper is to maximize the average system throughput while ensuring localization accuracy. Based on this objective, we design the fitness function of the PSO algorithm as follows:

$$F(\mathbf{X}_{i,\iota}) = U(\mathbf{X}_{i,\iota}) + \Gamma(\mathbf{X}_{i,\iota}), \quad (32)$$

where $U(\mathbf{X}_{i,\iota}) = \bar{R}$ is the utility function value of the i -th particle, which is positively correlated with the average system throughput. And $\Gamma(\mathbf{X}_{i,\iota})$ represents the penalty function, which is given by

$$\Gamma(\mathbf{X}_{i,\iota}) = \eta \cdot \sum_{n=1}^N \mathbf{1}(\delta_n > \delta_r) = \eta \cdot I_{\text{npos}}, \quad (33)$$

where η is a penalty coefficient, and I_{npos} denotes the number of users whose localization accuracy requirements are not satisfied.

The pseudocode of the PSO-JSAO method is provided in **Algorithm 2**. Initially, $N_{\text{particles}}$ particles are randomly generated, where each particle is characterized by position and velocity vectors consisting of four discrete components that

Algorithm 2 PSO-JSAO

Input: Current UAV positions $\mathbf{P}_l$, numbers of UAVs M , users N , sub-channels K , localization UAVs N_{loc} , particles $N_{\text{particles}}$, and PSO iterations N_{it} .

Output: Optimized sub-channel allocation and UAV association solution $\pi_l^{\max}$.

- 1: Randomly initialize the swarm $\{\mathbf{X}_{i,0}, \mathbf{V}_{i,0}\}_{i=1}^{N_{\text{particles}}}$.
- 2: Initialize $\mathbf{X}_{i,\text{pbest},0} \leftarrow \mathbf{X}_{i,0}$ for all particles and $\mathbf{X}_{\text{gbest},0}$ from the best particle in the PSO swarm.
- 3: **for** $\iota = 0, 1, \dots, N_{\text{it}} - 1$ **do**
- 4: **for** each particle $\mathbf{X}_{i,\iota}$ **do**
- 5: Calculate fitness value $F(\mathbf{X}_{i,\iota})$.
- 6: **if** $F(\mathbf{X}_{i,\iota}) > F(\mathbf{X}_{i,\text{pbest},\iota})$ **then**
- 7: $\mathbf{X}_{i,\text{pbest},\iota+1} \leftarrow \mathbf{X}_{i,\iota}$
- 8: **else**
- 9: $\mathbf{X}_{i,\text{pbest},\iota+1} \leftarrow \mathbf{X}_{i,\text{pbest},\iota}$
- 10: **if** $F(\mathbf{X}_{i,\text{pbest},\iota+1}) > F(\mathbf{X}_{\text{gbest},\iota})$ **then**
- 11: $\mathbf{X}_{\text{gbest},\iota+1} \leftarrow \mathbf{X}_{i,\text{pbest},\iota+1}$
- 12: **else**
- 13: $\mathbf{X}_{\text{gbest},\iota+1} \leftarrow \mathbf{X}_{\text{gbest},\iota}$
- 14: Update $\mathbf{V}_{i,\iota+1}$ and $\mathbf{X}_{i,\iota+1}$ using (30) and (31).
- 15: $\pi_l^{\max} \leftarrow X_{\text{gbest},N_{\text{it}}}$
- 16: **return** $\pi_l^{\max}$

jointly encode UAV associations and sub-channel allocations for communication and localization. The fitness value of each particle is then calculated. At each iteration, the velocity of each particle is updated according to its personal best position and the global best position. The particle's position is subsequently adjusted using the updated velocity, and its fitness value is re-evaluated. The personal best and global best values are updated accordingly. This iterative process continues until the maximum number of iterations is reached. Finally, the method outputs the optimized UAV association and sub-channel allocation solution for both communication and localization under fixed UAV deployment.

C. Dueling double DQN-UDO Method

In this section, we propose a UAV deployment method based on dueling double DQN to maximize the objective of problem $\mathcal{P}$. The adopted dueling double DQN combines the advantages of dueling DQN and double DQN. Specifically, dueling DQN decomposes the Q-function into the state-value function $V(S)$ and the advantage function $A(S, A)$, improving state-value discrimination and learning efficiency [6]. Meanwhile, double DQN mitigates Q-value overestimation through a dual-network architecture, enhancing the stability and accuracy of policy learning [6]. Therefore, the proposed method not only overcomes the local optima limitations of PSO but also significantly improves sample efficiency and policy generalization. As a result, it facilitates more robust and adaptive deployment of UAVs, thereby enhancing both average system throughput and localization accuracy for ground users.

Firstly, we demonstrate the design and definition of the state, action, and reward of the dueling double DQN.

State Space (S): The state space consists of all feasible UAV position vectors, and the current state is defined as

$$S_t = [\mathbf{p}_1(t), \mathbf{p}_2(t), \dots, \mathbf{p}_M(t)], \quad (34)$$

where $\mathbf{p}_m(t)$ is the position of U_m at step t .

Action Space (A): The action space is defined to determine both the UAV to be adjusted and its movement direction. At each step, the agent selects one UAV and determines the movement direction of this selected UAV, while the positions of the other UAVs remain unchanged. Therefore, the action at step t is denoted by

$$A_t = (m_t, o_t), \quad (35)$$

where $m_t \in \{1, 2, \dots, M\}$ denotes the index of the selected UAV, and $o_t \in \{0, 1, \dots, O\}$ determines the movement direction. Accordingly, the action space can be expressed as

$$\mathcal{A} = \{(m, o) \mid m \in \{1, 2, \dots, M\}, o \in \{0, 1, \dots, O\}\}. \quad (36)$$

For the selected U_{m_t} , its horizontal position is updated as

$$x_{m_t}^p(t+1) = \begin{cases} \text{clip}(x_{m_t}^p(t) + \Delta \cos \theta_{o_t}), & o_t \neq 0, \\ x_{m_t}^p(t), & o_t = 0, \end{cases} \quad (37)$$

$$y_{m_t}^p(t+1) = \begin{cases} \text{clip}(y_{m_t}^p(t) + \Delta \sin \theta_{o_t}), & o_t \neq 0, \\ y_{m_t}^p(t), & o_t = 0. \end{cases} \quad (38)$$

where Δ denotes the UAV movement step size, θ_{o_t} is the horizontal movement direction of the selected UAV. Specifically,

$$\theta_{o_t} = \frac{2\pi(o_t - 1)}{O} \quad (39)$$

where $o_t \in \{1, 2, \dots, O\}$, and $o_t = 0$ indicates that the selected UAV remains stationary. The $\text{clip}(\cdot)$ operation projects the updated coordinate onto the feasible deployment region. For any U_m with $m \neq m_t$, its position remains unchanged,

$$\mathbf{p}_m(t+1) = \mathbf{p}_m(t), \quad m \neq m_t. \quad (40)$$

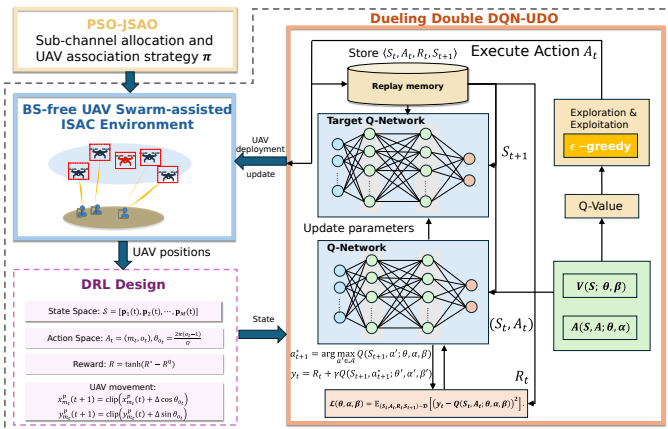


Fig. 5. The dueling double DQN-UDO method.

