IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 1 April 2024, accepted 15 April 2024, date of publication 24 April 2024, date of current version 3 May 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3392952

RESEARCH ARTICLE

Joint Caching, Communication, and Trajectory Optimization in Air-Ground Integrated Wireless Networks With Multiple UAVs and Multiple BSs

YI-MO LIN, SHIN-PING HUANG, MING-CHUN LEE, (Member, IEEE), AND YUE-RONG HUANG, (Student Member, IEEE)
Institute of Communications Engineering, National Yang Ming Chiao Tung University, Hsinchu 30010, Taiwan

Corresponding author: Ming-Chun Lee (mingchunlee@nycu.edu.tw)

This work was supported in part by the National Science and Technology Council (NSTC) of Taiwan under Grant 112-2221-E-A49-077-MY2 and Grant 112-2218-E-A49-020, and in part by the Industrial Technology Research Institute (ITRI) in Taiwan under Grant GD-SC-112-A000-02.

ABSTRACT Using unmanned aerial vehicles (UAVs) to improve the network performance has been widely discussed due to their high agility for deployment and better line-of-sight (LoS) channels. To further improve the efficiency of using UAVs to improve the network, this paper investigates the optimization of cache-enabled air-ground integrated networks where multiple UAVs and base stations (BSs) jointly serve users. Aiming to maximize the minimum user rate of the network, we first formulate two optimization problems. The first problem considers jointly optimizing the user scheduling and association, UAV power allocation, bandwidth allocation, and UAV trajectories when the caching placement is pre-determined, and the second problem further jointly optimizes the caching placement. To solve these two problems, we propose to decompose them into sub-problems, and then solve the sub-problems iteratively to obtain the solutions. The solution approaches of each sub-problem are provided. In addition, convergences of the proposed approaches are analyzed. Simulation results show that our proposed approaches perform well and outperform the reference schemes.

INDEX TERMS UAV communications, cache-aided wireless networks, joint caching and communication optimization, UAV trajectory optimization.

I. INTRODUCTION

Due to the emerging applications, e.g., high-definition video streaming, augmented/virtual reality, and artificial intelligence (AI)-aided applications, wireless data traffic has increased significantly and is expected to continue its increase [1]. However, improving the wireless network to meet such challenge under the conventional network framework is difficult. Therefore, many different new technologies have been proposed, e.g., millimeter-wave, heterogeneous networks, edge-caching and edge-computing, and non-terrestrial networks [2]. Among different technologies, the use of unmanned aerial vehicles (UAVs) to assist wireless networks has been deemed promising, as UAVs can serve as flexible and agile wireless moving platforms to provide services to users and improve the network performance [3], [4], [5], [6], [7].

Although UAVs can provide services to users by establishing high-quality user-UAV links, this is based on having the sufficiently good wireless backhaul to help provide connections between the UAVs and base stations (BSs) [8]. However, to establish high-quality wireless backhaul commonly requires additional bandwidth and infrastructure to be invested into the networks, which is costly. On the other hand, wireless caching has been widely discussed due to its ability to trading low-cost storage against off the

The associate editor coordinating the review of this manuscript and approving it for publication was Mohammad S. Khan.

high-cost bandwidth [9], [10], where the fundamental idea of caching at the wireless edge is to first use cheap storage to cache popular contents close to the users, and then whenever the contents are needed, the users can directly access the contents without resorting to the use of the backhaul and the database on the cloud. It has been shown in the literature that using caching at the wireless edge can significantly improve network performance [10], [11], [12], [13].

Observing the benefits of introducing caching into the wireless network, to relieve the burden of using the wireless backhaul by UAVs, caching at the UAVs has started to draw attentions [8], [14]. The idea of cache-aided UAV network is again to first let UAVs cache popular contents, and then when users request the contents, the UAVs can immediately provide the contents without using the high-cost wireless backhaul transmissions. This thus significantly improves the UAV network performance in practice and makes the UAV networks more practically appropriate in many different situations. Following this idea, cache-aided UAV networks have been widely discussed in recent years [8], [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27].

A. RELATED LITERATURE REVIEW

To improve the cache-aided UAV networks, optimization and design approaches for both caching and communication have been investigated [15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27], [28], [29]. Reference [15] considered that UAVs with caching can act as static aerial BSs to jointly serve users with the terrestrial BSs so that the minimum required rate of users associated to UAVs can be satisfied. Also in [16], considering UAVs to act as static aerial BSs, an approach that maximizes quality of experiences of users was proposed by jointly optimizing the UAV deployment, caching policy, and user association. With the same idea that UAVs can serve as aerial BSs, [17] proposed a scheme used in UAV networks, where UAVs can proactively transmit files to the selected ground nodes (GN) such that the selected GNs can first cache those files transmitted by UAVs, and then share them with neighboring users by using device-to-device (D2D) communications. On the other hand, an approach that jointly optimize the UAV deployment and caching content placement was proposed in [18] to maximize the throughput of multimedia data in IoT networks. In addition, in [19], when given the UAV deployment, the caching policies of UAVs and small-cell BSs as well as the interference management approach were investigated and designed for transmissions with better security. In [20], also considering that locations of UAVs are given, an approach based on the predicted user request distribution was proposed to optimize the resource allocation and cache placement. Although [15], [16], [17], [18], [19], and [20] studied caching and communication policies in cache-aided UAV networks, they assume that UAVs are with fixed locations, and thus the agility and moving ability of UAVs were not fully exploited.

To effectively exploit the agility and moving ability of UAVs, in [21], [22], [23], [24], [25], [26], [27], [28], and [29], trajectories of UAVs are jointly optimized with caching and communication. Specifically, in [21], when considering that the users can cache files, a joint UAV trajectory and scheduling optimization approach was proposed to maximize the minimum secrecy rate of users. Reference [22] considered using UAVs with caching to assist vehicle networks, and a joint caching, trajectory, and scheduling optimization approach was proposed via using the machine learning-based and meta-heuristic algorithms. However, the power and bandwidth optimization was not considered by it. In [23], with the use of non-orthogonal multiple access (NOMA), a reinforcement learning-based approach was proposed to jointly optimize the caching, user scheduling, and power allocation without the optimization of the trajectory of the UAV. Also considering the use of NOMA, [24] proposed an approach to jointly optimize the user association, power allocation, UAV trajectories, and caching placement without considering the influence of possible user scheduling and links between BSs and users. Also, [28] proposed an approach to jointly optimize the UAV trajectories, caching placement and transmit power without considering the scheduling and bandwidth allocation among UAVs and BSs. To deal with the environmental uncertainties, a deep reinforcement learning approach was proposed in [29] to jointly design the caching, scheduling, and UAVs power allocation, and trajectories for latency minimization. Different from the above papers that focused on using UAVs to directly serve users, [25], [26], and [27] worked on cache-aided UAV-assisted D2D networks and investigated the caching policy optimization for UAVs and devices as well as the optimization of UAV trajectories.

