

Multi-layer UAV Deployment and Proactive Caching Optimization for Timely Information Delivery in Emergency Services

Lingfeng Shen, Ming Li, Jiaxin Xu, Lingyun Li, Guanghui Wang*, Fang Zuo, Yansong Liu, and Xin He

Abstract: In disaster emergency response, providing timely situational awareness information presents a significant challenge, particularly when ground infrastructure is compromised. To address this issue, this study introduces a multilayered, heterogeneous unmanned aerial vehicle (UAV)-assisted network designed for efficient information delivery. Low latency is achieved through the integrated optimization of proactive caching and three-dimensional (3D) UAV deployment. The joint optimization problem is formulated as a complex mixed-integer nonlinear programming (MINLP) problem, which is difficult to solve. Therefore, we decompose the primary issue into two sub-problems: One focusing on the caching strategy and the other on the UAV deployment strategy. The caching sub-problem is an non-deterministic polynomial-time hard (NP-hard) knapsack problem. To effectively balance the transmission speed and completeness of emergency information, this work develops two novel caching algorithms: Hot information priority (HIP) and diversified information priority (DIP). Regarding UAV deployment, the 3D placement problem is inherently non-convex. We further divide it into horizontal position optimization and vertical height optimization. The horizontal position optimization can be reformulated as a convex optimization problem. The vertical height optimization problem is theoretically proven to be quasiconvex and can be efficiently solved. Simulation results demonstrate that the proposed framework markedly reduces the latency of information dissemination in UAV-assisted emergency networks.

Key words: proactive caching; multi-layer unmanned aerial vehicles (UAVs); cross-layer caching; three-dimensional (3D) deployment; tethered UAV; quadrotor UAV

1 Introduction

In recent years, natural disasters such as earthquakes,

- Lingfeng Shen, Ming Li, Jiaxin Xu, Lingyun Li, Guanghui Wang, Fang Zuo, and Xin He are with School of Software, Henan University, Kaifeng 475004, China. E-mail: ielfshen@henu.edu.cn; liminglm@henu.edu.cn; xjx123@henu.edu.cn; lly174916@henu.edu.cn; gwang@vip.henu.edu.cn; zuofang@henu.edu.cn; hxsyjkf@foxmail.com.

- Yansong Liu is with RWTH Aachen University, Aachen 52062, Germany. E-mail: yansong.liu@rwth-aachen.de.

* To whom correspondence should be addressed.

Manuscript received: 2025-12-01; revised: 2026-03-12; accepted: 2026-05-30

floods, and wildfires have become increasingly frequent worldwide. The destruction of terrestrial communication infrastructure, including optical fiber cables, often leads to severe network disruptions, creating “communication islands” in affected areas^[1–3]. This situation impedes the dissemination of critical disaster information to external parties and hinders the reception of rescue directives internally, presenting formidable challenges to emergency response^[4]. In June 2025, an unprecedented flood event struck the Suifen River basin in Dongning City, Heilongjiang Province, China. The disaster caused extensive damage, leading to the total collapse of terrestrial communication networks, power grids, and

transportation routes. Consequently, the affected region was plunged into a state of total isolation. The communication breakdown prevented affected individuals from quickly accessing vital information, such as locations of shelters and evacuation plans, drastically diminishing the efficiency of deploying rescue resources and directly jeopardizing lives and property^[5].

Prompt and efficient distribution of disaster and command information is essential for emergency rescue^[6, 7]. Unmanned aerial vehicles (UAVs) have emerged as a pivotal alternative solution for emergency communication constraints, owing to their advantageous characteristics: Affordability, rapid deployment capability, and superior mobility^[8]. Reference [9] introduced a UAV-assisted collaborative communication emergency network for rural and semi-urban environments, achieving extensive coverage and extremely dependable emergency communication by employing UAVs as relays and airborne base stations. Reference [10] offered several efficient UAV deployment strategies for four distinct environments: suburban, urban, dense urban, and high-rise dense urban locations. This method enhances the deployment of UAVs in contexts subject to dynamic changes and differing disaster area coverage needs, ensuring prompt and precise information dissemination while securing coverage for terrestrial users. Reference [11] examined large-scale UAV deployment tactics in extremely dynamic emergency rescue situations, employing aggregative games to represent limits on strategy updates and decision-making errors, markedly improving emergency coverage performance.

Despite the numerous significant advantages of UAVs in emergency communications, they have limited payload capacity^[12]. The problem of optimal information storage within limited cache capacity has emerged as a pressing challenge^[13, 14]. The heterogeneity and uncertainty of demands in emergency networks significantly increase the complexity of cache strategy design^[15]. Literature^[16] proposed an active joint strategy for UAV trajectory planning and cache optimization, effectively addressing the practical problem of incomplete user content demand distribution information by introducing a data-driven approach based on first- and second-order statistics. Unlike conventional approaches dependent on complete probability distribution assumptions, this

research developed network optimization solutions exhibiting robust distribution characteristics, delivering substantially improved UAV network service performance under conditions of demand uncertainty. In Ref. [17], researchers proposed a multi-objective optimization problem for task offloading, computational resource allocation, and UAV trajectory control in a multi-UAV assisted mobile edge computing system. They addressed the coupling of these decision variables by decomposing the problem into three parts, then solving each separately to obtain the respective decisions.

In addition to the aforementioned caching strategies, proactive caching strategies based on 6G edge-cloud collaboration are gaining prominence^[18]. This research direction has already yielded fruitful results. Literature Ref. [19] focuses on proactive caching based on distributed deep reinforcement learning (DRL), exploring content prediction and content decision methods. The researchers jointly optimized content prediction using the temporal convolution network (TCN) and content access and replacement strategies based on the distributional deep Q network (DDQN) by considering the actual dynamic attributes of user requests. They constructed a novel TCN-based strategy using an attention model named temporal convolution sequence network (TCSN), improving prediction accuracy. They then developed a multi-element fusion algorithm called particle swarm optimization-guided (PG)-DDQN by combining a generative adversarial network (GAN) with prioritized experience replay (PER) based on DDQN. Another work proposed federated integrated privacy-centric (FIPC), a proactive caching scheme combining federated learning with implicit feedback^[20], which predicts information popularity from historical user request patterns and pre-caches the top-N items with the highest recommendation scores. Experimental results demonstrate that the proposed FIPC scheme outperforms baseline caching algorithms and significantly reduces overall service latency.

Moreover, because of the restricted energy capacity of UAVs, they will cease functioning once their power is exhausted, resulting in service disruptions that compromise the continuity and efficacy of rescue operations^[21]. Designing energy-efficient emergency information distribution strategies is thus of considerable significance. This necessitates network

architecture designs that meet low power consumption and minimal cache cost requirements^[22]. Existing literature proposed algorithms for UAV deployment optimization and policy learning, improving coverage, reducing inter-UAV interference, and enhancing energy efficiency.

This study proposes a hierarchical, heterogeneous network for the distribution of emergency information assisted by UAVs. The system achieves energy-efficient and low-latency information dissemination using an integrated methodology that merges proactive caching with optimized three-dimensional (3D) UAV deployment. The main contributions of this work are summarized as follows:

- To address the limited energy capacity of rotary-wing UAVs, this paper proposes a hierarchical service architecture based on the collaboration between tethered UAVs (TUs) and quadrotor UAVs (QUs). This architecture leverages a multi-layer UAV system equipped with proactive caching technology to facilitate efficient information distribution in emergency networks. We jointly optimize the 3D deployment of UAVs and proactive caching strategies, minimizing information dissemination latency while reducing both the operational duration and energy consumption of the QUs.