Reward (R): The reward represents the feedback received by the agent upon taking a specific action in the current environment. To effectively translate the optimization objective into the reward expression, the following considerations are taken into account:

- 1) In problem $\mathcal{P}$, localization accuracy is imposed as a strict constraint. Therefore, to ensure that the localization requirements of users are satisfied, the reward function should be designed to exhibit a negative correlation with the number of users that violate the localization accuracy threshold.
- 2) Since the primary objective of the optimization is to maximize the average system throughput, the reward should exhibit a positive correlation with the average system throughput when the localization accuracy constraints for all users are satisfied.
- 3) To improve the learning efficiency and accelerate the convergence of the dueling double DQN algorithm, the reward function is normalized within a predefined bounded range.
- 4) Given that the DRL-based UAV position optimization is performed based on the initial solution obtained from PSO, the reward should also incorporate the performance baseline provided by the PSO solution.

Taking into account the constraints and optimization goals discussed in 1), 2), and 3), we formulate the following piecewise-defined reward function to guide the learning process effectively.

$$R^* = \begin{cases} \frac{1}{1 + e^{\vartheta \cdot \bar{R} + \varpi}}, & I_{\text{npos}} = 0, \\ r_{I_{\text{npos}}}, & 1 \leq I_{\text{npos}} \leq 4, \\ -1, & \text{otherwise.} \end{cases} \quad (41)$$

Let $r_{I_{\text{npos}}} \in \{r_1, r_2, r_3, r_4\}$ denote the reward coefficients corresponding to different values of I_{npos} , respectively. ϑ and ϖ are the scaling and offset coefficients, respectively, which control its slope and shift. Based on the above considerations and taking into account case 4), we design the final reward function as:

$$R = \tanh(R^* - R^0), \quad (42)$$

where R^0 represents the initial reward, which is obtained from the solution provided by the current PSO-JSAO and the current position, calculated according to equation (41). And $\tanh(\cdot)$ is used to restrict the reward value within the range $[-1, 1]$.

In dueling double DQN, the Q-network is divided into two parts:

$$Q(S_t, A_t; \theta, \alpha, \beta) = V(S_t; \theta, \beta) + A(S_t, A_t; \theta, \alpha) - \frac{1}{|\mathcal{A}|} \sum_{a' \in \mathcal{A}} A(S_t, a'; \theta, \alpha). \quad (43)$$

Here, $V(S_t; \theta, \beta)$ denotes the value function and $A(S_t, A_t; \theta, \alpha)$ denotes the advantage function, θ is the network parameters of the layers before the output layer, and α and β are the parameters of the two streams of fully-connected layers.

At each step, the agent obtains the current network state from the environment, denoted as S_t . Then the agent selects the current action corresponding to the ϵ -greedy strategy, which is expressed as:

$$A_t = \begin{cases} \arg \max_{a' \in \mathcal{A}} Q(S_t, a'; \theta, \alpha, \beta), & \text{with probability } 1 - \epsilon, \\ \text{random action}, & \text{with probability } \epsilon. \end{cases} \quad (44)$$

Then A_t is applied to the UAV-assisted ISAC network, resulting in R_t and S_{t+1} , and the tuple $\langle S_t, A_t, R_t, S_{t+1} \rangle$ is subsequently stored in the experience replay buffer $\mathcal{D}$. Once $\mathcal{D}$ accumulates sufficient experiences, the agent randomly selects a batch of samples to compute the target Q-value:

$$y_t = R_t + \gamma Q \left(S_{t+1}, \arg \max_{a' \in \mathcal{A}} Q(S_{t+1}, a'; \theta, \alpha, \beta); \theta', \alpha', \beta' \right). \quad (45)$$

where $(\theta', \alpha', \beta')$ denote the parameters of the target network and γ is the discount factor. The loss function of dueling double DQN is defined as:

$$\mathcal{L}(\theta, \alpha, \beta) = \mathbb{E}_{(S_t, A_t, R_t, S_{t+1}) \sim \mathcal{D}} \left[(y_t - Q(S_t, A_t; \theta, \alpha, \beta))^2 \right]. \quad (46)$$

Moreover, network parameters θ, α, β are updated periodically after a certain number of deployments. Given that the neural network parameters θ, α, β are randomly initialized, it is essential to encourage sufficient exploration of the action

Algorithm 3 Dueling double DQN-UDO

Input: Current solution $\pi_l^{\max}$, UAV positions $\mathbf{P}_l$ at iteration l , initial reward R_l^0 , number of episodes and steps $EP, STEP$, replay buffer capacity $D_{\max}$, batch size N_b , and target-network update frequency N_u .

Output: Optimized UAV positions $\mathbf{P}_l^{\max}$.

- 1: Initialize Q-network with (θ, α, β) and the target Q-network with $(\theta', \alpha', \beta')$.
 - 2: Set $(\theta', \alpha', \beta') \leftarrow (\theta, \alpha, \beta)$, global_step $\leftarrow 0$, and replay buffer $\mathcal{D}$.
 - 3: Initialize $\mathbf{P}_l^{\max} \leftarrow \mathbf{P}_l$, $R_l^{\max} \leftarrow R_l^0$.
 - 4: **for** $ep = 1, 2, \dots, EP$ **do**
 - 5: Reset state S_t based on $\mathbf{P}_l$.
 - 6: **for** $t = 1, 2, \dots, STEP$ **do**
 - 7: Select action A_t by the ϵ -greedy policy.
 - 8: Execute A_t , obtain next state S_{t+1} , and calculate the reward R_t by (42).
 - 9: Store $\langle S_t, A_t, R_t, S_{t+1} \rangle$ in $\mathcal{D}$.
 - 10: **if** global_step $\geq D_{\max}$ **then**
 - 11: Sample a mini-batch from $\mathcal{D}$ and update (θ, α, β) .
 - 12: **if** global_step mod $N_u = 0$ **then**
 - 13: $(\theta', \alpha', \beta') \leftarrow (\theta, \alpha, \beta)$.
 - 14: $S_t \leftarrow S_{t+1}$, global_step $\leftarrow$ global_step + 1.
 - 15: **if** $R_t > R_l^{\max}$ **then**
 - 16: $R_l^{\max} \leftarrow R_t$ and update $\mathbf{P}_l^{\max}$.
 - 17: **return** $\mathbf{P}_l^{\max}$.
-

during the initial training phase. To this end, we adjust the exploration rate across episodes to achieve a balance between exploration and exploitation:

$$\epsilon = \max\{0.01, (1 - \frac{episode}{500})\}, \quad (47)$$

where *episode* denotes the current training episode.

The pseudocode of dueling double DQN-UDO method is presented in **Algorithm 3**. Firstly, the dueling double DQN-UDO initializes the Q-network and target Q-network parameters and other parameters. Then, the initial reward R^0 is calculated based on the current UAV positions as well as the given sub-channel allocation and UAV association solution. For each *episode*, the UAV environment is reset based on the initial positions. During each step of an *episode*, Q-values for all possible actions are obtained from the Q-network, and an ϵ -greedy strategy is applied to determine the movement directions of the UAVs. The average system throughput and localization accuracy are then evaluated under the current UAV deployment, and a corresponding reward R is computed. The resulting experience tuple, consisting of the current state, action, reward, and next state, is stored in the replay buffer. Once sufficient data has been collected, mini-batches are sampled from the buffer to update the Q-network parameters, and the target network is periodically synchronized. Upon completion of all training iterations, the final state is decoded into UAV coordinates, and the optimized UAV deployment is derived.