B. CONTRIBUTIONS

Although there exist studies for cache-aided UAV networks, to the best of our knowledge, the studies are still incomplete. Especially, we see that there is no existing study that considers cache-aided UAV networks with multiple UAVs and multiple BSs, and then studies the joint caching, communication, and trajectory optimization, where the UAV-user links and BS-user links are jointly considered. In fact, existing works with UAV trajectory optimization consider either a single UAV with multiple BSs or a single BS with multiple UAVs and do not consider the interactions between UAV-user links and BS-user links. In addition, even for the cases of single UAV and multiple BSs or single BS and multiple UAVs, the joint optimization of caching, communication, and trajectory for networks where users are served by UAVs and BSs has not been comprehensively studied. We note that [21], [23], [24] worked with some specific schemes or purposes, e.g., NOMA and security, and [25], [26], [27] worked on the UAV-aided D2D networks. The only exception is [22]. However, it does not consider the optimization of power and bandwidth allocation and the association and scheduling of users. Finally, most studies on the joint caching, communication, and trajectory optimization were focusing on optimizing the overall network performance, e.g., the sum-rate of users, without considering the minimum rate among all users except for [28]. However, the cooperation between UAV and ground BSs was not considered in [28], indicating that the fairness between users has not been well addressed in the cache-aided UAV networks Therefore, to resolve the aforementioned issues, we in this paper aim to investigate the joint caching, communication, and trajectory optimization that aims to maximize the minimum user rate of the cache-aided UAV networks where multiple UAVs and multiple BSs can jointly serve the users. We stress that such air-ground integrated networks can improve the network performance by bring contents closer to users and harmonizing the resources among UAVs and BS.

In this paper, we consider the cache-aided UAV network with multiple UAVs and multiple BSs that jointly serve the users. We consider maximizing the minimum user rate by jointly optimizing the caching placement, user scheduling and association with the UAVs and BSs, UAV power allocation, bandwidth allocation among UAVs and BSs, and UAV trajectories. Furthermore, unlike most existing works that assume a UAV to take off and land at the same location (or BS), we allow UAVs to take off and land at different locations, making the trajectories more flexibly optimized. To develop the optimization approaches, we first formulate two problems, namely, the cache-aware joint communication and trajectory (CJCT) optimization problem and the joint caching, communication, and trajectory (JCCT) optimization problem, where the difference between them lies in that the CJCT optimization problem considers that the caching placement is pre-determined and cannot be optimized, corresponding to the cases that the update of caching can only happen infrequently in the off-peak hours for minimizing the update cost. On the other hand, the JCCT optimization problem considers that the caching placement can be optimized, corresponding to the cases that the sufficiently strong backhaul is equipped by the BSs at which the UAVs stop, and thus the update of caching can be more frequently conducted.

To develop the optimization approaches, we start with solving the CJCT problem. By the observation that the CJCT problem can be decomposed into sub-problems via using the concept of the block coordinate descent (BCD) method [30], we decompose the problem into the user scheduling and association sub-problem, UAV power allocation sub-problem, system bandwidth allocation sub-problem, and UAV trajectory optimization sub-problem. Then, the overall optimization approach is to solve the sub-problems iteratively until convergence, where the solution approaches of each sub-problem are also proposed via exploiting the techniques of linear integer programming and the successive convex approximation (SCA) [31]. Based on the optimization framework for solving the CJCT problem, we develop the approach for solving the JCCT problem. Specifically, since the only difference between the CJCT and JCCT problems is that the JCCT problem needs to additionally optimize the caching placement, we again decompose the JCCT problem using the BCD method, while the user scheduling and association sub-problem obtained from the decomposition is in this case changed to the user scheduling, association, and caching placement sub-problem, and other sub-problems remain the same. Consequently, we discuss the approach for solving the user scheduling, association, and caching placement sub-problem, and then obtain the overall optimization approach for the JCCT problem. The convergence analysis of the proposed approaches for solving the CJCT and JCCT problems is conducted, showing that our approaches can monotonically increase the minimum user rate until convergence. In addition, computer simulations are also conducted to evaluate our proposed approaches. Results show that our approaches perform well and significantly outperform all the reference schemes.

C. ORGANIZATION

The remainder of this paper is organized as follows. In Sec. II, the network model considered in this paper is provided. Sec. III discusses the proposed CJCT and JCCT problem formulations. The optimization approaches for the CJCT and JCCT problems are developed and presented in Sec. IV. Then, the convergence analysis of our proposed approaches is provided in Sec. V. Simulations results are provided in Sec. VI. Finally, the conclusions and references are provided at the end of this paper.

II. NETWORK MODEL

In this section, the network model for the cache-aided UAV networks considered in this paper is presented.

A. UAV-ASSISTED NETWORK MODEL

In this paper, we consider a UAV-assisted wireless network with K base stations (BSs) and M UAVs, where each UAV has the storage that can cache files. We assume that K BSs act as the content servers and charging stations for UAVs to recharge their batteries and refresh their storages. We assume that there are U users in the network for requesting services and assume that users request files from a library containing F files. We assume that the BSs have all files, and thus user requests can always be served by the BSs as long as the corresponding BS-user links can be established. On the other hand, the UAVs can act as aerial BSs that provide services to users. However, due to the high cost (or lack) of the backhaul, UAVs can provide services to requests only if the requested files are cached by the UAVs.

We assume that a frequency reuse scheme is used by the BSs so that the interference between BSs can be ignored. In addition, we consider that UAVs shared the same bandwidth for communications. As a result, interference between UAVs could exist when they transmit signals to users at the same time-slot. We assume that the UAVs and BSs use different bands to serve users, and thus the interference between BSs and UAVs does not exist. We assume that a service period consists of N time-slots, and the UAVs can fly out of the BSs to provide services during the service period. However, the UAVs need to come back to the BSs for recharging and/or refreshing their storages at the end of the service period. Note that different from some studies in the literature that assume that the UAVs must come back to their original BSs, we assume that the UAVs can land at any BSs at the end of the service period as long as the BSs have enough parking space for the UAVs to land.