- The joint optimization of UAV 3D deployment and cross-layer proactive caching constitutes a complex mixed-integer problem. To manage this complexity, we decompose the original problem into two sub-problems: One for optimizing the information caching strategy and another for determining the optimal UAV positions.

- By establishing that minimizing information delivery latency is equivalent to minimizing the miss request rate, we developed two novel caching strategies: Hot information priority (HIP) and diversified information priority (DIP). These effectively solve the non-deterministic polynomial hard (NP-hard) cross-layer information caching challenge.

- The 3D deployment problem for UAVs is divided into horizontal position optimization and hovering altitude optimization. The horizontal position optimization is reformulated as a convex optimization problem, while the optimization of the QU's hovering altitude is proven to be quasiconvex. Both subproblems can be efficiently solved by the convex (CVX) toolbox.

The remainder of this paper is organized as follows.

The system model and the original optimization problem formulation are presented in Section 2. The proactive caching strategy design is proposed in Section 3. The QUs' 3D deployment strategy design is proposed subsequently in Section 4. Numerical results evaluating the performance of the proposed schemes are presented in Section 5. Concluding remarks are given in Section 6.

2 System Model and Problem Description

In this work, we propose a multi-layer UAV architecture equipped with proactive caching technology to minimize information dissemination latency in emergency networks, as illustrated in Fig. 1. A large number of users distributed over a wide area are organized into clusters, with each cluster served by a QU acting as a low-altitude platform. However, the QU's limited onboard cache capacity can become a bottleneck, potentially failing to satisfy all user information requests. To overcome this limitation, a TU is introduced as a high-altitude platform. The TU, capable of covering all clusters, serves as a larger-capacity cache. Furthermore, it is equipped with a reliable free-space optical (FSO) communication module for high-speed transmission with the emergency command center (information server), thereby satisfying ground users' diverse information demands.

2.1 Channel model

For UAV-assisted large-scale communication, clustering can overcome the shortcoming of the UAV's limited service capacity and coverage, which is an effective and mature way to improve UAV

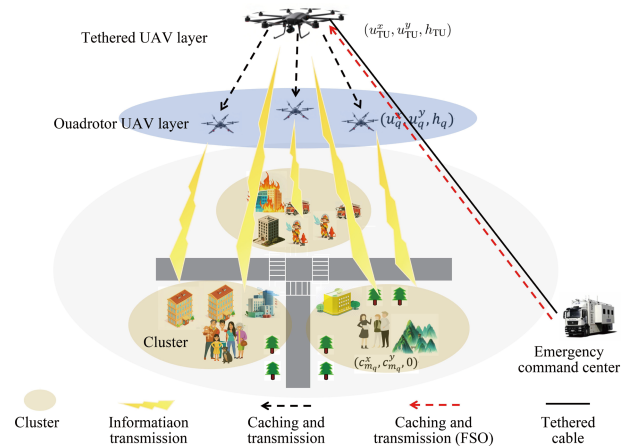


Fig. 1 Multi-layer proactive information caching model.

communication quality. Besides, in the event of a disaster, individuals typically congregate, either spontaneously or guided, at predetermined emergency shelters. Building upon this observation, this paper posits that ground users are effectively clustered.

As illustrated in Fig. 1, this network consists of Q QUs and one TU, where Q denotes the number of QUs. The QU index set is denoted as $\mathcal{Q} = \{1, 2, \dots, Q\}$, where $q \in \mathcal{Q}$ is the QU index. Suppose the q^{th} QU serves M_q users. The user set served by the q^{th} QU is $\mathcal{M}_q = \left\{ \sum_{i=0}^{q-1} M_i + 1, \dots, \sum_{i=0}^q M_i \right\}$, where $M_0 = 0$ is a dummy variable for the convenience of notation. There are $M = \sum_{q=1}^Q M_q$ users in total, and all users are distributed on the ground. Define the height of the ground as 0, and the flying heights of the QUs and TU as h_q and h_{TU} , respectively. The coordinates of the TU and the q -th QU are $u_{TU} = (u_{TU}^x, u_{TU}^y, h_{TU})$ and $u_q = (u_q^x, u_q^y, h_q)$, respectively. The coordinates of the m_q -th ground user in the q -th group can be expressed as $c_{m_q} = (c_{m_q}^x, c_{m_q}^y, 0)$, where $m_q \in \mathcal{M}_q$.

The different circumstances influence the characteristics of the air-to-ground (A2G) pathloss channel. The A2G link can be modeled by considering the occurrence probability of either line-of-sight (LoS) or non-line-of-sight (NLoS) conditions. The pathloss of LoS and NLoS can be described as

$$L_{LoS} = 20 \lg d + 20 \lg f + 20 \lg \frac{4\pi}{c} + \eta_{LoS} \quad (1)$$

and

$$L_{NLoS} = 20 \lg d + 20 \lg f + 20 \lg \frac{4\pi}{c} + \eta_{NLoS} \quad (2)$$

where d is the physical distance between the transmitter and the receiver, f is the carrier frequency, and c is the speed of light. η_{LoS} and η_{NLoS} are the additional loss on top of the free space pathloss.

The LoS and NLoS occurrence probability can be given by^[23]

$$\Pr_{LoS}(\theta) = \frac{1}{1 + g \exp(-h[\theta - g])} \quad (3)$$

$$\Pr_{NLoS}(\theta) = 1 - \Pr_{LoS}(\theta) \quad (4)$$

where g and h are constants related to the environment, such as the building's height and density. θ is the elevation angle, $\theta = \arcsin\left(\frac{h_q}{d}\right)$.

Orthogonal resource blocks are allocated to different users to avoid inter-user interference. The

instantaneous achievable transmission rate between the q^{th} QU and its users can be expressed as

$$R_{QUs}(u_q) = \Pr_{LoS}(\theta_q) \log_2 \left(1 + \frac{P_q}{\lambda_{LoS} \sigma^2 \|u_q - c_{m_q}\|_2^2} \right) + \Pr_{NLoS}(\theta_q) \log_2 \left(1 + \frac{P_q}{\lambda_{NLoS} \sigma^2 \|u_q - c_{m_q}\|_2^2} \right) \quad (5)$$

where p_q is the transmitting power of the q^{th} QU, σ^2 denotes noise power, and $\lambda_{LoS} = 10^{\frac{L_{LoS}}{10}}$ and $\lambda_{NLoS} = 10^{\frac{L_{NLoS}}{10}}$ are the LoS and NLoS pathloss factors, respectively.

Similarly, the achievable instantaneous transmission rate between TU and its users becomes

$$R_{TU} = \Pr_{LoS}(\theta_{TU}) \log_2 \left(1 + \frac{P_{TU}}{\lambda_{LoS} \sigma^2 \|u_{TU} - c_{m_q}\|_2^2} \right) + \Pr_{NLoS}(\theta_{TU}) \log_2 \left(1 + \frac{P_{TU}}{\lambda_{NLoS} \sigma^2 \|u_{TU} - c_{m_q}\|_2^2} \right) \quad (6)$$

where p_{TU} is the transmitting power of the TU, and θ_{TU} is the elevation angle between the TU and the ground user.

2.2 Information request model

It is postulated that the UAVs possess awareness of the locations of all affected individuals. Within each decision interval, the users' positions and the information to be disseminated remain unchanged. Therefore, we can design the hovering position of QUs and the cross-layer caching strategy between QUs and TU to improve the efficiency of information dissemination. The information library contains N different files, where N denotes the number of files, which is indexed $\{1, 2, \dots, N\}$. For simplicity, we assume that all files have the same normalized size.