Remark 1: CAIO combines the global search capability of heuristic optimization with the dynamic decision-making ability of DRL, and iteratively converges to a stable suboptimal solution since the objective value is non-decreasing and upper bounded by finite system resources [43]. This design mitigates the dimensionality explosion and slow convergence issues of multi-agent reinforcement learning and other single-stage methods.

Remark 2: The proposed method shows strong potential for practical deployment in emergency rescue scenarios. UAV swarms serve as supplementary devices to replace damaged ground BSs, enabling efficient communication and localization. To reduce time and energy costs, an offline training and online deployment strategy is used. Additionally, by utilizing fixed-wing UAVs with greater endurance and deploying the algorithm on the energy-rich head UAV, task execution time and service stability are significantly improved. DRL demonstrates adaptability in complex environments, while the ToA method's simplicity and efficiency have led to its widespread real-world use. Overall, the method offers high adaptability, low energy consumption, and reduced deployment costs, making it promising for emergency rescue scenarios.

D. Complexity Analysis

1) *Time Complexity:* For the PSO-JSAO method, the time complexity is $O(N_{\text{particles}} \times N_{\text{it}})$. For the dueling double DQN-UDO method, the time complexity is $O((EP \times STEP) \times N_b \times F(\Theta))$, where Θ represents all learnable parameters of the dueling double DQN, and $F(\Theta)$ denotes the floating point

TABLE III
KEY SIMULATION PARAMETERS

Parameters	Value
Deployment region of UAVs and users	[600 m, 600 m]
The initial altitude of the UAVs	60 m
Number of sub-channels K	[6, 50]
Frequency f	2 GHz
Number of localization UAVs per user N_{loc}	3,4,5
Power of communication and localization $p^{\text{com}}, p^{\text{loc}}$	1 W, 0.1 W
Bandwidth B	100 kHz
Channel parameters $\alpha, \mathbf{b}$	9.61, 0.16 [37]
Attenuation factors $\mu^{\text{LoS}}, \mu^{\text{NLoS}}$	1, 20 [37]
localization RMSE and accuracy ϵ_r / δ_r	1 m / 1.6 m [29]
UAV step size Δ	10 m
SINR threshold $r_{\text{th}}^{\text{com}}, r_{\text{th}}^{\text{loc}}$	3 dB [37]
Maximum number of outer iterations l_{max} and stopping threshold ϱ	10, 0.001
PSO parameters $N_{\text{it}}, N_{\text{particles}}, \omega, c_1, c_2, V_{\text{max}}$	100, 200, 0.6, 1.5, 1.5, 2
Penalty coefficient η	-10^9
The replay buffer capacity D_{max}	50000
Batch size N_b	256
Learning rate	0.0001
Discount factor γ	0.9
Numbers of episodes and steps $EP, STEP$	700, 200
Number of UAVs movement directions O	4
Reward coefficients r_1, r_2, r_3, r_4	-0.4, -0.6, -0.8, -0.9
Scaling and offset coefficients of reward ϑ, ϖ	$-2.4812 \times 10^{-6}, 5.8123$
Front-end bandwidth and DLL coefficient B_{fe}, a_{DLL}	200 kHz [35], 0.2 [36]
hidden layer	[128, 64]
Optimizer of DRL	Adam optimizer
Activation function of DRL	ReLU

operations required for a single forward pass of the dueling double DQN. Therefore, the overall time complexity of the CAIO method can be expressed as $O(l_{\text{max}} \times (N_{\text{it}} \times N_{\text{particles}} + EP \times STEP \times N_b \times F(\Theta)))$.

2) *Space Complexity:* For the PSO-JSAO method, the space complexity is $O(N_{\text{particles}} \times N)$. For the dueling double DQN-UDO method, the space complexity is $O(\Theta) + O(D_{\text{max}} \times (|S| + |A| + 1)) \approx O(\Theta) + O(D_{\text{max}} \times |S|) = O(\Theta + D_{\text{max}} M)$. Therefore, the overall space complexity of the CAIO method is $O(N_{\text{particles}} \times N + \Theta + D_{\text{max}} \times M)$.

VI. SIMULATION RESULTS

A. Simulation Settings

We conduct simulation studies for the proposed algorithm on a Debian-based computer equipped with a 5.70 GHz Intel i9 CPU, 64 GB of RAM, and an NVIDIA GeForce RTX 4090 D GPU with 24 GB of VRAM. The code is publicly available at [44]. The simulation parameters are summarized in Table III. In particular, the SINR threshold $r_{\text{th}}^{\text{com}} = r_{\text{th}}^{\text{loc}} = 3$ dB, while the RMSE and localization accuracy thresholds are set to 1 and 1.6, respectively. The reward coefficients $r_1 = -0.4, r_2 = -0.6, r_3 = -0.8, r_4 = -0.9$, and the learning rate is 10^{-4} . For the PSO, $\omega = 0.6, c_1 = c_2 = 1.5$, and $V_{\text{max}}^{(\text{ca})} = V_{\text{max}}^{(\text{pa})} = V_{\text{max}}^{(\text{cs})} = V_{\text{max}}^{(\text{ps})} = V_{\text{max}} = 2$. The adopted neural network consists of two hidden layers with 128 and 64 hidden units, respectively, using ReLU activation and Adam

optimization. Different network scales with varying numbers of UAVs and users are considered to validate the robustness and stability of the proposed algorithm, namely Scenario 1, Scenario 2, Scenario 3 and Scenario 4, corresponding to user-UAV configurations of [8, 12], [12, 15], [18, 22] and [60, 30], respectively. For Scenario 4, $B = 200$ kHz is used for easier observation, while for all other scenarios $B = 100$ kHz. The number of sub-channels varies between 6 and 50, with different ranges used for each scenario in the result figures to aid analysis. The initial positions of the UAVs are located within a $600 \text{ m} \times 600 \text{ m}$ area, with an initial altitude of 60 m unless otherwise specified. Each UAV moves horizontally according to the selected action with a fixed step size of $\Delta = 10 \text{ m}$, and any updated position outside the deployment region is clipped to the nearest feasible boundary point.

For the sake of brevity in the figure legend, dueling double DQN is denoted as tDQN in the figure. To validate the effectiveness of the proposed method, we compare it with the following baselines:

- To evaluate PSO-JSAO, DE and ant colony optimization (ACO) are used as baselines for sub-channel allocation and UAV association, while tDQN is used for UAV deployment optimization. The resulting methods are denoted as PSO-tDQN, DE-tDQN, and ACO-tDQN, respectively.
- To assess the effectiveness of UAV deployment, a PSO-based method and multi-agent DQN (MADQN) are employed as the benchmarks, in which PSO and DQN are utilized for sub-channel allocation, UAV association and UAV deployment, respectively. These methods are denoted as PSO-SAD and MADQN-CAIO in the result figures.
- To demonstrate the generality of the proposed method, we further extend the CAIO to include altitude optimization, referred to as 3D-CAIO, with altitude ranging from 30 m to 100 m. We design the action space³ and height out-of-bounds penalty⁴ to guide the learning process.

B. Fairness of Baseline Hyperparameter Settings

To ensure fairness, we further compare the baselines under the same network settings with multiple hyperparameter groups. The results in Fig. 6a and Fig. 6b show that CAIO and PSO still achieve the best performance. The best-performing settings are adopted in the subsequent experiments, and the detailed parameter groups are provided in Appendix B.