To represent the locations of BSs, UAVs, and users, the Cartesian coordinate system is adopted. We assume that the BSs are located on the ground, and thus the vertical coordinate value is 0. Then, the remaining two-dimensional (2D) coordinates of BS k is expressed by $\mathrm{q}_{\mathrm{BS},k} = \left[x_{\mathrm{BS},k}, y_{\mathrm{BS},k}\right]^T$. We assume also users are located on the ground within a given area during the service period. Thus, the 2D coordinates of user u at time-slot n is expressed by $\mathrm{q}_{\mathrm{US},u}[n] = \left[x_{\mathrm{US},u}[n], y_{\mathrm{US},u}[n]\right]^T$ and the vertical coordinate is with value equal to 0. Then, due to certain flying policy, we assume that all UAVs fly at a fixed flight altitude [6], [22], denoted as $H > 0$. Then, the 2D coordinates of UAV m at time-slot n is expressed as $\mathrm{q}_m[n] = [x_m[n], y_m[n]]^T$. We assume that each UAV takes off from a BS at the beginning of the service period. Therefore, the initial values of the 2D coordinates of UAVs must be corresponding to the values of the 2D coordinates of some BSs. It follows that when considering that UAV m is from BS k at the beginning, we have the following UAV trajectory constraints:

$$\sum_{m=1}^{M} \delta_{m,k} = M_{\max}, \forall k; \quad \sum_{k=1}^{K} \delta_{m,k} = 1, \forall m,$$

$$\mathrm{q}_m[0] = \mathrm{q}_{\mathrm{BS},k}, \forall m, k; \quad \mathrm{q}_m[N] = \sum_{k=1}^{K} \delta_{m,k}\mathrm{q}_{\mathrm{BS},k}, \forall m, k,$$

where $\delta_{m,k} \in \{0, 1\}$ is a binary indicator to indicate whether UAV m will land at BS k at the end of the service period and $M_{\max}$ is the maximum number of UAVs that can land at BS k. Note that the constraint $\sum_{k=1}^{K} \delta_{m,k} = 1$ is to indicate that a UAV can only land in a BS. We assume that the maximum moving distance of a UAV for each time-slot is $d_{\max}$. Thus, for UAVs, we have the following moving distance constraints to satisfy in all the time-slots:

$$\|\mathrm{q}_m[n+1] - \mathrm{q}_m[n]\|^2 \le d_{\max}^2, \quad \forall m, n.$$

B. COMMUNICATION MODEL

By the aforementioned network model, the distance between user u and UAV m and the distance between user u and BS k at time-slot n are given by

$$d_{u,m}[n] = \sqrt{\left\|\mathrm{q}_m[n] - \mathrm{q}_{\mathrm{US},u}[n]\right\|^2 + H^2}$$

and

$$d_{u,k}[n] = \sqrt{\left\|\mathrm{q}_{\mathrm{BS},k} - \mathrm{q}_{\mathrm{US},u}[n]\right\|^2},$$

respectively. We assume that the channels between UAVs and users are dominated by the LoS path without small-scale fading [32]. Then, the channel power gain between UAV m and user u at time-slot n is given as [6]:

$$h_{u,m}[n] = \rho_0 d_{u,m}^{-2},$$

where ρ_0 denotes the reference channel gain at the distance $d_0 = 1$ m. We consider a regular terrestrial large-scale channel model for communication links between BSs and users. Therefore, the large-scale path loss between BS k and user u is given as:

$$PL_{u,k}(d)[\mathrm{dB}] = 20\log_{10}\left(\frac{4\pi}{\lambda_c}d_0\right) + 10\epsilon\log_{10}(d_{u,k}) + \Xi,$$

where ϵ is the path-loss exponent and Ξ is the shadowing effect coefficient. Note that since Ξ is to represent the large-scale shadowing effect, the shadowing effect for the same user-BS link is invariant as the locations of users and BSs are fixed. We assume that the small-scale fading in terrestrial links can again be mitigated by some diversity scheme, and thus the small-scale fading effect is ignored [33]. We denote $p_m[n]$ as the downlink transmit power of UAV m in time-slot n, where $0 \le p_m[n] \le P_{\max}$. The signal-to-interference-plus-noise ratio (SINR) for user u when served by UAV m is then expressed as:[1]

$$\gamma_{u,m}[n] = \frac{p_m[n]h_{u,m}[n]}{\sum_{j=1,j\neq M}^{M} p_j[n]h_{u,j}[n] + \sigma^2},$$

where σ^2 is the noise power at the receiver. Also, we denote p_k as the downlink transmit power of BS k. It follows that the signal-to-noise ratio (SNR) for user u when served by BS k is:

$$\gamma_{u,k}[n] = \frac{p_k \cdot 10^{\frac{-PL_{u,k}(d)}{10}}}{\sigma^2}.$$

C. CACHING AND REQUEST MODELS

In this paper, we consider that UAVs can serve users only if the requested files by users are cached by UAVs. We then denote $r_{u,f}$ and $c_{m,f}$ as the binary indicators to indicate whether user u requests file f and whether UAV m cache file f, respectively, i.e.,

$$r_{u,f} = \begin{cases} 1, & \text{if file } f \text{ is requested by user } u, \\ 0, & \text{otherwise,} \end{cases}$$

and

$$c_{m,f} = \begin{cases} 1, & \text{if file } f \text{ is cached by UAV } m, \\ 0, & \text{otherwise.} \end{cases}$$

[1]We note here we consider the constant noise power perspective where the overall noise is a composition of different types of noise at the receiver and the noise power is restricted by the receiver. Nevertheless, another perspective that the noise power is proportional to the bandwidth can also be accommodated in our proposed design approach via using the fact that the bandwidth optimization in this case is simply a concave optimization [34] with Sub-problem 3 in Sec. IVA.3.

We assume that user u can at most have R_u requests for a service period. Thus, we have $\sum_{f=1}^{F} r_{u,f} \leq R_u, \forall u$. In addition, since the cache capacity of a UAV is limited, we have the the caching capacity constraint given as:

$$\sum_{f=1}^{F} c_{m,f} \leq C, \forall m,$$

where C is the maximum number of files that can be cached by a UAV. We note that although we consider that caching capacity of UAVs is the same in this paper for simplicity, the extension to considering that different UAVs have different cache capacities is straightforward.

D. USER SCHEDULING AND ASSOCIATION MODEL

To reduce the interference between different communication links, we conduct the user scheduling and association. We denote $\alpha_{u,m}[n]$ as the binary scheduling and association indicator to indicate whether user u is served by UAV m at time-slot n. We consider that each UAV can only serve a single user at a time-slot and each user can only be served by a single UAV at a time-slot. Consequently, the following scheduling and association constraints are considered:

$$\sum_{u=1}^{U} \alpha_{u,m}[n] \leq 1, \forall m, n; \quad \sum_{m=1}^{M} \alpha_{u,m}[n] \leq 1, \forall u, n;$$

$$\alpha_{u,m}[n] \in \{0, 1\}, \forall u, m, n.$$

Subsequently, since user u can be served by UAV m only if the requested the file f by the user is cached by the UAV, we have the following constraint to satisfy:

$$\sum_{f=1}^{F} (r_{u,f} \cdot c_{m,f}) \geq \alpha_{u,m}[n], \forall u, m, n,$$

indicating that user u can be scheduled to be served by UAV m only if one of the requested files of user u is cached by UAV m.