In emergency scenarios, information requests typically demonstrate a pronounced "long-tail effect", wherein a few critical pieces of information are frequently sought, while the majority receive comparatively fewer requests. The Zipf distribution provides an effective model for this uneven request pattern. Consequently, in this paper, we employ the Zipf distribution to simulate the popularity of information. The request probability for the n^{th} information in the area covered by the q^{th} QU can be expressed as

$$P_{f_q}^n = \frac{n^{-\alpha^q}}{\sum_{n=1}^N n^{-\alpha^q}}, n \in \{1, 2, \dots, N\}, q \in \{1, 2, \dots, Q\} \quad (7)$$

The Zipf exponent α^q determines the skewness in the request pattern. In general, $\alpha^q > 0$ and the greater the α^q , the request are more concentrated.

The overall popularity of the n^{th} information over all Q clusters is denoted by P_f^n , which is calculated as

$$P_f^n = \sum_{q=1}^Q M_q P_{f_q}^n / M = \frac{\sum_{q=1}^Q M_q n^{-\alpha^q}}{\sum_{n=1}^N M n^{-\alpha^q}} \quad (8)$$

The caching capacity of the TU and QUs are denoted by S_{TU} and S_{QUs} , respectively. We define the information caching indicator variables $\mu_n \triangleq \{\mu_n^q \in \{1, 0\}\}$ indicating whether the n^{th} information is cached at the q^{th} QUs, and $v_n \triangleq \{v_n \in \{1, 0\}\}$ indicating whether the n^{th} information is cached at the TU.

The hit ratio of the q^{th} QUs and TU are defined as

$$r_q(\mu_n) = \frac{\sum_{n=1}^N P_{f_q}^n \mu_n^q}{\sum_{n=1}^N P_{f_q}^n} \quad (9)$$

$$r(v_n) = \frac{\sum_{n=1}^N P_f^n v_n}{\sum_{n=1}^N P_f^n} \quad (10)$$

2.3 Delay model

Assume that all of the files have the same normalized size, represented by γ . The delay to obtain the requested information through three transmission modes can be expressed as $t_{\text{QUs}}(\mathbf{u}_q) = \frac{\gamma}{R_{\text{QUs}}}$, $t_{\text{TU}} = \frac{\gamma}{R_{\text{TU}}}$, and $t_{\text{miss}} = t_{\text{TU}} + t_{\text{fso}}$, where $t_{\text{QUs}}(\mathbf{u}_q)$ and t_{TU} represent the delay to obtain the requested information from QUs and TU, respectively, and t_{miss} indicates that neither QU nor TU caches the information requested by the user. In that case, the user obtains the file from the information server through the TU relay. t_{fso} is the delay for the information server to transfer files to TU via FSO links.

It is assumed that the delay for users to obtain information from QU is less than that from TU. When the QU and TU cache the information requested by the user at the same time, the QU responds to the user's request. The average transmission delay of the requested information is denoted by $T(\mathbf{u}_q, \mu_n, v_n)$ and calculated as

$$\begin{aligned} T(\mathbf{u}_q, \mu_n, v_n) = & r_q(\mu_n)(1 - r(v_n))t_{\text{QUs}}(\mathbf{u}_q) \\ & + r(v_n)(1 - r_q(\mu_n))t_{\text{TU}} \\ & + (1 - r(v_n))(1 - r_q(\mu_n))t_{\text{miss}} \\ & + r_q(\mu_n)r(v_n)t_{\text{QUs}}(\mathbf{u}_q) \end{aligned} \quad (11)$$

2.4 Problem formulation

We formulate the following optimization problem to minimize the information dissemination delay by optimizing the cross-layer caching strategy and the QUs' 3D deployment strategy.

$$\mathbf{P}: \text{minimize}_{\{\mathbf{u}_q, \mu_n, v_n\}} \max\{T(\mathbf{u}_q, \mu_n, v_n)\} \quad (12)$$

$$\text{s.t. } \mu_n^q, v_n \in \{1, 0\} \quad (12a)$$

$$\sum_{n=1}^N \mu_n^q \leq S_{\text{QUs}} \quad (12b)$$

$$\sum_{n=1}^N v_n \leq S_{\text{TU}} \quad (12c)$$

$$h_{\min} \leq h_q \leq h_{\max} \quad (12d)$$

Constraints (12b) and (12c) are the caching capacity constraints of QUs and TU, respectively. Constraint (12d) sets the lower and upper limits, denoted by $h_{\min}$ and $h_{\max}$, respectively, to the flying height of QUs. Obviously, both the cross-layer information caching strategy, and the QUs' hovering position will affect the delay. As a result, the problem in Formula (12) is a non-convex mixed integer optimization problem, which is challenging to solve.

To solve this problem, we decompose it into two sub-problems. We first design the information caching strategy. The first sub-problem can be described as a NP-hard knapsack problem. Minimizing the information delivery delay can be equivalently transformed as minimizing the missing request ratio. Then the NP-hard knapsack problem can be solved by designing the caching principles that minimize the missing request ratio. The second sub-problem optimizes the 3D hovering position of QUs, which is

still non-convex. We divide the 3D deployment problem of UAVs into horizontal position optimization and flying height optimization. The QUs' horizontal position optimization is transformed into a convex optimization problem, while the corresponding hovering height optimization is proved to be a quasiconvex problem. The detailed design and solution are proposed in Section 3 and Section 4, respectively.

3 Cross-Layer Caching Strategy Design

In this part, we design the information caching strategy. The first sub-problem can be written as

$$\text{SP1: } \underset{\{\mu_n, \nu_n\}}{\text{minimize}} \quad \max\{T(\mu_n, \nu_n)\} \quad (13)$$

$$\text{s.t. } \mu_n^q, \nu_n \in \{1, 0\} \quad (13a)$$

$$\sum_{n=1}^N \mu_n^q \leq S_{\text{QUs}} \quad (13b)$$

$$\sum_{n=1}^N \nu_n \leq S_{\text{TU}} \quad (13c)$$

where μ_n and ν_n denote the caching indicator variables of files at QUs and TU, respectively, and $T(\mu_n, \nu_n)$ is the transmission delay of the requested information.

Each file has a size and value in terms of reducing delay. The caching capacity of QUs and TU is limited. This constitutes a typical knapsack problem. In this work, the size of each file is assumed to be the same. Therefore, we only need to consider how to define the files' caching value. Next, we transform **SP1** and define the value of each file accordingly.

By definition, we know that $t_{\text{miss}} > t_{\text{TU}} > t_{\text{QUs}}$. Combining Eq. (11), we can see that the larger the $r_q(\mu_n)$, $r(\nu_n)$, the smaller the $T(\mu_n, \nu_n)$. We propose Proposition 1 as follows:

Proposition 1 Minimizing the information delivery delay can be equivalently transformed into minimizing the missing request ratio $1 - r_q - r$.

Proof First, assume that the positions of all nodes have been fixed, therefore, $t_{\text{QUs}}(\mathbf{u}_q)$, t_{TU} and t_{miss} are constants. For convenience, we rewrite Eq. (11) as

$$T(r_q^*, r^*) = r_q^*(1 - r^*)t_{\text{QUs}} + r^*(1 - r_q^*)t_{\text{TU}} + (1 - r^*)(1 - r_q^*)t_{\text{miss}} + r_q^*r^*t_{\text{QUs}} \quad (14)$$

Assuming that r_q^* and r^* are the hit ratios corresponding to the minimal missing request ratio $1 - r_q - r$ and the minimal system delay. Otherwise, there must be Δ_{r_q} and Δ_r to satisfy

$$\begin{aligned} T(r_q^\Delta, r^\Delta) &= r_q^\Delta(1 - r^\Delta)t_{\text{QUs}} + r^\Delta(1 - r_q^\Delta)t_{\text{TU}} \\ &\quad + (1 - r^\Delta)(1 - r_q^\Delta)t_{\text{miss}} + r_q^\Delta r^\Delta t_{\text{QUs}} \\ &\leq T(r_q^*, r^*) \end{aligned} \quad (15)$$

where r_q^Δ and r^Δ denote the reduced hit ratios of QUs and TU, respectively; $r_q^\Delta = r_q^* - \Delta_{r_q}$ and $r^\Delta = r^* - \Delta_r$. Here, Δ_{r_q} and Δ_r denote the reductions from the optimal hit ratios r_q^* and r^* , respectively.