C. Convergence of the Proposed Method

We investigate the convergence behavior of the CAIO over 200k training steps. The four curves in Fig. 7 correspond to the cases of $N_{\text{loc}} = 3, K = 14$ in Scenario 1 (red), $N_{\text{loc}} = 3, K = 21$ in Scenario 2 (blue), $N_{\text{loc}} = 3, K = 28$

³Redefine $o_t \in \{0, 1, \dots, O, O+1, O+2\}$, $o_t = 0$ indicates that the selected UAV remains stationary; $o_t \in \{1, 2, \dots, O\}$ indicates horizontal movement with azimuth angle $\theta_{o_t} = 2\pi(o_t - 1)/O$; $o_t = O+1$ and $o_t = O+2$ indicate upward and downward vertical movements, respectively.

⁴In (41), I_h is defined as the number of UAVs exceeding the height limit. If $I_h \neq 0$, then $R^* = -I_h$.

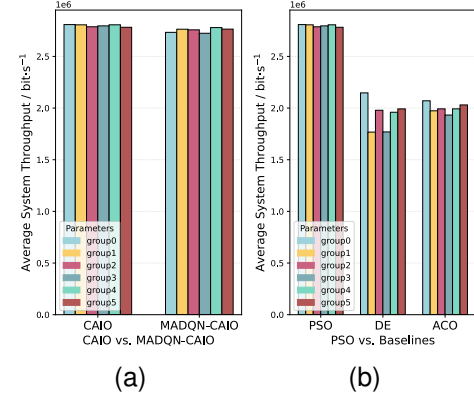


Fig. 6. Fairness evaluation of baseline hyperparameter settings. (a) CAIO and MADQN-CAIO under different parameter groups. (b) PSO, DE, and ACO under different parameter groups.

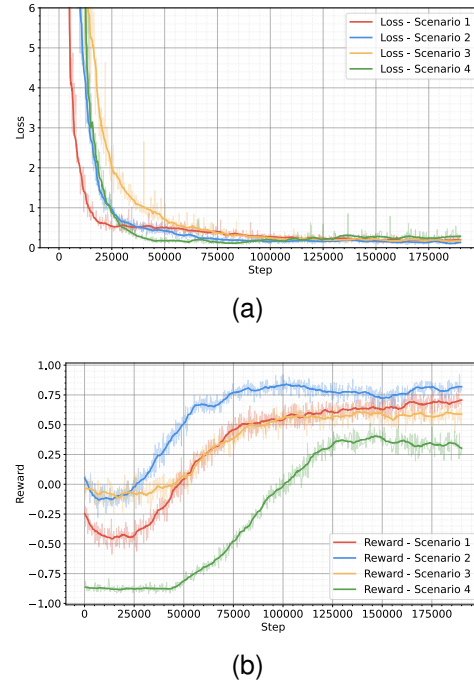


Fig. 7. Convergence performance of the CAIO. (a) Loss. (b) Reward.

in Scenario 3 (yellow), and $N_{\text{loc}} = 3, K = 50$ in Scenario 4 (green). As shown in Fig. 7, the Q-network loss decreases rapidly within the first 50k steps, while the reward increases steadily and converges around 100k steps, indicating fast policy learning, stable value estimation, and high sample efficiency of CAIO. We further record the memory usage, training time, and inference time for each network scenario, with the corresponding results summarized in Table IV, where the reported inference time is the total time measured over 100 episodes. As the network scale and the number of sub-channels increases, training time, inference time and memory usage grow accordingly, which is consistent with the complexity analysis in Section V-D. Notably, Scenario 4 considers a large-scale network with 60 users and 30 UAVs. As shown in Fig. 7 and Table IV, CAIO still converges stably in large-scale network settings, demonstrating the scalability of the proposed

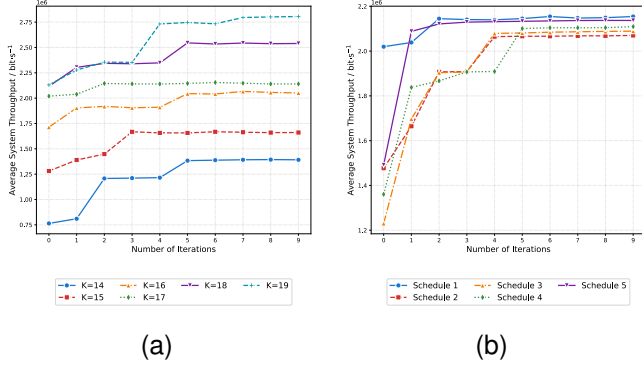


Fig. 8. Average system throughput versus the number of outer iterations for the CAIO (network Scenario 2, $N_{loc} = 4$). (a) Performance under different sub-channel settings. (b) Performance under different alternation schedules between PSO-JSAO and tDQN-UDO.

Note: Schedule 1 denotes one PSO-JSAO update and one tDQN-UDO update (used in CAIO); Schedule 2 denotes two PSO-JSAO updates and one tDQN-UDO update; Schedule 3 denotes one PSO-JSAO update and two tDQN-UDO updates; Schedule 4 denotes two tDQN-UDO updates and one PSO-JSAO update; Schedule 5 denotes one tDQN-UDO update and two PSO-JSAO updates.

TABLE IV

MEMORY USAGE, TRAINING TIME AND INFERENCE TIME (100 EPISODES) UNDER DIFFERENT PARAMETER CONFIGURATIONS.

Network	N_{loc}, K	Memory usage		Training time	Inference time
		PSO	DRL		
Scenario 1	(3, 14)	7.44 MB	19.19 MB	204.56 s	74.94 s
Scenario 2	(3, 21)	8.75 MB	23.78 MB	226.72 s	110.10 s
Scenario 3	(3, 28)	11.16 MB	34.49 MB	451.69 s	170.79 s
Scenario 4	(3, 50)	26.91 MB	46.74 MB	1202.5 s	582.27 s

method.

D. Communication Performance Analysis

In this section, we present simulation results to demonstrate the communication performance of the proposed method. Fig. 8a shows the performance of the average system throughput with the number of outer iterations under different values of the sub-channel number K . Under the fixed alternation schedule (one PSO-JSAO update followed by one dueling double DQN-UDO update), we can observe that the average system throughput increases rapidly in the early iterations

and then gradually stabilizes, confirming that the proposed alternating process converges to a stable suboptimal solution. Moreover, the converged throughput increases with the number of sub-channels, since a larger K offers more spectrum resources and allocation flexibility, thereby reducing interference and improving communication performance. Fig. 8b compares alternation schedules with different update frequencies and execution orders of PSO-JSAO and dueling double DQN-UDO. The proposed CAIO method adopts the Schedule 1 (one PSO-JSAO update and one tDQN-UDO update). It can be seen that all schedules converge to stable solutions and the Schedule 1 achieves the best final throughput. This is because a single update of each method can already provide a stable suboptimal solution, whereas more frequent updates may cause over-optimization on outdated intermediate variables. In addition, executing PSO-JSAO before dueling double DQN-UDO first optimizes key variables, including sub-channel allocation and UAV association, providing a reliable basis for subsequent UAV deployment optimization and improving overall performance. In total, in Scenario 2, the CAIO achieves maximum average system throughput gains of 82.20%, with average gains of 16.90% for $N_{loc} = 4$.