Similarly, we denote $\beta_{u,m}[n]$ as the binary scheduling and association indicator to indicate whether user u is served by BS k at time-slot n. We again consider that a BS can only serve a single user at a time-slot and that each user can only be served by a single BS at a time-slot. It follows that the scheduling and association constraints for BSs are:

$$\sum_{u=1}^{U} \beta_{u,k}[n] \leq 1, \forall k, n; \quad \sum_{k=1}^{K} \beta_{u,k}[n] \leq 1, \forall u, n;$$

$$\beta_{u,k}[n] \in \{0, 1\}, \forall u, k, n.$$

We note that since BSs and UAVs use different bands to serve users and carrier aggregation is possible for the cellular network [35], there is no scheduling and association constraints between UAVs and BSs. In other words, a user can be served by a BS and a UAV simultaneously at a time-slot.

E. SYSTEM BANDWIDTH ALLOCATION MODEL AND USER RATE

We assume that the total bandwidth of our network is B_{total}. Since the BSs adopt the frequency reuse scheme, the frequency resources used by different BSs are orthogonal. We then denote $B_{\text{BS}_k}[n]$ as the bandwidth used by BS k in time-slot n. In addition, we denote $B_{\text{UAV}}[n]$ as the bandwidth used by the UAV in the time-slot n. Note that since UAVs share the same bandwidth, there is no need to split the bandwidth used by UAVs. It follows that when considering the bandwidth allocation between BSs and UAVs, the relevant constraint for the system is:

$$B_{\text{UAV}}[n] + \sum_{k=1}^{K} B_{\text{BS}_k}[n] \leq B_{\text{total}}, \forall n.$$

Using the aforementioned models, we can compute the achievable rate of user u in time-slot n by summing up the rates of UAVs and BSs provided to user u in the time-slot. Specifically, the achievable rate provided by UAV m to the user u at time-slot n is:

$$R_{u,m}^{UAV}[n] = \alpha_{u,m}[n] B_{\text{UAV}}[n] \log_2 \left(1 + \gamma_{u,m}[n]\right).$$

Similarly, the achievable rate provided by BS k to user u at time-slot n is:

$$R_{u,k}^{BS}[n] = \beta_{u,m}[n] B_{\text{BS}_k}[n] \log_2 \left(1 + \gamma_{u,k}[n]\right).$$

It follows that the total achievable rate provided to user u by the network at time-slot n is:

$$R_u^{sum}[n] = \sum_{m=1}^{M} R_{u,m}^{UAV}[n] + \sum_{k=1}^{K} R_{u,k}^{BS}[n].$$

Consequently, the sum-rate of user u provided by the network during the service period is:

$$R_u^{net} = \sum_{n=1}^{N} \left[\sum_{m=1}^{M} R_{u,m}^{UAV}[n] + \sum_{k=1}^{K} R_{u,k}^{BS}[n] \right].$$

III. PROPOSED CACHING, COMMUNICATION, AND TRAJECTORY OPTIMIZATION PROBLEM FORMULATIONS

In this section, we formulate two optimization problems that maximize the minimum achievable user rate of the network. The first problem, called CJCT problem, considers that the files cached in the UAVs are pre-determined and fixed, and then optimizes the user scheduling and association, UAV power allocation, bandwidth allocation for UAVs and BSs, and the flying trajectories of UAVs. This problem is more feasible for the cases where files cached in UAVs is infrequently updated due to overhead reduction. On the other hand, the second problem, called JCCT problem, considers that the cached files are jointly optimized with the communications and trajectories. Such consideration is more suitable for the cases where the cache update cost is not a concern. In the following, we first introduce the CJCT problem. Then, the JCCT problem is provided.

A. CACHE-AWARE JOINT COMMUNICATION AND TRAJECTORY OPTIMIZATION PROBLEM

In this subsection, the details of the CJCT problem are presented. We consider that the user requests $r_{u,f}, \forall u, f$ and the UAV caching placement $c_{m,f}, \forall m, f$ are given and known by UAVs and BSs. Then, the optimization problem that maximizes the minimum user rate in a service period is formulated as:

$$\max_{\substack{\mathrm{q}_m[n],\delta_{m,k},\\ \alpha_{u,m}[n],\beta_{u,k}[n],\\ p_m[n],B_{\mathrm{UAV}}[n],\\ B_{\mathrm{BS}_k}[n],\forall m,n,k}} \min_u \quad R_u^{net} \tag{1a}$$

$$\text{subject to} \quad 0 \le \|\mathrm{q}_m[n+1]-\mathrm{q}_m[n]\|^2 \le d_{\max}^2, \forall m, n, \tag{1b}$$

$$\sum_{k=1}^{K} \delta_{m,k}\mathrm{q}_{\mathrm{BS},k} = \mathrm{q}_m[N], \forall m, \sum_{k=1}^{K} \delta_{m,k} = 1, \forall m, \tag{1c}$$

$$\sum_{m=1}^{M} \delta_{m,k} \le M_{\max}, \forall k, \delta_{m,k} \in \{0,1\}, \forall m, k, \tag{1d}$$

$$0 \le p_m[n] \le P_{\max}, \forall m, n, \tag{1e}$$

$$\sum_{u=1}^{U} \alpha_{u,m}[n] \le 1, \forall m, n; \sum_{m=1}^{M} \alpha_{u,m}[n] \le 1, \forall u, n, \tag{1f}$$

$$\alpha_{u,m}[n] \in \{0,1\}, \forall u, m, n, \tag{1g}$$

$$\sum_{u=1}^{U} \beta_{u,k}[n] \le 1, \forall k, n; \sum_{k=1}^{K} \beta_{u,k}[n] \le 1, \forall u, n, \tag{1h}$$

$$\beta_{u,k}[n] \in \{0,1\}, \forall u, k, n, \tag{1i}$$

$$\sum_{f=1}^{F} \left(r_{u,f} \cdot c_{m,f}\right) \ge \alpha_{u,m}[n], \forall u, m, n, \tag{1j}$$

$$B_{\mathrm{UAV}}[n] + \sum_{k=1}^{K} B_{\mathrm{BS}_k}[n] \le B_{\mathrm{total}}, \forall n, \tag{1k}$$