However, we find that

$$\begin{aligned} T(r_q^\Delta, r^\Delta) &= (r_q^* - \Delta_{r_q})t_{\text{QUs}} + (r^* - \Delta_r)(1 - r_q^* + \Delta_{r_q})t_{\text{TU}} \\ &\quad + (1 - r^* + \Delta_r)(1 - r_q^* + \Delta_{r_q})t_{\text{miss}} \\ &= r_q^*t_{\text{QUs}} + r^*(1 - r_q^*)t_{\text{TU}} + (1 - r^*)(1 - r_q^*)t_{\text{miss}} \\ &\quad - \Delta_{r_q}t_{\text{QUs}} + r^*\Delta_{r_q}t_{\text{TU}} - \Delta_r(1 - r_q^*)t_{\text{TU}} \\ &\quad + \Delta_r(1 - r_q^*)t_{\text{miss}} + \Delta_{r_q}(1 - r^*)t_{\text{miss}} \\ &\geq T(r_q^*, r^*) + \Delta_r(1 - r_q^*)(t_{\text{miss}} - t_{\text{TU}}) \\ &\quad + \Delta_{r_q}(1 - r^*)(t_{\text{miss}} - t_{\text{TU}}) \end{aligned} \quad (16)$$

Obviously

$$\begin{aligned} &\Delta_r(1 - r_q^*)(t_{\text{miss}} - t_{\text{TU}}) + \Delta_{r_q}(1 - r^*)(t_{\text{miss}} - t_{\text{TU}}) \\ &= (\Delta_r + \Delta_{r_q} - \Delta_r r_q^* - \Delta_{r_q} r^*)(t_{\text{miss}} - t_{\text{TU}}) \geq 0 \end{aligned} \quad (17)$$

Combining Formulas (16) and (17), $T(r_q^\Delta, r^\Delta) \geq T(r_q^*, r^*)$, which contradicts Formula (15). Therefore, there is no Δ_{r_q} or Δ_r that can make the delay lower. ■

The optimization problem (13) can be transformed as

$$\text{SP1-1: } \underset{\{\mu_n, \nu_n\}}{\text{minimize}} \quad \max\{1 - r_q - r\} \quad (18)$$

$$\text{s.t. } \mu_n^q, \nu_n \in \{1, 0\} \quad (18a)$$

$$\sum_{n=1}^N \mu_n^q \leq S_{\text{QUs}} \quad (18b)$$

$$\sum_{n=1}^N \nu_n \leq S_{\text{TU}} \quad (18c)$$

It can be seen from Eqs. (9), (10) and Formula (18), minimizing the missing request ratio is equal to maximize r_q and r . We use popularity as the value of each file in the knapsack problem. Therefore, when the caching strategies of QUs and TU are independent, caching S_{QUs} and S_{TU} files with the highest popularity in QUs and TU respectively is the optimal solution for the knapsack problem. This means that the NP-hard knapsack problem can be solved by designing the caching principles that minimize the missing request ratio. We next detail the two principles for cross-layer information caching.

1) HIP caching strategy: The information library consists of N different files. The file index set is defined as $\mathcal{N}$. The information that needs to be disseminated varies among different clusters. Each QU sorts the information library's files according to their popularity in each cluster. Let $\mathcal{I}_q^n$ be the sorted index set of the N files in the q^{th} cluster. The $\mathcal{I}_q^n$ can be regarded as the result of substitution and sorting of $\mathcal{N}$. The popularity of the files gradually decreases with the file index i_q^n ($i_q^n \in \mathcal{I}_q^n$), where i_q^n denotes the sorted file index in the q^{th} cluster, from small to large in each cluster. Then each QU only needs to cache files in a descending order of file popularity in the cluster until the upper limit of the caching capacity is reached.

The probability for the i_q^n information in the area covered by the q^{th} QU can be expressed as

$$P_{fq}^{i_q^n} \triangleq P_{fq}^n \quad (19)$$

The hit ratio of the q^{th} QUs is

$$r_q^{HIP} = \frac{\sum_{i_q^n=1}^{S_{QUs}} P_{fq}^{i_q^n}}{\sum_{n=1}^N P_{fq}^n} \quad (20)$$

Similarly, we use $\mathcal{I}_T^n$ to represent the sorted index set of the N files in the TU's service area, where i_T^n denotes the sorted file index in the TU's service area. The popularity of the i_T^n information in total can be expressed as

$$P_f^{i_T^n} \triangleq P_f^n \quad (21)$$

The hit ratio of the TUs is

$$r^{HIP} = \frac{\sum_{i_T^n=1}^{S_{TU}} P_f^{i_T^n}}{\sum_{n=1}^N P_f^n} \quad (22)$$

The average delay can be calculated as

$$T^{HIP} = r_q^{HIP}(1 - r_q^{HIP})t_{QUs} + r_q^{HIP}(1 - r_q^{HIP})t_{TU} + (1 - r_q^{HIP})(1 - r_q^{HIP})t_{miss} + r_q^{HIP}r_q^{HIP}t_{QUs} \quad (23)$$

When QU and TU do not cache duplicate information, HIP is the optimal caching strategy. However, there will be some waste of storage space. Next, we propose another caching strategy that does not cache files repeatedly. Although the optimality of

this strategy cannot be guaranteed, we may obtain a higher request hit ratio under special conditions.

2) DIP caching strategy: In order to ensure diversity of the cached information in the QUs and the TU, we propose a new caching strategy in which QUs and TU do not cache the same information repeatedly. From Eqs. (7) and (20), the requests missed by QUs and $\sum_{q=1}^Q (M_q - r_q^{HIP} M_q)$, which form a new cluster to design the information caching strategy on the TU. Then, the popularity of the n^{th} information in the TU becomes

$$P_f^{DIP^n} = \sum_{q=1}^Q \frac{M_q(1 - r_q^{HIP})P_{fq}^n}{M - \sum_{q=1}^Q M_q r_q^{HIP}} \triangleq P_f^{DIP i_T^n} \quad (24)$$

where $i_T^n \in \mathcal{I}_T^n$ denotes the sorted file index in the TU's service area.

The hit ratio of the TUs is

$$r^{DIP} = \frac{\sum_{i_T^n=1}^{S_{TU}} P_f^{DIP i_T^n}}{\sum_{n=1}^N P_f^{DIP^n}} \quad (25)$$

The average delay is

$$T^{DIP} = r_q^{HIP}t_{QUs} + r^{DIP}t_{TU} + (1 - r^{DIP} - r_q^{HIP})t_{miss} \quad (26)$$

Now the cross-layer caching strategy of multi-layer UAVs has been obtained. Next we optimize the 3D positions of QUs.