Fig. 9 shows the average system throughput of the CAIO method under different network scenarios and sub-channel numbers. Blue circles, orange triangles, and green squares correspond to scenarios with $N_{loc} = 3, 4, 5$, respectively. Since localization accuracy is a strict constraint, throughput results are omitted when the accuracy requirement is not satisfied. Across all scenarios and N_{loc} , throughput increases with the number of sub-channels, as more sub-channels provide additional network resources and improve communication performance under the localization constraint. And the throughput curve for $N_{loc} = 5$ saturates earlier than those for $N_{loc} = 3$ and $N_{loc} = 4$, demonstrating that higher anchor density improves resource utilization and accelerates saturation. Moreover, in Scenario 4, the performance slightly decreases due to the larger network scale, as fewer average resources are available per node under the same number of sub-channels. Multiple experiments with different random seeds were also conducted, and the resulting variations remain within a reasonable range, confirming the reliability and stability of the results. Across the four network scenarios, compared with $N_{loc} = 3$ and $N_{loc} = 4$, $N_{loc} = 5$ achieves average throughput gains of 58.19%, 35.41%, 62.94%, 79.6% and 14.88%, 7.09%, 16.77%,

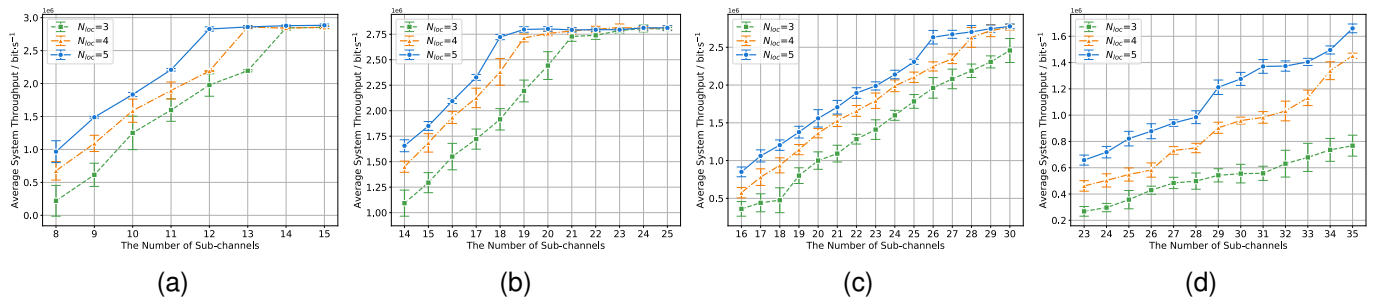


Fig. 9. The average system throughput of CAIO under different sub-channels. (a) Scenario 1. (b) Scenario 2. (c) Scenario 3. (d) Scenario 4

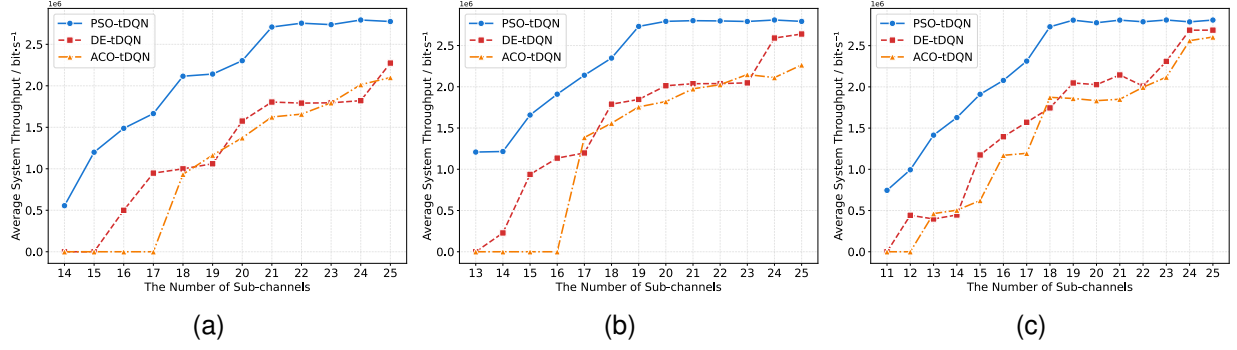


Fig. 10. The average system throughput achieved by different UAV-association and sub-channel-allocation methods in Scenario 2. (a) $N_{\text{loc}} = 3$. (b) $N_{\text{loc}} = 4$. (c) $N_{\text{loc}} = 5$.

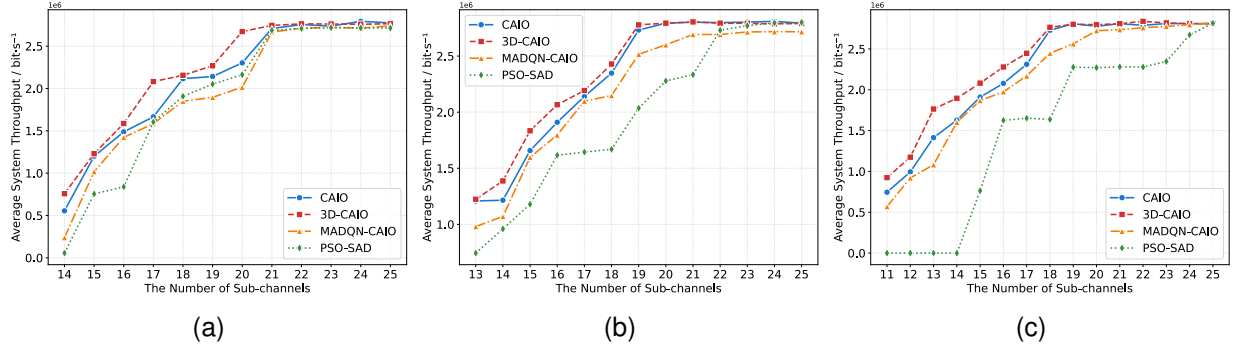


Fig. 11. The average system throughput performance achieved by different joint optimization algorithms in Scenario 2. (a) $N_{\text{loc}} = 3$. (b) $N_{\text{loc}} = 4$. (c) $N_{\text{loc}} = 5$.

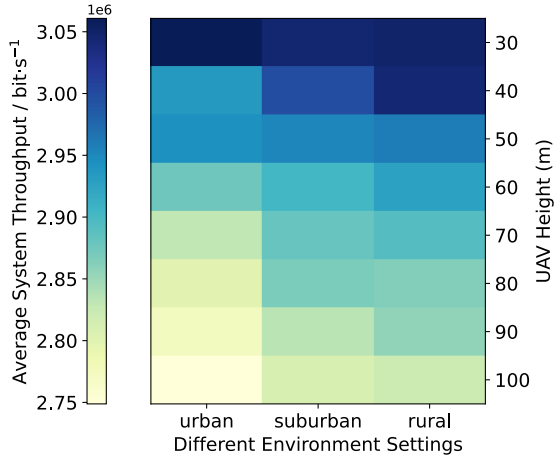


Fig. 12. The average system throughput performance under different environments and heights.

44.9%, respectively. Relative to $N_{\text{loc}} = 3$, $N_{\text{loc}} = 4$ yields gains of 35.47%, 23.86%, 37.36%, 64.9%.

Fig. 10 shows the average system throughput in network Scenario 2 under PSO-tDQN, DE-tDQN, and ACO-tDQN with tDQN-UDO for UAV deployment, where throughput values not meeting the localization accuracy constraint are set to zero. PSO-tDQN consistently outperforms the baselines, exhibiting strong throughput growth under limited sub-channels and rapid

convergence as resources increase. This is because PSO preserves high-quality feasible UAV association and sub-channel allocation solutions through personal-best and global-best guidance, while balancing exploration and exploitation via its inertia and learning terms. DE-tDQN and ACO-tDQN underperform in low-resource regions, particularly for $N_{\text{loc}} = 3$ and 4, where limited sub-channels hinder satisfying localization requirements. With sufficient resources, all methods meet the constraints, but PSO-tDQN achieves the highest throughput, improving over DE-tDQN and ACO-tDQN by 69.57% and 67.72% at $N_{\text{loc}} = 3$, 43.23% and 50.41% at $N_{\text{loc}} = 4$, and 70.14% and 86.79% at $N_{\text{loc}} = 5$, respectively, highlighting its superior balance between communication performance and localization accuracy.