where (1b) is the moving distance constraint; (1c) and (1d) are the landing position constraints; (1e) is the transmit power constraint; (1f)-(1j) are the scheduling and association constraints for users; and (1k) is the system bandwidth allocation constraint. We note that because the interference between BSs and UAVs are avoided by using frequency reuse scheme and orthogonal bands, the optimization of the BS transmit power is straightforward – the larger transmit power always gives better user rate. Therefore, the transmit power optimization for BSs is not included in this paper. By introducing a slack variable η, the problem in (1) can be equivalently written as:

$$\max_{\substack{\eta,\mathrm{q}_m[n],\delta_{m,k},\\ \alpha_{u,m}[n],\beta_{u,k}[n],\\ p_m[n],B_{\mathrm{UAV}}[n],\\ B_{\mathrm{BS}_k}[n],\forall m,n,k}} \quad \eta \tag{2a}$$

$$\text{subject to} \quad R_u^{net} \ge \eta, \forall u, \tag{2b}$$

$$(1b) - (1k). \tag{2c}$$

Such reformulation will be used frequently later in this paper.

B. JOINT CACHING, COMMUNICATION, AND TRAJECTORY OPTIMIZATION PROBLEM FORMULATION

In this subsection, given the user requests $r_{u,f}, \forall u, f$, we formulate the JCCT optimization problem that again maximizes the minimum user rate in a service period as:

$$\max_{\substack{c_{m,f},\mathrm{q}_m[n],\delta_{m,k},\\ \alpha_{u,m}[n],\beta_{u,k}[n],\\ p_m[n],B_{\mathrm{UAV}}[n],\\ B_{\mathrm{BS}_k}[n],\forall m,n,k,f}} \min_u R_u^{net} \tag{3a}$$

$$\text{subject to} \sum_{f=1}^{F} c_{m,f} \le C, \forall m, \tag{3b}$$

$$c_{m,f} \in \{0,1\}, \forall m, f, \tag{3c}$$

$$(1b) - (1k). \tag{3d}$$

Obviously, we observe that the difference between (1) and (3) lies in that constraints (3b) and (3c) are additional included for caching placement optimization, where the numbers of files that can be cached are upper bound by C. In other words, the difference between (1) and (3) lies only in that the files to be cached in UAVs need to be jointly optimized with the communications and UAV trajectories in (3). Similar to (2), we can introduce a slack variable η to reformulate (3) as:

$$\max_{\substack{\eta,c_{m,f},\mathrm{q}_m[n],\delta_{m,k},\\ \alpha_{u,m}[n],\beta_{u,k}[n],\\ p_m[n],B_{\mathrm{UAV}}[n],\\ B_{\mathrm{BS}_k}[n],\forall m,n,k}} \quad \eta \tag{4a}$$

$$\text{subject to} \quad R_u^{net} \ge \eta, \forall u, \tag{4b}$$

$$(3b), (3c), (1b) - (1k). \tag{4c}$$

We can see that (1) and (3) are non-convex mixed-integer problems which are difficult to be solved by standard solvers. Thus, in the following sections, we propose solution approaches to effectively solve them by exploiting the BCD method [30] and the SCA [31]. We remark that the harmony between the caching placement, user scheduling, association, UAV power allocation, bandwidth allocation, and UAV trajectories is the most critical part to the successful optimization. To achieve this, the key factors to consider are the SINR values of each user as well as the fairness issue between users. In other words, we need the SINR values for each scheduled link to be good. In addition, as our goal in this paper is to maximize the minimum user rate, we also need users to be scheduled with high fairness within the overall transmission period.

IV. PROPOSED CACHING, COMMUNICATION, AND TRAJECTORY OPTIMIZATION APPROACHES

In this section, we propose approaches that solve the formulated problems. To do this, our idea is to first use BCD method to decompose the problems into sub-problems. Then, different solution approaches are proposed to conquer different sub-problems, respectively. Finally, following the principle of BCD method, we then iteratively solve the sub-problems until convergence. Note that the rationale of the decomposition is that when the variables to optimize in different sub-problems are "separable" in constraints [30], iteratively optimizing variables in a sub-problem when fixing other variables can the monotonically improve of the objective function of the original problem if the sub-problem can be effective solved. This will be discussed in more detail along with the convergence analysis in the next section.

Following the aforementioned idea, when considering the CJCT problem in (1), we see that the problem can be decomposed into four sub-problems: user scheduling and association, UAV transmit power allocation, system bandwidth allocation, and UAV trajectory optimization sub-problems, where the user scheduling and association sub-problem is to optimize $\alpha_{u,m}[n], \forall u, m, n$ and $\beta_{u,k}[n], \forall u, k, n$; the power allocation sub-problem is to optimize $p_m[n], \forall m, n$; the system bandwidth allocation sub-problem is to optimize $B_{\text{UAV}}[n], \forall n$ and $B_{\text{BS}_k}[n], \forall k, n$; and the UAV trajectory optimization sub-problem is to optimize $q_m[n], \forall m, n$ and $\delta_{m,k}, \forall m, k$. Similarly, when considering the JCCT problem in (3), we see that the problem can be decomposed into four sub-problems: user scheduling, association, and cache placement, UAV transmit power allocation, system bandwidth allocation, and UAV trajectory sub-problems, where the latter three sub-problems are the same sub-problems as those mentioned above, while the user scheduling, association, and cache placement sub-problem is to jointly optimize $c_{m,f}, \forall m, f$, $\alpha_{u,m}[n], \forall u, m, n$, and $\beta_{u,k}[n], \forall u, k, n$.

In the following, the solution approaches for solving the corresponding sub-problems of the CJCT problem will first be discussed along with the overall iterative algorithm for solving the CJCT problem. Then, the solution approach for solving the JCCT problem will be provided.

A. PROPOSED SOLUTION APPROACH FOR CACHE-AWARE JOINT COMMUNICATION AND TRAJECTORY OPTIMIZATION PROBLEM

In this subsection, we first provide the details of the sub-problems obtained from decomposing the CJCT problem with their corresponding solution approaches. Then, the overall algorithm is presented at the end of this subsection.