4 3D Deployment Strategy Design for QUs

After the caching strategy is obtained, the hit ratio of the information request and the transmission scheme are determined. Minimizing the maximum system delay can be equivalent to maximize the minimum instantaneous transmission rate between QUs and users. The optimization problem **SP2** can be formulated as

$$\mathbf{SP2}: \underset{\{u_q\}}{\text{maximize}} \quad \min R_{QUs}(u_q) \quad (27)$$

$$\text{s.t.} \quad h_{\min} \leq h_q \leq h_{\max} \quad (27a)$$

Due to the complex form of Formula (27), problem **SP2** is still a non-convex and complex problem. To solve this problem, we decompose the 3D deployment problem of QUs into optimizing the UAVs' horizontal positions and hovering heights.

Proposition 2 By introducing a lower bound of R , the 3D hovering position optimization of the UAV can be equivalently decomposed into the horizontal position optimization and the hovering height.

Proof Obviously, the $\log_2\left(1 + \frac{p_q}{\lambda\sigma^2 d^2}\right)$ is convex with respect to λ . Let $d(u_q)^2 = \|u_q - c_{m_q}\|_2^2$, based on Jensen's inequality, the lower bound of the expected throughput accordingly comes as

$$\begin{aligned} R_{\text{QUS}}(u_q) &\geq \log_2 \left(1 + \frac{p_q \left(\frac{\Pr_{\text{LoS}}(\theta_q)}{\lambda_{\text{LoS}}} + \frac{\Pr_{\text{NLoS}}(\theta_q)}{\lambda_{\text{NLoS}}} \right)}{\sigma^2 d(u_q)^2} \right) \\ &\geq \log_2 \left(1 + \frac{p_q}{\lambda(\theta_q)\sigma^2 d(u_q)^2} \right) \\ &\triangleq \bar{R}_{\text{QUS}}(u_q) \end{aligned} \quad (28)$$

where $\lambda(\theta_q) = 10^{\frac{\Pr_{\text{LoS}}(\theta_q)\eta_{\text{LoS}} + \Pr_{\text{NLoS}}(\theta_q)\eta_{\text{NLoS}} + 20\lg f + 20\lg \frac{4\pi}{c}}{10}}$.

First, we fix the flying height of QUs and optimize their horizontal position. The elevation angle of the m_q user to the q^{th} QU is

$$\theta_q = \arcsin \left(\frac{h_q}{\|u_q - c_{m_q}\|_2} \right) \quad (29)$$

Obviously when the QUs' flying height is fixed, the smaller the $\|u_q - c_{m_q}\|_2$, the larger the θ_q . Besides, it can be seen that $\lambda(\theta_q)$ is a monotonically decreasing function of θ_q . Combining Formula (28) and Eq. (29), we can get a conclusion that the instantaneous transmission rate between users and QUs is a monotonically decreasing function of $\|u_q - c_{m_q}\|_2$. Therefore, the horizontal position optimization problem of QUs can be equivalently described as

$$\text{SP2-1: minimize } \chi_{\{u_q, \chi\}} \quad (30)$$

$$\text{s.t. } \|u_q - c_{m_q}\|_2^2 \leq \chi \quad (30a)$$

where χ is an auxiliary variable for constructing a standard convex optimization problem. The problem in Formula (30) can be solved easily with the existing tools such as the CVX.

Next, we solve the optimal flying height of QUs. Using χ^* to express the optimal solution of Formula (30). The optimization problem of the flying height of QUs can be expressed as

$$\text{SP2-2: minimize}_{h_q} \lambda(\theta_q^*)(\chi^*)^2 + (h_q)^2 \quad (31)$$

$$\text{s.t. } h_{\min} \leq h_q \leq h_{\max} \quad (31a)$$

where $\theta_q^* = \arcsin(h_q / \sqrt{\chi^*})$. ■

Proposition 3 Problem SP2-2 is a quasiconvex optimization problem.

Proof Define $g(h_q) = \lg(\lambda(\theta_q)((\chi^*)^2 + (h_q)^2))$. We obtain Eq. (32), where

$$\begin{aligned} \frac{dg(h_q)}{dh_q} &= \frac{\eta_{\text{LoS}} - \eta_{\text{NLoS}}}{10} \frac{\exp(-h[\theta_q - g]) \frac{gh\sqrt{\chi^*}}{\chi^* + h_q^2}}{(1 + g \exp(-h[\theta_q - g]))^2} \\ &\quad + \frac{2h_q}{\ln 10 (\chi^* + h_q^2)} = \frac{z_1(h_q)}{z_2(h_q)}, \end{aligned} \quad (32)$$

$$\begin{aligned} z_1(h_q) &= (\eta_{\text{LoS}} - \eta_{\text{NLoS}}) \exp(-h[\theta_q - g]) gh \sqrt{\chi^*} \\ &\quad + \frac{2h_q(1 + g \exp(-h[\theta_q - g]))^2}{\ln 10}, \end{aligned} \quad (33)$$

$$z_2(h_q) = 10(1 + g \exp(-h[\theta_q - g]))^2 (\chi^* + h_q^2) \quad (34)$$

$$\begin{aligned} \frac{dz_1(h_q)}{dh_q} &= \sqrt{\chi^*} (\eta_{\text{LoS}} - \eta_{\text{NLoS}}) gh \exp(gh - h\theta_q^*) \left(-h \frac{d\theta_q^*}{dh_q} \right) \\ &\quad + \frac{2(1 + g \exp(gh - h\theta_q^*))^2}{\ln 10} \\ &\quad + \frac{4h_q(1 + g \exp(gh - h\theta_q^*))}{\ln 10} * g \exp(gh - h\theta_q^*) \left(-h \frac{d\theta_q^*}{dh_q} \right) \\ &= \sqrt{\chi^*} (\eta_{\text{LoS}} - \eta_{\text{NLoS}}) gh \exp(gh - h\theta_q^*) \left(-h \frac{d\theta_q^*}{dh_q} \right) \\ &\quad + \frac{2h_q(1 + g \exp(gh - h\theta_q^*))^2}{\ln 10} \left(-h \frac{d\theta_q^*}{dh_q} \right) \\ &\quad + \frac{2h_q(1 + g \exp(gh - h\theta_q^*))^2}{\ln 10} \left(-h \frac{d\theta_q^*}{dh_q} \right) \\ &\quad + \frac{2(1 + g \exp(gh - h\theta_q^*))^2}{\ln 10} \\ &\quad + \frac{2h_q(1 + g \exp(gh - h\theta_q^*))}{\ln 10} \left(h \frac{d\theta_q^*}{dh_q} \right) \\ &= z_1(h_q) + \frac{2h_q(1 + g \exp(gh - h\theta_q^*))^2}{\ln 10} \left(-h \frac{d\theta_q^*}{dh_q} \right) \\ &\quad + \frac{2(1 + g \exp(gh - h\theta_q^*))^2}{\ln 10} \\ &\quad + \frac{2h_q(1 + g \exp(gh - h\theta_q^*))}{\ln 10} \left(h \frac{d\theta_q^*}{dh_q} \right) \end{aligned} \quad (35)$$

when $z_1(h_q) = 0$

$$\frac{dz_1(h_q)}{dh_q} = \frac{2h_q(1 + g \exp(gh - h\theta_q^*))^2}{\ln 10} \left(-h \frac{d\theta_q^*}{dh_q} \right)$$

$$\begin{aligned}
& + \frac{2(1 + g \exp(gh - h\theta_q^*))^2}{\ln 10} \\
& + \frac{2h_q(1 + g \exp(gh - h\theta_q^*))}{\ln 10} \left(h \frac{d\theta_q^*}{dh_q} \right) \\
& = \frac{2(1 + g \exp(gh - h\theta_q^*))^2}{\ln 10} \left(1 - \frac{\sqrt{\chi^*} h_q h}{\chi^* + h_q^2} \right) \\
& + \frac{2h_q(1 + g \exp(gh - h\theta_q^*))}{\ln 10} \left(h \frac{d\theta_q^*}{dh_q} \right) \\
& \geq \frac{2(1 + g \exp(gh - h\theta_q^*))^2}{\ln 10} \left(1 - \frac{h}{2} \right) \\
& + \frac{2h_q(1 + g \exp(gh - h\theta_q^*))}{\ln 10} \left(h \frac{d\theta_q^*}{dh_q} \right) \geq 0 \quad (36)
\end{aligned}$$

$$\frac{d^2 g(h_q)}{dh_q^2} = \frac{dz_1(h_q)z_2(h_q) - z_1(h_q)dz_2(h_q)}{z_2(h_q)^2} \quad (37)$$