Fig. 11 compares average system throughput between CAIO (solid lines), 3D-CAIO (dashed lines), MADQN-CAIO (dotted lines) and the PSO-SAD (dashed-dotted lines) under varying numbers of sub-channels and N_{loc} in Scenario 2. We can observe that CAIO and 3D-CAIO consistently outperform MADQN and PSO-SAD for $N_{\text{loc}} = 3, 4, 5$, with the performance gap most pronounced under limited sub-channels or fewer localization nodes, demonstrating the effectiveness of jointly optimizing resource allocation, UAV association, and deployment. Meanwhile, it also illustrates the applicability of the proposed method across multiple optimization dimensions, demonstrating that the proposed method is highly general and adaptable. We then observe that the performance of MADQN

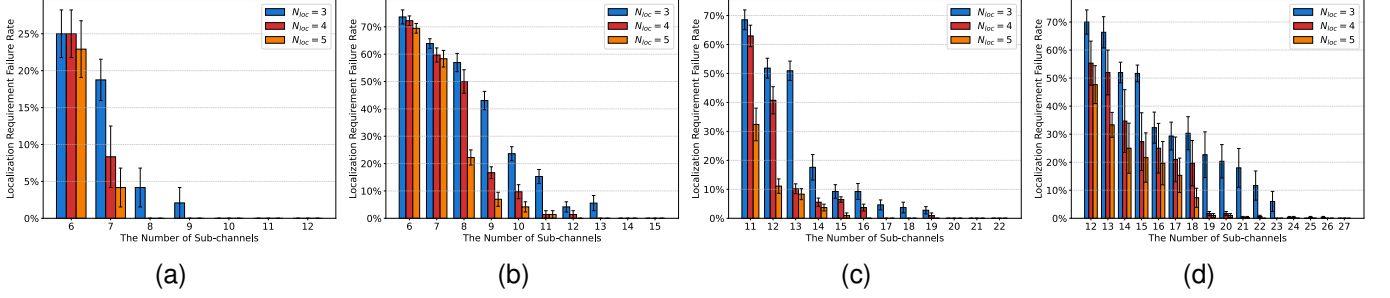


Fig. 13. The percentage of users that do not meet the localization accuracy changes with the number of sub-channels. (a) Scenario 1. (b) Scenario 2. (c) Scenario 3. (d) Scenario 4.

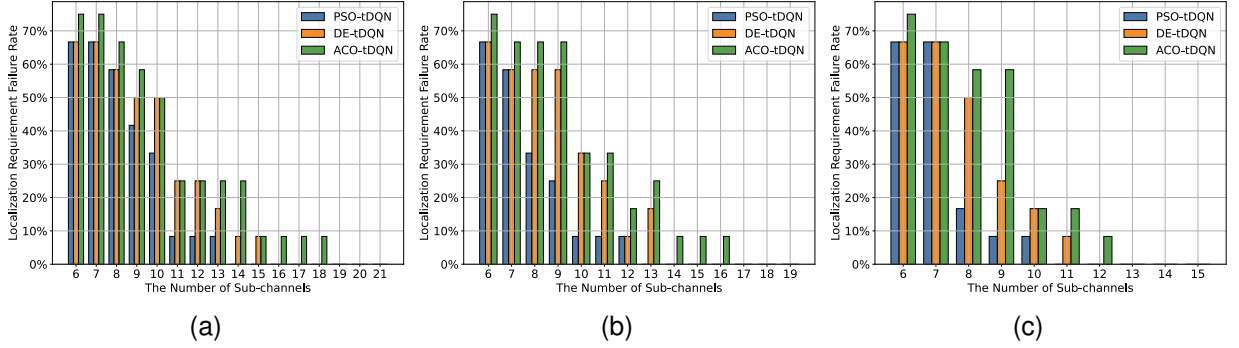


Fig. 14. The percentage of users that do not meet the localization accuracy changes with the number of sub-channels under Scenario 2. (a) $N_{loc} = 3$. (b) $N_{loc} = 4$. (c) $N_{loc} = 5$.

is slightly lower than that of CAIO, suggesting that the increased complexity of multi-agent interactions may introduce additional non-stationarity and slow down convergence under the considered settings. Regarding the CAIO and 3D-CAIO, as shown in the figure, when the number of sub-channels is small, indicating limited network resources, 3D-CAIO achieves better performance, demonstrating that 3D optimization is more effective in utilizing the available resources and enhancing system performance under resource constraints. Overall, averaged over $N_{loc} = 3, 4, 5$, CAIO outperforms MADQN-CAIO and PSO-SAD by 10.69% and 49.19%, respectively. 3D-CAIO outperforms CAIO by 6.17% for $N_{loc} = 3, 4, 5$.

Fig. 12 shows the variation of average system throughput of CAIO under different environments and heights, where darker colors represent higher throughput and better communication performance. It is evident that in all three environments, as the height increases, the average system throughput decreases, indicating that the effect of altitude on channel gain has a greater impact than the LoS probability, resulting in a reduction in communication performance. In different environmental scenarios, the system throughput in urban and suburban areas is lower compared to rural areas, primarily due to NLoS propagation caused by buildings and obstacles. Therefore, system throughput is influenced by both altitude and environmental factors.

E. Localization Performance Analysis

This section evaluates the localization performance of the proposed method using the percentage of users failing to

meet the localization accuracy requirement. Fig. 13 shows this metric across four network scenarios with varying numbers of localization UAVs ($N_{loc} = 3, 4, 5$) and sub-channels. For clarity, x-axis values corresponding to full compliance are omitted. Overall, the percentage of unsatisfied users decreases markedly as the number of sub-channels increases, highlighting the importance of resource availability for localization performance. Specifically, in Scenario 1 (Fig. 13a), high localization accuracy is achieved even with few sub-channels, owing to the sufficient resource availability. In Scenario 2 and Scenario 3 (Fig. 13b and Fig. 13c), the percentage of unsatisfied users is higher, and performance differences among different N_{loc} become more pronounced, indicating that increasing the number of localization anchors effectively improves localization accuracy in resource-constrained or dense user environments. In Scenario 4 (Fig. 13d), the percentage of unsatisfied users significantly increases due to the limited network resources. Overall, in resource-limited or complex scenarios, CAIO significantly enhances localization performance through joint optimization of sub-channel allocation, UAV association and deployment. Additionally, simulations with various random seeds were performed to highlight the robustness and reliability of the proposed method in localization performance.

Fig. 14 shows the localization performance of the proposed method and baselines under different numbers of sub-channels for $N_{loc} = 3, 4, 5$. For $N_{loc} = 3$, many users cannot satisfy the localization requirement when the number of sub-channels is small. As K increases, the number of unsatisfied users decreases steadily and reaches zero. Notably, for $N_{loc} = 5$,

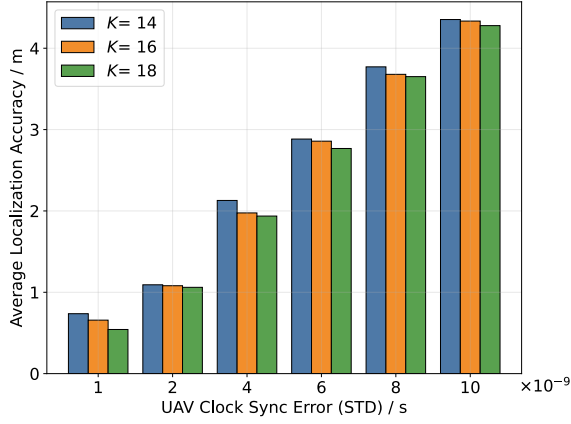


Fig. 15. The average localization accuracy under different clock synchronization errors STD and varying sub-channel numbers ($N_{loc} = 5$, network Scenario 2, $\delta_r = 5$ m).