1) SUB-PROBLEM 1: USER SCHEDULING AND ASSOCIATION SUB-PROBLEM

For sub-problem 1, we consider optimizing the user scheduling and association to UAVs and BSs while considering other variables are fixed. As a result, based on (1) and (2), the user scheduling and association sub-problem can be formulated as:

$$\begin{aligned} \max_{\substack{\eta,\alpha_{u,m}[n],\\ \beta_{u,k}[n],\forall u,m,k,n}} \quad & \eta \\ \text{subject to } & R_u^{net} \geq \eta, \forall u, \\ & (1f), (1g), (1h), (1i), (1j). \end{aligned} \tag{5}$$

We can then see that (5) is a linear integer programming (LIP). Thus, the standard LIP solver can be used to solve it. In addition, to further reduce the complexity, we can relax $\alpha_{u,m}[n], \forall u, m, n$ and $\beta_{u,k}[n], \forall u, k, n$ to be within 0 and 1, and then solve the relaxed problem which is as simple as a linear programming (LP):

$$\begin{aligned} \max_{\substack{\eta,\alpha_{u,m}[n],\\ \beta_{u,k}[n],\forall u,m,k,n}} \quad & \eta \\ & R_u^{net} \geq \eta, \forall u, \\ & (1f), (1h), (1j) \\ & 0 \leq \alpha_{u,m}[n] \leq 1, \forall u, m, n, \\ & 0 \leq \beta_{u,k}[n] \leq 1, \forall u, k, n. \end{aligned} \tag{6}$$

It then turns out that (6) can be solved with very low complexity. Finally, with the solution of (6), we can obtain the final integer solution of $\alpha_{u,m}[n], \forall u, m, n$ and $\beta_{u,k}[n], \forall u, k, n$ by using rounding. Note that due to the constraints $\sum_{u=1}^{U} \alpha_{u,m}[n] \leq 1, \forall m, n$, $\sum_{m=1}^{M} \alpha_{u,m}[n] \leq 1, \forall u, n$, $\sum_{u=1}^{U} \beta_{u,k}[n] \leq 1$, and $\sum_{k=1}^{K} \beta_{u,k}[n] \leq 1, \forall u, n$, such rounding can directly lead to a feasible solution of (5). By empirical experience, we see that solving (6) and conducting rounding can give the similar as that obtained by solving (5) using a stand LIP solver in almost all cases.

2) SUB-PROBLEM 2: UAV POWER ALLOCATION SUB-PROBLEM

For the UAV power allocation sub-problem, we consider optimizing the powers of UAVs while considering other variables are fixed. Thus, the UAV power allocation sub-problem is formulated as:

$$\begin{aligned} \max_{\eta,p_m[n],\forall m,n} \quad & \eta \\ \text{subject to } & \sum_{n=1}^{N} \left\{ \sum_{m=1}^{M} \alpha_{u,m}[n] B_{\text{UAV}}[n] \log_2\left(1+\gamma_{u,m}[n]\right) \right. \\ & \left. + \sum_{k=1}^{K} R_{u,k}^{BS}[n] \right\} \geq \eta, \forall u, \\ & 0 \leq p_m[n] \leq P_{\max}, \forall m, n, \end{aligned} \tag{7}$$

where

$$\gamma_{u,m}[n] = \frac{p_m[n] h_{u,m}[n]}{\sum_{j=1, j\neq M}^{M} p_j[n] h_{u,j}[n] + \sigma^2}$$

is a function of the power allocation. We see that (7) is non-convex because of $\gamma_{u,m}[n]$. Therefore, we use the technique of SCA which obtains a convexified problem by approximating $\hat{R}_{u,m}[n]$ with a convex function. Then, by iteratively solving the convexified problem, we can obtain a local optimum of (7) according to the theory of SCA [31]. Specifically, to obtain a convexified problem, we see that

$$\log_2\left(1+\gamma_{u,m}[n]\right) = \log_2\left(1+\frac{p_m[n]h_{u,m}[n]}{\sum_{j=1,j\neq m}^{M} p_j[n]h_{u,j}[n]+\sigma^2}\right)$$
$$= \log_2\left(\sum_{j=1}^{M} p_j[n]h_{u,j}[n]+\sigma^2\right) - \hat{R}_{u,m}[n] \quad (8)$$

where

$$\hat{R}_{u,m}[n] = \log_2\left(\sum_{j=1,j\neq m}^{M} p_j[n]h_{u,j}[n]+\sigma^2\right).$$

Then noticing that the linearization at any point is always a global upper of a concave function, we can obtain an upper bound at the given local point $p_j^i[n]$ as:

$$\hat{R}_{u,m}[n] \leq \sum_{j=1,j\neq m}^{M}\left(S_{u,j}[n]\left(p_j[n]-p_j^i[n]\right)\right) + \log_2\left(\sum_{j=1,j\neq m}^{M} p_j^i[n]h_{u,j}[n]+\sigma^2\right) \triangleq \hat{R}_{u,m}^{ub}[n], \quad (9)$$

where $p_j^i[n]$ is the given power allocation obtained from the previous round of optimization (or initialization) for conducting the linearization; and

$$S_{u,j}[n] = \frac{h_{u,j}[n]\log_2(e)}{\sum_{l=1,l\neq m}^{M} p_l^i[n]h_{u,l}[n]+\sigma^2}, \quad \forall u,j,n.$$

By using (8) and (9), the problem in (7) can be convexified as:

$$\max_{\eta,p_m[n],\forall m,n} \ \eta$$
$$\text{subject to} \ \sum_{n=1}^{N}\left\{\sum_{m=1}^{M}\left(\alpha_{u,m}[n]B_{\mathrm{UAV}}[n]R_{u,m}^{UAV,Polb}[n]\right) + \sum_{k=1}^{K} R_{u,m}^{BS}[n]\right\} \geq \eta, \forall u,$$
$$0 \leq p_m[n] \leq P_{\max}, \forall m,n, \quad (10)$$

where

$$R_{u,m}^{UAV,Polb}[n] = \log_2\left(\sum_{j=1}^{M} p_j[n]h_{u,j}[n]+\sigma^2\right) - \hat{R}_{u,m}^{ub}[n]$$

is a lower bound of $\log_2\left(1+\gamma_{u,m}[n]\right)$. Problem (10) is then a convex optimization problem. Thus, it can be solved by the standard convex solver, e.g, CVX. Furthermore, as will be shown in detail later in next section, iteratively solving (10) can guarantee the monotonic improvement of the objective function value in (7).