From Eq. (34) and Formula (36), we know that $z_2(h_q) > 0$, and $\frac{dz_1(h_q)}{dh_q} \geq 0$ when $z_1(h_q) = 0$. It means that when $\frac{dg(h_q)}{dh_q} = 0$, $\frac{d^2 g(h_q)}{dh_q^2} \geq 0$. The function $g(h_q)$ is therefore quasiconvex. The problem **SP2-2** is a quasiconvex optimization problem. ■

According to Proposition 3, problem **(SP2-2)** is a quasiconvex optimization problem, which can be transformed into a set of feasible solution problems and solved by the CVX. Now the optimal solution for the 3D deployment of QUs has been obtained.

5 Simulation Result

This section simulates the performance of multi-layer UAVs equipped with proactive caching technology in reducing the information delivery delay. The related system parameters for simulation are summarized in Table 1.

Let the number of users be $M = 500$. To simplify the

Table 1 Parameter values of the system setting for simulations.

Parameter	Value
Number of QUs	$Q = 9$
Power spectral density of AWGN	-174 dBm/Hz
Communication bandwidth	1 MHz
Number of users	$M = 500$
Additional loss on the LoS channel	$\eta_{\text{LoS}} = 1.6$ dB
Additional loss on the NLoS channel	$\eta_{\text{NLoS}} = 23$ dB
LoS occurrence probability parameter	$g = 9.6, h = 0.28$
Files size	$\gamma = 5$ Mb

experimental setup, we assume that all users adhere to a zero-inflated generalized Poisson distribution and are distributed within a square region of side length 1200 meters. The region is divided into $Q = 9$ clusters. The flying altitude of TU is set to $h_{\text{TU}} = 400$ m. The 3D deployment positions of QUs are obtained using the method proposed in Section 4. We first observe the impact of the size of the information library on the transmission delay when the caching capacities of QUs and TU are fixed at $S_{\text{QUs}} = 100$ and $S_{\text{TU}} = 200$, respectively. The Zipf exponent is set as $\alpha = 1$. The simulation results are shown in Fig. 2.

Figure 2 indicates that when the caching capacities of QUs and TU are limited, the delay in information delivery is positively correlated with the total number of files in the information library. The comparison schemes include the dominant most popular content (MPC), largest content diversity (LCD), and random caching (RC) techniques^[24, 25]. The latency increases with N in the HIP, DIP, and the comparison methods. Both HIP and DIP significantly outperform the three comparison methods. DIP performs better when the number of files is small because the DIP cache enhances information diversity on QUs, improving the hit ratio of user requests. HIP replicates popular files, which makes it slightly inferior to DIP in this scenario. As the number of files N increases, the weight of each individual file gradually decreases, reducing the negative impact of repeated caching. Caching files with higher popularity is more effective in minimizing information transmission delays, causing HIP's delay to eventually decrease below that of DIP.

Next, we examine the correlation between system delay and the information popularity distribution parameter α . The caching capacity of QUs is given as $S_{\text{QUs}} = 100$. The TU's caching capacity is $S_{\text{TU}} = 200$. Figures 3 and 4 illustrate the effect of α on system latency for both large and small information library

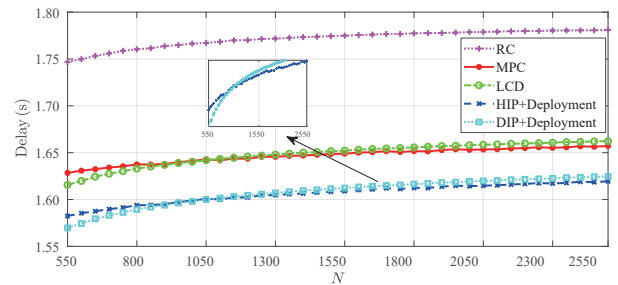


Fig. 2 Relationship between the information delivery delay and the number of files.

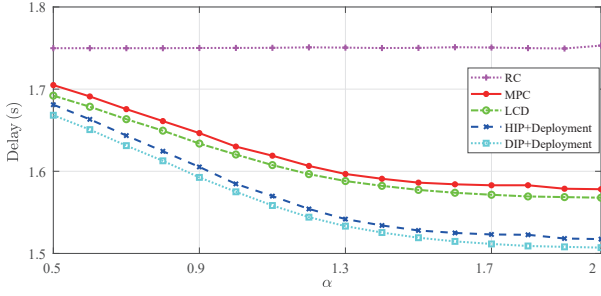


Fig. 3 Relationship between the information delivery delay and α , $N = 600$.

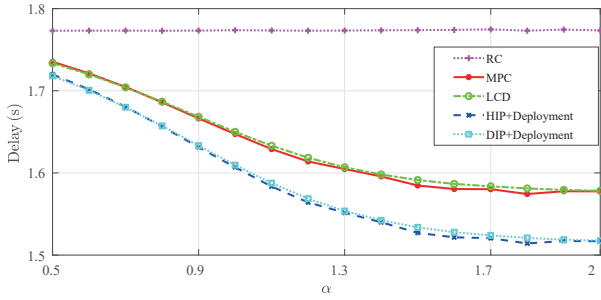


Fig. 4 Relationship between the information delivery delay and α , $N = 1400$.

sizes, respectively. The overall file count in Fig. 3 is $N = 600$, while in Fig. 4, it is $N = 1400$.

Figures 3 and 4 illustrate that α exerts negligible influence on the transmission latency of the RC approach. However, as α increases, the latency of HIP and DIP has progressively decreased. The latency of both the HIP and DIP is inferior to that of the MPC and LCD. When N is small, the efficiency of DIP surpasses that of HIP. When N is large, Fig. 4 illustrates that for relatively small values of α , the delay associated with the DIP strategy is inferior to that of the HIP technique. Nonetheless, when α is large, the HIP exhibits a reduced latency compared to the DIP. This can be explained as a higher concentration of user requests correlating with the increased applicability of the HIP, determined by the prioritization of popularity as a proactive caching approach. Conversely, when users seek scattered information, storing diversified contents is more efficient in reducing information transmission latency.

We jointly examine the influence of α and N on the system delay as depicted in Fig. 5. Consistent with the findings illustrated in Figs. 2, 3, and 4, Fig. 5 demonstrates that when the total number of files N and α are small, the delay of the DIP is at the lowest level. As α or N increases, the performance of HIP

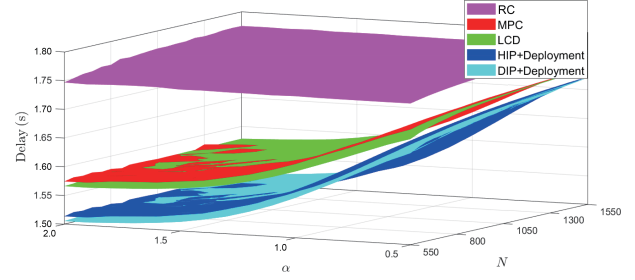


Fig. 5 Effect of the number of files and alpha on the information delivery delay.

progressively surpasses that of DIP, establishing it as a superior proactive caching approach.