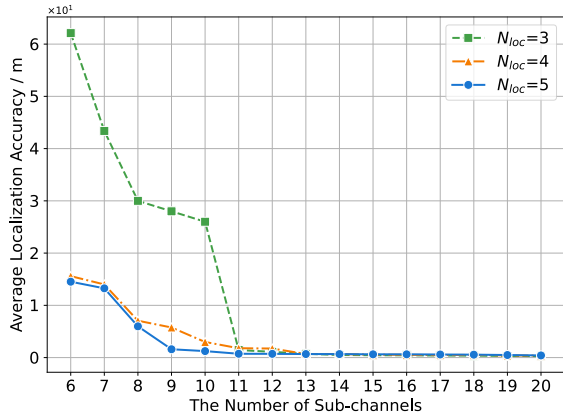


Fig. 16. The average localization accuracy of the CAIO versus the number of sub-channels in Scenario 2.

high localization accuracy is maintained even with limited sub-channels, and PSO-tDQN consistently outperforms DE-tDQN and ACO-tDQN. Overall, PSO-tDQN achieves up to 25% improvement in the percentage of users satisfying the localization constraint, with average gains of 8.3% and 12.82%, confirming its effectiveness in sub-channel allocation and UAV association under different environments.

Fig. 15 shows the relationship between the average localization accuracy and the UAV clock synchronization error STD. The results show that the UAV clock synchronization errors directly deteriorate the ToA-based localization accuracy, but the impact of clock synchronization errors is limited. Meanwhile, increasing the available spectrum resources, corresponding to a larger K , can partially mitigate the adverse impact of clock synchronization errors. This is because more sub-channels help reduce co-channel interference and enhance the quality of ToA measurements, thereby improving the localization accuracy. Therefore, clock synchronization errors are neglected to simplify the model and focus on the joint optimization of spectrum allocation, UAV association, and UAV deployment.

Fig. 16 shows the relationship between the number of sub-

channels and the average localization accuracy for three different settings: $N_{loc} = 3, 4, 5$. As the number of sub-channels increases, the average localization accuracy improves, not merely the percentage of users failing to meet the localization requirement. The average localization accuracy stabilizes when $K \geq 12$ and exhibits a trend consistent with the previous simulation results.

VII. CONCLUSION

This paper investigated joint spectrum, UAV association, and deployment optimization in a UAV swarm-assisted ISAC network for emergency rescue, aiming to maximize system average throughput while ensuring localization accuracy. A UAV swarm-assisted integrated communication and localization method was proposed, along with a joint optimization model decomposed into two subproblems. The PSO-JSAO method optimized sub-channel allocation and UAV association, while tDQN-UDO further optimized UAV deployment. The joint iterative CAIO method alternately solved the subproblems to enhance both communication and localization performance.

Simulation results demonstrate that PSO-tDQN achieves an average throughput gain of 64.64% over DE-tDQN and ACO-tDQN, and that increasing the number of localization anchors substantially improves throughput. Additionally, the proposed method increases the proportion of users meeting localization accuracy requirements by up to 25%. Overall, the proposed method effectively balances communication performance and localization accuracy, providing reliable support for efficient emergency rescue operations. In the future, we will consider dynamic scenarios (such as mobile users) and UAV energy consumption, to further enhance the method's adaptability and robustness, enabling better deployment in real-world environments. In addition, inter-cluster spectrum sharing will be explored to enhance system performance in dense scenarios such as smart logistics and smart city applications.

REFERENCES

- [1] X. Zhu, J. Liu, L. Lu, T. Zhang, T. Qiu, C. Wang, and Y. Liu, "Enabling intelligent connectivity: A survey of secure ISAC in 6G networks," *IEEE Communications Surveys & Tutorials*, vol. 27, no. 2, pp. 748–781, 2025.
- [2] X. Luo, Q. Lin, R. Zhang, H.-H. Chen, X. Wang, and M. Huang, "ISAC—a survey on its layered architecture, technologies, standardizations, prototypes and testbeds," *IEEE Communications Surveys & Tutorials*, vol. 28, pp. 485–526, 2026.
- [3] C. Jiang, X. Zhou, H. Chen, and T. Liu, "UAV positioning using GNSS: A review of the current status," *Drones*, vol. 10, no. 2, 2026, Art. no. 91.
- [4] Y. Cui, C. Wang, Q. Hu, B. Xu, X. Song, Z. Yuan, and Y. Zhu, "A novel positioning method for UAV in GNSS-denied environments based on mechanical antenna," *IEEE Transactions on Industrial Electronics*, vol. 71, no. 10, pp. 13 461–13 469, 2024.
- [5] S. Javed, A. Hassan, R. Ahmad, W. Ahmed, R. Ahmed, A. Saadat, and M. Guizani, "State-of-the-art and future research challenges in UAV swarms," *IEEE Internet of Things Journal*, vol. 11, no. 11, pp. 19 023–19 045, 2024.
- [6] W. Zhao, S. Cui, W. Qiu, Z. He, Z. Liu, X. Zheng, B. Mao, and N. Kato, "A survey on DRL based UAV communications and networking: DRL fundamentals, applications and implementations," *IEEE Communications Surveys & Tutorials*, vol. 28, pp. 3911–3941, 2026.
- [7] M. Ahmed, A. A. Soofi, F. Khan, S. Raza, W. U. Khan, L. Su, F. Xu, and Z. Han, "Toward a sustainable low-altitude economy: A survey of energy-efficient RIS-UAV networks," *IEEE Internet of Things Journal*, vol. 12, no. 24, pp. 51 951–51 975, 2025.