3) SUB-PROBLEM 3: SYSTEM BANDWIDTH ALLOCATION SUB-PROBLEM

For the system bandwidth allocation optimization problem, we consider optimizing the system bandwidth allocated to BSs and UAVs while fixing other variables. The optimization problem is then formulated as:

$$\max_{\substack{\eta,B_{\mathrm{UAV}}[n],\\ B_{\mathrm{BS}_k}[n],\forall k,n}} \ \eta$$
$$\text{subject to} \ \sum_{n=1}^{N}\left\{\sum_{m=1}^{M}\alpha_{u,m}[n]B_{\mathrm{UAV}}[n]\log_2\left(1+\gamma_{u,m}[n]\right) + \sum_{k=1}^{K}\beta_{u,k}[n]B_{\mathrm{BS}_k}[n]\log_2\left(1+\gamma_{u,k}[n]\right)\right\} \geq \eta, \forall u,$$
$$B_{\mathrm{UAV}}[n] + \sum_{k=1}^{K} B_{\mathrm{BS}_k}[n] \leq B_{\mathrm{total}}, \forall n. \quad (11)$$

We can observe that (11) is simply a standard LP, and thus its optimal solution can be efficiently obtained by using the standard LP solver.

4) SUB-PROBLEM 4: UAV TRAJECTORY OPTIMIZATION SUB-PROBLEM

For the UAV trajectory optimization sub-problem, we consider optimizing the trajectories of UAVs while fixing the others. The optimization sub-problem is formulated as:

$$\max_{\substack{\eta,\mathrm{q}_m[n],\\ \delta_{m,k},\forall m,k,n}} \ \eta \quad (12a)$$
$$\text{subject to} \ \sum_{n=1}^{N}\left[\sum_{m=1}^{M} R_{u,m}^{UAV}[n] + \sum_{k=1}^{K} R_{u,k}^{BS}[n]\right] \geq \eta, \forall u, \quad (12b)$$
$$0 \leq \|\mathrm{q}_m[n+1]-\mathrm{q}_m[n]\|^2 \leq d_{\max}^2, \forall m,n, \quad (12c)$$
$$\mathrm{q}_m[N] = \sum_{k=1}^{K}\delta_{m,k}\mathrm{q}_{\mathrm{BS},k}, \forall m, \ \sum_{k=1}^{K}\delta_{m,k} = 1, \forall m, \quad (12d)$$
$$\sum_{m=1}^{M}\delta_{m,k} \leq M_{\max}, \forall k, \ \delta_{m,k} \in \{0,1\}, \forall m,k, \quad (12e)$$

where $R_{u,m}^{UAV}[n]$ is a function of the trajectories according to the models in Sec. II. We note that (12) is a non-convex optimization problem due to (12b). Thus, to solve it, the SCA technique is again used. Specifically, we observe that:

$$\begin{aligned}
R_{u,m}^{UAV}[n] &= \log_2\left(1+\gamma_{u,m}[n]\right)\\
&= \log_2\left(1+\frac{\frac{p_m[n]\rho_0}{\|q_m[n]-q_{US,u}[n]\|^2+H^2}}{\sum_{j=1,j\neq m}^{M}\frac{p_j[n]\rho_0}{\|q_j[n]-q_{US,u}[n]\|^2+H^2}+\sigma^2}\right)\\
&= \widetilde{R}_{u,m}[n]\\
&\quad -\log_2\left(\sum_{j=1,j\neq m}^{M}\frac{p_j[n]\rho_0}{\|q_j[n]-q_{US,u}[n]\|^2+H^2}+\sigma^2\right),
\end{aligned}$$

where

$$\widetilde{R}_{u,m}[n] = \log_2\left(\sum_{j=1}^{M}\frac{p_j[n]\rho_0}{\|q_j[n]-q_{US,u}[n]\|^2+H^2}+\sigma^2\right).$$

Subsequently, we introduce slack variables $Q_{u,m}[n]$, $\forall u, m, n$, and then convert (12b) as:

$$\begin{aligned}
&\sum_{n=1}^{N}\left\{\sum_{m=1}^{M}\left(\alpha_{u,m}[n]B_{\text{UAV}}[n]\left(\widetilde{R}_{u,m}[n]-\log_2\left(\sum_{j=1,j\neq m}^{M}\right.\right.\right.\right.\\
&\quad\left.\left.\left.\left.\times\frac{p_j[n]\rho_0}{Q_{u,j}[n]+H^2}+\sigma^2\right)\right)\right)+\sum_{k=1}^{K}R_{u,k}^{BS}[n]\right\}\geq\eta,\ \forall u.
\end{aligned}\tag{13}$$

It follows that, problem (12) can be equivalently expressed as:

$$\begin{aligned}
\max_{\substack{\eta, q_m[n],\\ Q_{u,m}[n],\forall u,m,n}} \quad & \eta\\
\text{subject to}\quad & Q_{u,m}[n]\leq\|q_m[n]-q_{US,u}[n]\|^2,\ \forall u,m,n,\\
& (12c), (12d), (12e), (13).
\end{aligned}\tag{14}$$

It can be observed that the term $\log_2\left(\sum_{j=1,j\neq m}^{M}\frac{p_j[n]\rho_0}{Q_{u,j}[n]+H^2}+\sigma^2\right)$ in (13) is now convex with respect to $Q_{u,j}[n]$. However, the problem is still non-convex because $\widetilde{R}_{u,m}[n]$ in (13) is still non-convex with respect to $q_m[n]$ and $Q_{u,m}[n] \leq \|q_m[n]-q_{US,u}\|^2$ is a non-convex constraint.

To further resolve the non-convexity, we apply the SCA technique to them. Specifically, we let $q_m^i[n], \forall m, n$ be a given trajectories of UAVs for linearization. Then, we use the first-order Taylor expansion to approximate $\widetilde{R}_{u,m}[n]$ at the given trajectories $q_m^i[n], \forall m, n$ and obtain the following lower bound for it:

$$\begin{aligned}
\widetilde{R}_{u,m}[n] \geq \sum_{j=1}^{M}&\left\{\left(-C_{u,j}^i[n]\right)\left(\|q_j[n]-q_{US,u}[n]\|^2\right.\right.\\
&\left.\left.-\left\|q_j^i[n]-q_{US,u}[n]\right\|^2\right)\right\}+D_{u,m}^i[n]\triangleq\widetilde{R}_{u,m}^{\text{lb}}[n]
\end{aligned}\tag{15}$$

where $C_{u,j}^i[n]$ and $D_{u,m}^i[n]$ are some constants given as:

$$\begin{aligned}
C_{u,j}^i[n] &= \frac{\frac{p_j[n]\rho_0}{\left(\left\|q_j^i[n]-q_{US,u}[n]\right\|^2+H^2\right)^2}\log_2(e)}{\sum_{l=1}^{M}\frac{p_l[n]\rho_0}{\left\|q_l^i[n]-q_{US,u}[n]\right\|^2+H^2}+\sigma^2};\\
D_{u,m}^i[n] &= \log_2\left(\sum_{l=1}^{M}\frac{p_l[n]\rho_0}{\left\|q_l^i[n]-q_{US,u}[n]\right\|^2+H^2}+\sigma^2\right).
\end{aligned}\tag{16}$$