We subsequently examine the effect of the caching capacity of QUs and TU on the delay. Let $N = 600$ and $\alpha = 0.5$. Figure 6 illustrates that when the total number of files is small and user requests are scattered, the information delivery delay of the DIP is minimal. As the caching capacity of QUs expands, the information transmission delay of the five caching strategies all decreases.

In Fig. 7, we examine the impact of TU's caching capacity on information delivery delay when $N = 600$, $\alpha = 0.5$. As illustrated in Fig. 6, the information transmission delay of all caching strategies progressively decreases with the growth in TU's caching capacity. The delays of the DIP and HIP are less significant than those of the comparison strategies,

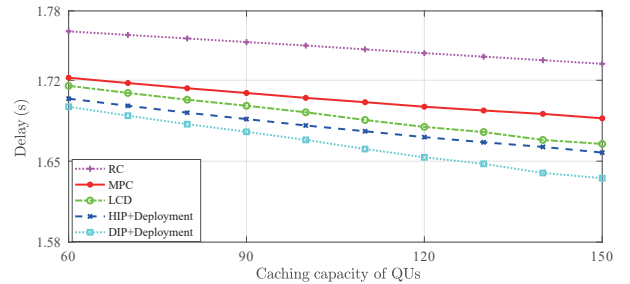


Fig. 6 Relationship between the information delivery delay and S_{QU} , $N = 600$.

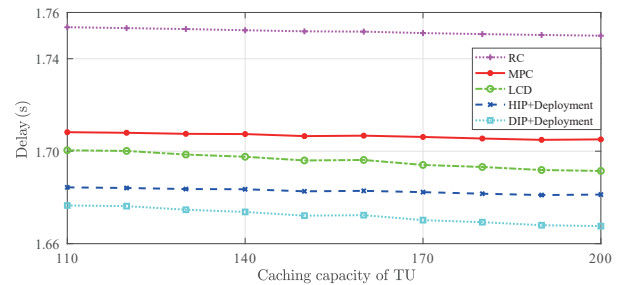


Fig. 7 Relationship between the information delivery delay and S_{TU} , $N = 600$.

with the DIP being the most beneficial caching strategy in this specific case.

Let $N = 1400$, $\alpha = 1.5$. Figures 8 and 9 illustrate the impact of the caching capacity of QUs and TUs on information delivery delay. As illustrated in Figs. 6 and 7, the DIP and HIP exhibit significantly reduced delays compared to the comparison strategies. Figures 8 and 9 demonstrate that when N and α are large, an increase in the caching capacity of QUs and TUs results in a decrease of the information delivery delay. In this case, the HIP exhibits the lowest delay, while the DIP performs inferiorly compared to the HIP.

The conclusion can be summarized as follows: When user requests concentrate, the popularity of a single file considerably influences the information request hit ratio, making HIP an efficient proactive caching approach. On the contrary, DIP is a better caching strategy.

6 Conclusion

This paper proposes a multilayered, heterogeneous UAV-assisted framework for emergency catastrophe mitigation, designed to minimize information dissemination delay. To address the joint design of cross-layer information caching and QUs' 3D deployment, we decouple this complex problem into two subproblems: One focused on cross-layer

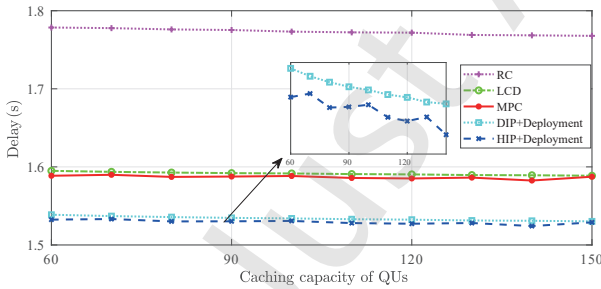


Fig. 8 Relationship between the information delivery delay and S_{QU} , $N = 1400$.

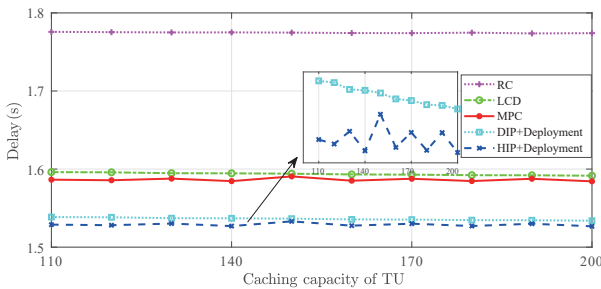


Fig. 9 Relationship between the information delivery delay and S_{TU} , $N = 1400$.

information caching and the other on QUs' 3D deployment. To solve the NP-hard information caching problem, we prove that minimizing delay is equivalent to minimizing the missing request ratio. Based on this insight, we develop two novel caching principles, HIP and DIP, to derive optimal caching strategies. The 3D deployment problem is then divided into its horizontal and vertical dimensions. The optimization of the QUs' horizontal positions is formulated as a convex optimization problem, while the corresponding hovering height optimization is proven to be quasiconvex. Both subproblems are efficiently solved using the CVX toolbox. Simulation results demonstrate that our proposed proactive caching strategies, when combined with the optimized UAV deployment, can significantly reduce information dissemination delay.

Acknowledgment

This work was financially supported by the National Natural Science Foundation of China (Nos. 92467103 and 92367302), the Henan Provincial Department of Transportation (No. 2023-3-2), the Henan Provincial Science and Technology Research Project (Nos. 252102210176 and 252102210232), and the Foundation of Henan Educational Committee (Nos. 25A510015 and 25A520008).

Declaration of competing interest

The authors have no competing interests to declare that are relevant to the content of this article.

References

- [1] Y. Guan, S. Zou, H. Peng, W. Ni, Y. Sun, and H. Gao, Cooperative UAV trajectory design for disaster area emergency communications: A multiagent PPO method, *IEEE Internet Things J.*, vol. 11, no. 5, pp. 8848–8859, 2024.
- [2] D. Wang, Y. Cao, K. -Y. Lam, Y. Hu, and O. Kaiwartya, Authentication and key agreement based on three factors and PUF for UAV-assisted post-disaster emergency communication, *IEEE Internet Things J.*, vol. 11, no. 11, pp. 20457–20472, 2024.
- [3] M. H. Bashir, M. Ahmad, D. R. Rizvi, and A. A. A. El-Latif, Efficient CNN-based disaster events classification using UAV-aided images for emergency response application, *Neural Comput. Appl.*, vol. 36, no. 18, pp. 10599–10612, 2024.
- [4] T. Bose, N. Chatur, and A. Adhya, Energy-efficient disjoint multi-UAV deployment for emergency communication in disaster-affected areas, in *Proc. IEEE*