- [8] X. Cheng, R. Jiang, H. Sang, G. Li, and B. He, "Joint optimization of multi-UAV deployment and user association via deep reinforcement learning for long-term communication coverage," *IEEE Transactions on Instrumentation and Measurement*, vol. 73, 2024, Art. no. 5503613.
- [9] Y. He, Y. Liu, and J. Xie, "A distributed approach for user association and UAV deployment in QoE-aware multi-UAV networks," *IEEE Transactions on Network and Service Management*, vol. 22, no. 5, pp. 4394–4406, 2025.
- [10] S. Geng, Z. Wei, J. Zhao, F. Shen, J. Joung, and S. Sun, "Joint deployment and resource allocation for service provision in multi-UAV-assisted wireless networks," *IEEE Internet of Things Journal*, vol. 11, no. 22, pp. 37 269–37 286, 2024.
- [11] L. Ge, X. Liang, H. Zhang, P. Dong, J. Liao, and J. Wang, "Joint resource allocation and trajectory design for resilient multi-UAV communication networks," *IEEE Wireless Communications Letters*, vol. 13, no. 4, pp. 994–998, 2024.
- [12] N. Khan, A. Ahmad, A. Wakeel, Z. Kaleem, B. Rashid, and W. Khalid, "Efficient UAVs deployment and resource allocation in UAV-relay assisted public safety networks for video transmission," *IEEE Access*, vol. 12, pp. 4561–4574, 2024.
- [13] W. Sun, H. Wang, Z. Qin, and Z. Qin, "Joint optimization of deployment, user association, channel, and resource allocation for fairness-aware multi-UAV network," *IET Communications*, vol. 18, no. 13, pp. 799–816, 2024.
- [14] C. Wang, D. Deng, L. Xu, and W. Wang, "Resource scheduling based on deep reinforcement learning in UAV assisted emergency communication networks," *IEEE Transactions on Communications*, vol. 70, no. 6, pp. 3834–3848, 2022.
- [15] G. Chen, X. B. Zhai, and C. Li, "Joint optimization of trajectory and user association via reinforcement learning for UAV-aided data collection in wireless networks," *IEEE Transactions on Wireless Communications*, vol. 22, no. 5, pp. 3128–3143, 2023.
- [16] Y. Song, Q. Zeng, G. Chen, G. Yuan, and H. Duan, "Multi-agent proximal policy optimization based efficient user association and resource allocation in UAV-assisted heterogeneous cellular networks," *Computer Communications*, vol. 238, 2025, Art. no. 108172.
- [17] S. Qu, Y. Zhao, and X. Li, "Joint energy beamforming and bandwidth allocation for cooperative localization in passive sensor networks," *Internet of Things*, vol. 25, 2024, Art. no. 101073.
- [18] G. Fatima, P. Stoica, A. Aubry, A. De Maio, and P. Babu, "Optimal placement of the receivers for multistatic target localization," *IEEE Transactions on Radar Systems*, vol. 2, pp. 391–403, 2024.
- [19] A. Famili and A. Stavrou, "CARTA: Coordinated arrangement of receivers for target acquisition," in *2024 IFIP Networking Conference (IFIP Networking)*. IEEE, 2024, pp. 323–331.
- [20] A. Famili, S. Sun, T. Atalay, and A. Stavrou, "STARMAP: Spaceborne target acquisition radar with meta-RL assisted placement," *IEEE Open Journal of the Communications Society*, vol. 6, pp. 6218–6241, 2025.
- [21] A. M. Graff and T. E. Humphreys, "Ziv-Zakai-optimal OFDM resource allocation for time-of-arrival estimation," *IEEE Transactions on Wireless Communications*, vol. 24, no. 8, pp. 6886–6901, 2025.
- [22] A. Saikko, J. Talvitie, J. S  , J. Pirskanen, and M. Valkama, "Positioning and tracking in DECT-2020 NR with proactive anchor selection for range, angle, and RSS measurements," *IEEE Journal of Indoor and Seamless Positioning and Navigation*, vol. 3, pp. 70–80, 2025.
- [23] M. Jes  s-Azabal, M. Zheng, and V. N. Soares, "Dynamic energy-aware anchor optimization for contact-based indoor localization in MANETs," *Information*, vol. 16, no. 10, 2025, Art. no. 855.
- [24] D. Ebrahimi, S. Sharafeddine, P.-H. Ho, and C. Assi, "Autonomous UAV trajectory for localizing ground objects: A reinforcement learning approach," *IEEE Transactions on Mobile Computing*, vol. 20, no. 4, pp. 1312–1324, 2021.
- [25] M. Kato, T. K. Rodrigues, T. Abe, and T. Suganuma, "Exploiting radio frequency characteristics with a support unmanned aerial vehicle to improve wireless sensor location estimation accuracy," *IEEE Internet of Things Journal*, vol. 11, no. 24, pp. 39 570–39 578, 2024.
- [26] S. Bi, J. Yu, Z. Yang, X. Lin, and Y. Wu, "Joint 3-D deployment and resource allocation for UAV-assisted integrated communication and localization," *IEEE Wireless Communications Letters*, vol. 12, no. 10, pp. 1672–1676, 2023.
- [27] Z. Yang, S. Bi, and Y.-J. A. Zhang, "Deployment optimization of dual-functional UAVs for integrated localization and communication," *IEEE Transactions on Wireless Communications*, vol. 22, no. 12, pp. 9672–9687, 2023.
- [28] J. Yang, J. Lee, and J. Lim, "Joint placement and communication optimization of UAV base stations in GPS-denied environments," *Journal of Communications and Networks*, vol. 26, no. 5, pp. 490–501, 2024.
- [29] Z. Wang, R. Liu, Q. Liu, J. S. Thompson, and M. Kadoch, "Energy-efficient data collection and device positioning in UAV-assisted IoT," *IEEE Internet of Things Journal*, vol. 7, no. 2, pp. 1122–1139, 2020.
- [30] Q. Zhu, R. Liu, Z. Wang, Q. Liu, and C. Chen, "Sensing-communication co-design for UAV swarm-assisted vehicular network in perspective of Doppler," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 2, pp. 2578–2592, 2024.
- [31] Y. Tao, Z. Gao, Z. Li, Z. Wan, T. Li, C. Zhu, L. Chen, G. Wen, D. Zheng, and D. Niyato, "Pseudo-random TDM-MIMO FMCW based millimeter-wave sensing and communication integration for UAV swarm," *IEEE Internet of Things Journal*, vol. 13, no. 1, pp. 549–564, 2026.
- [32] Y. Xie, L. Wang, Z. Chang, L. Xu, S. Bi, and Z. Han, "GeoAgg-HSAC: an RL-based framework for trajectory and resource optimization in mountainous UAV integrated localization and communication networks," *IEEE Transactions on Wireless Communications*, vol. 25, pp. 6507–6522, 2026.
- [33] Q. Zhu, R. Liu, Q. Liu, and C. Chen, "Resource allocation for UAV swarm-assisted green ISAC networks via multi-agent RL," *IEEE Transactions on Green Communications and Networking*, vol. 9, no. 3, pp. 1354–1367, 2025.
- [34] M. Zhu, Z. Wei, C. Qiu, W. Jiang, H. Wu, and Z. Feng, "Joint data collection and sensor positioning in multi-UAV-assisted wireless sensor network," *IEEE Sensors Journal*, vol. 23, no. 19, pp. 23 664–23 675, 2023.
- [35] L. Yin, J. Cao, K. Lin, Z. Deng, and Q. Ni, "A novel positioning-communication integrated signal in wireless communication systems," *IEEE Wireless Communications Letters*, vol. 8, no. 5, pp. 1353–1356, 2019.
- [36] L. Yin, J. Cao, Q. Ni, Y. Ma, and S. Li, "Design and performance analysis of multi-scale NOMA for future communication-positioning integration system," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 4, pp. 1333–1345, 2022.
- [37] 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, 2020.
- [38] H. Guo, Y. Wang, J. Liu, and C. Liu, "Multi-UAV cooperative task offloading and resource allocation in 5G advanced and beyond," *IEEE Transactions on Wireless Communications*, vol. 23, no. 1, pp. 347–359, 2024.
- [39] I. Jarraya, A. Al-Batati, M. B. Kadri, M. Abdelkader, A. Ammar, W. Boulila, and A. Koubaa, "GNSS-denied unmanned aerial vehicle navigation: Analyzing computational complexity, sensor fusion, and localization methodologies," *Satellite Navigation*, vol. 6, 2025, Art. no. 9.
- [40] B. Zhang, X. Shao, Y. Wang, G. Sun, and W. Yao, "R-LVIO: Resilient LiDAR-visual-inertial odometry for UAVs in GNSS-denied environment," *Drones*, vol. 8, no. 9, 2024, Art. no. 487.
- [41] S. A. Zekavat and R. M. Buehrer, Eds., *Handbook of Position Location: Theory, Practice, and Advances*, 2nd ed. Hoboken, NJ, USA: Wiley-IEEE Press, 2019.
- [42] J. W. Betz and K. R. Kolodziejski, "Generalized theory of code tracking with an early-late discriminator Part I: Lower bound and coherent processing," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 45, no. 4, pp. 1538–1556, 2009.
- [43] H. Xiao, X. Hu, W. Zhang, W. Wang, K.-K. Wong, and K. Yang, "Energy-efficient STAR-RIS enhanced UAV-enabled MEC networks with bi-directional task offloading," *IEEE Transactions on Wireless Communications*, vol. 24, no. 4, pp. 3258–3272, 2025.
- [44] Z. Yang, "UAV communication-positioning joint optimization," GitHub repository, 2026, accessed: May 25, 2026. [Online]. Available: https://github.com/sheepon123/MN_25092601



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.



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.



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.



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.



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.

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.