- 21st India Council Int. Conf. (INDICON)., Kharagpur, India, 2024, pp. 1–6.
- [5] C. Wang, D. Deng, L. Xu, and W. Wang, Resource scheduling based on deep reinforcement learning in UAV assisted emergency communication networks, *IEEE Trans. Commun.*, vol. 70, no. 6, pp. 3834–3848, 2022.
- [6] Y. Zhang, Y. Hong, M. Guizani, S. Wu, P. Zhang, and R. Liu, A multi-layer information dissemination model and interference optimization strategy for communication networks in disaster areas, *IEEE Trans. Veh. Technol.*, vol. 73, no. 1, pp. 1239–1252, 2024.
- [7] Y. Li, Y. Bi, J. Wang, Z. Li, H. Zhang, and P. Zhang, Unmanned aerial vehicle assisted communication: Applications, challenges, and future outlook, *Cluster Comput.*, vol. 27, no. 9, pp. 13187–13202, 2024.
- [8] J. Li, C. Yi, J. Chen, Y. Shi, T. Zhang, X. Li, R. Wang, and K. Zhu, A reinforcement learning-based stochastic game for energy-efficient UAV swarm-assisted MEC with dynamic clustering and scheduling, *IEEE Trans. Green Commun. Netw.*, vol. 9, no. 1, pp. 255–270, 2025.
- [9] N. L. Prasad and B. Ramkumar, 3-D deployment and trajectory planning for relay based UAV assisted cooperative communication for emergency scenarios using Dijkstra's algorithm, *IEEE Trans. Veh. Technol.*, vol. 72, no. 4, pp. 5049–5063, 2023.
- [10] A. Hassan, S. Javed, R. Ahmad, W. Ahmed, A. Kiani, and M. M. Alam, Optimizing the deployment of multi-UAVs in emergency response networks, in *Proc. 19th Biennial Baltic Electronics Conf. (BEC)*, Tallinn, Estonia, 2024, pp. 1–6.
- [11] J. Wu, Q. Chen, H. Jiang, H. Wang, Y. Xie, W. Xu, P. Zhou, Z. Xu, L. Chen, B. Li et al., Joint power and coverage control of massive UAVs in post-disaster emergency networks: An aggregative game-theoretic learning approach, *IEEE Trans. Netw. Sci. Eng.*, vol. 11, no. 4, pp. 3782–3799, 2024.
- [12] J. Miao, S. Bai, S. Mumtaz, Q. Zhang, and J. Mu, Utility-oriented optimization for video streaming in UAV-aided MEC network: A DRL approach, *IEEE Trans. Green Commun. Netw.*, vol. 8, no. 2, pp. 878–889, 2024.
- [13] K. Shi, B. Li, Y. Wang, B. Ran, P. Kong, X. Huang, D. Ma, and F. Liang, Content popularity prediction with integrated features in UAV-assisted wireless caching networks, in *Proc. IEEE 22nd Int. Conf. Communication Technology (ICCT)*, Nanjing, China, 2022, pp. 573–577.
- [14] İ. Demirci and Ö. Korçak, Proactive edge caching with LSTM-based popularity prediction, in *Proc. 7th Int. Balkan Conf. Communications and Networking (BalkanCom)*, Ljubljana, Slovenia, 2024, pp. 218–223.
- [15] G. T. Maale, G. Sun, N. A. E. Kuadey, T. Kwantwi, R. Ou, and G. Liu, DeepFESL: Deep federated echo state learning-based proactive content caching in UAV-assisted networks, *IEEE Trans. Veh. Technol.*, vol. 72, no. 9, pp. 12208–12220, 2023.
- [16] X. Li, J. Liu, N. Zhao, and X. Wang, UAV-assisted edge caching under uncertain demand: A data-driven distributionally robust joint strategy, *IEEE Trans. Commun.*, vol. 70, no. 5, pp. 3499–3511, 2022.
- [17] G. Sun, Y. Wang, Z. Sun, Q. Wu, J. Kang, D. Niyato, and V. C. M. Leung, Multi-objective optimization for multi-UAV-assisted mobile edge computing, *IEEE Trans. Mob. Comput.*, vol. 23, no. 12, pp. 14803–14820, 2024.
- [18] N. Lin, T. Wu, L. Zhao, A. Hawbani, S. Wan, and M. Guizani, An energy effective RIS-assisted multi-UAV coverage scheme for fairness-aware ground terminals, *IEEE Trans. Green Commun. Netw.*, vol. 9, no. 1, pp. 164–176, 2025.
- [19] C. Wu, Z. Xu, X. He, Q. Lou, Y. Xia, and S. Huang, Proactive caching with distributed deep reinforcement learning in 6G cloud-edge collaboration computing, *IEEE Trans. Parallel Distrib. Syst.*, vol. 35, no. 8, pp. 1387–1399, 2024.
- [20] Z. Qian, G. Li, T. Qi, C. Dai, and J. Hu, Proactive video caching based on federated learning and implicit feedback in mobile edge computing, in *Proc. GLOBECOM 2024 - 2024 IEEE Global Communications Conf.*, Cape Town, South Africa, 2024, pp. 5319–5324.
- [21] Q. Li, P. Si, Y. Zhang, J. Wang, D. Zhang, and F. R. Yu, UAV altitude, relay selection, and user association optimization for cooperative relay-transmission in UAV-IRS-based THz networks, *IEEE Trans. Green Commun. Netw.*, vol. 8, no. 2, pp. 815–826, 2024.
- [22] M. Ali Jamshed, F. Ayaz, A. Kaushik, C. Fischione, and M. Ur-Rehman, Green UAV-enabled Internet-of-Things network with AI-assisted NOMA for disaster management, in *Proc. IEEE 35th Int. Symp. on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Valencia, Spain, 2024, pp. 1–6.
- [23] R. Jin, L. Yang, and H. Zhang, Performance analysis of temporal correlation in finite-area UAV networks with LoS/NLoS, in *Proc. IEEE Wireless Communications and Networking Conf. (WCNC)*, Seoul, Korea (South), 2020, pp. 1–6.
- [24] G. Sun, Y. Wang, Z. Sun, Q. Wu, J. Kang, D. Niyato, and V. C. M. Leung, Multi-objective optimization for multi-UAV-assisted mobile edge computing, arXiv preprint arXiv: 2404.15292, 2024.
- [25] Q. Yang, H. Deng, and L. Wang, Partitioning caching: Keep popular and diverse content in ICN in-network caching, *Int. J. Innov. Comput. Inf. Control*, vol. 15, no. 5, pp. 1655–1672, 2019.



and federated learning.

Lingfeng Shen received the PhD degree in information and communication engineering from Zhengzhou University, China, in 2022. He is currently with School of Software, Henan University, China. His current research interests include unmanned aerial vehicle (UAV) communications, Internet of Things (IoT),



Ming Li received the BEng degree in Internet of Things engineering from Henan University of Animal Husbandry and Economy, China, in 2024. He is currently pursuing the MS degree in cyberspace security at Henan University, China. His research interest includes low-cost localization in IoT.



Jiaxin Xu is currently an undergraduate student in network engineering at Henan University, China. Her current research interests include UAV communications.



Lingyun Li received the BEng degree in Internet of Things engineering from Huanghuai University, China, in 2023. He is currently pursuing the MS degree in network and information security at Henan University, China. His research interests include localization, IoT, and joint optimization.



Guanghui Wang received the PhD degree in information networks from Nanjing University of Posts and Telecommunications, China, in 2019. He is currently an associate professor with School of Software, Henan University, China. He was a visiting research student with Department of Computer Science, University of Victoria, Canada, from 2016 to 2018. His current research interests include blockchain, federated learning, security and privacy preservation, and IoT communications.



Fang Zuo received the PhD degree in computer application technology from East China Normal University, China, in 2013. He is currently a professor with School of Software, Henan University, China. His research interests include computer networks, peer-to-peer networks, network optimization theory, algorithmic game theory, joint optimization, and mobile opportunistic networks.



Yansong Liu is currently a master's student studying software engineering at RWTH Aachen University, Germany. His research interests include UAV communications.



Xin He received the PhD degree in computer science from Xi'an Jiaotong University, China, in 2011. He is currently a professor with School of Software, Henan University, China. His research interests include computer networks, opportunistic networks, mobile computing, cloud computing, big data processing, and trajectory optimization.