To deal with the constraint $Q_{u,m}[n] \leq \|q_m[n]-q_{US,u}[n]\|^2$, we use the first-order Taylor expansion to approximate it with respect to the given $q_m^i[n]$. We then obtain:

$$\begin{aligned}
&\|q_m[n]-q_{US,u}[n]\|^2\\
&\geq\left\|q_m^i[n]-q_{US,u}[n]\right\|^2\\
&\quad+2\left(q_m^i[n]-q_{US,u}[n]\right)^T\left(q_m[n]-q_m^i[n]\right).
\end{aligned}\tag{17}$$

After the above convexifications, problem (14) can then be convexified as:

$$\begin{aligned}
\max_{\substack{\eta, q_m[n],\\ Q_{u,m}[n],\forall u,m,n}} \quad & \eta\\
\text{subject to}\quad & \sum_{n=1}^{N}\left[\sum_{m=1}^{M}R_{u,m}^{UAV,lb}[n]+\sum_{k=1}^{K}R_{u,k}^{BS}[n]\right]\geq\eta,\ \forall u,\\
& Q_{u,m}[n]\leq\left\|q_m^i[n]-q_{US,u}[n]\right\|^2+2\cdot\\
& \left(q_m^i[n]-q_{US,u}[n]\right)^T\left(q_m[n]-q_m^i[n]\right),\\
& \forall u, m, n,\\
& (12c), (12d), (12e),
\end{aligned}\tag{18}$$

where

$$\begin{aligned}
R_{u,m}^{UAV,lb}[n] = \alpha_{u,m}[n]B_{\text{UAV}}[n]&\left(\widetilde{R}_{u,m}^{\text{lb}}[n]\right.\\
&\left.-\log_2\left(\sum_{j=1,j\neq m}^{M}\frac{p_j[n]\rho_0}{Q_{u,j}[n]+H^2}+\sigma^2\right)\right).
\end{aligned}$$

Now, problem (18) is a convex optimization problem if we ignore the integer constraints of $\delta_{m,k}$, $\forall m, k$ in (13). However, thanks to the development of the solver for the mixed-integer convex optimization, (18) with the linear integer constraints can be solved by the solver, e.g., CVX, efficiently if the number of UAVs is not very large. On the other hand, we can also relax $\delta_{m,k}$ to be within zero and one to obtain a pure convex problem after optimization, and then conduct the overall optimization until convergence. Then, if the stopping locations are not at the BSs, we move the stopping locations of UAVs to the nearest BSs. We note that it is also possible that we manually determine the landing locations of UAVs, e.g., many works in the literature let the UAVs land at the same locations as where they take off. In this

Algorithm 1 Proposed Algorithm for Solving Cache-Aware Joint Communication and Trajectory Optimization Problem

1: Given users requests $r_{u,f}, \forall u, f$ and UAVs caching placements $c_{m,f} \forall m, f$
2: Set $i = 0$. Initialize $p_m^i[n], \forall m, n$, $q_m^i[n], \forall m, n$, $B_{\text{UAV}}^i[n], \forall n$, and $B_{\text{BS}_k}^i[n], \forall k, n$.
3: **repeat**
4: Given $p_m^i[n], \forall m, n$, $q_m^i[n], \forall m, n$, $B_{\text{UAV}}^i[n], \forall n$, and $B_{\text{BS}_k}^i[n], \forall k, n$, solve sub-problem 1 in (5) using the solver or using (6) along with rounding. Then, update $\alpha_{u,m}^{i+1}[n], \forall u, m, n$ and $\beta_{u,k}^{i+1}[n], \forall u, k, n$.
5: Given $\alpha_{u,m}^{i+1}[n], \forall u, m, n$, $\beta_{u,k}^{i+1}[n], \forall u, k, n$, $q_m^i[n], \forall m, n$, $B_{\text{UAV}}^i[n], \forall n$, and $B_{\text{BS}_k}^i[n], \forall k, n$, solve sub-problem 2 in (7) via iteratively solving (10). Then, update $p_m^{i+1}[n], \forall m, n$.
6: Given $\alpha_{u,m}^{i+1}[n], \forall u, m, n$, $\beta_{u,k}^{i+1}[n], \forall u, k, n$, $p_m^i[n], \forall m, n$, $q_m^i[n], \forall m, n$, solve sub-problem 3 using (11). Then, update $B_{\text{UAV}}^i[n], \forall n$ and $B_{\text{BS}_k}^i[n], \forall k, n$.
7: Given $\alpha_{u,m}^{i+1}[n], \forall u, m, n$, $\beta_{u,k}^{i+1}[n], \forall u, k, n$, $p_m^{i+1}[n], \forall m, n$, $B_{\text{UAV}}^i[n], \forall n$, and $B_{\text{BS}_k}^i[n], \forall k, n$, solve sub-problem 4 via iteratively solving (18). Then, update $q_m^i[n], \forall m, n$.
8: Update $i = i + 1$.
9: **until** The improvement is below a predefined threshold ϵ.

case, (18) reduces simply to a convex optimization problem. By iteratively solving (18), the monotonic improvement of the objective function value for (12) can be guaranteed. This will be discussed in the next section.

5) OVERALL ALGORITHM

With the proposed solution approaches for each sub-problem, we now present the overall algorithm for solving the formulated CJCT problem. The idea of the overall algorithm is simply to iteratively solve the sub-problems. The overall algorithm is summarized in Alg. 1. We note that although the BCD method suggests that we conduct the inner loops that iteratively solve (10) and (18), respectively, for their own sub-problems until convergence first, and then iteratively solve the sub-problems in the outer loop, in practice, the convergence of the overall algorithm can be accelerated by only solving (10) and (18) a couple of iteration (even a single iteration) when conducting the inner loop. In addition, the order for solving the sub-problem might be changed if someone wants to fine-tuning the algorithm for solving a specific realization of the problem. However, by our empirical experiences, such fine-tuning commonly provides minor improvement, and thus might not be necessary.

B. PROPOSED SOLUTION APPROACH FOR JOINT CACHING, COMMUNICATION, AND TRAJECTORY OPTIMIZATION PROBLEM

In this subsection, given user requests $r_{u,f}, \forall u, f$, we discuss the approach for solving the JCCT problem. Specifically, by using the similar ideas and approaches in Sec. IV-A, we develop a solution approach consisting of solving four sub-problems: user scheduling, association, and cache placement, UAV power allocation, system bandwidth allocation, and UAVs trajectory optimization sub-problems. Then, since the latter three sub-problems are the same sub-problems as in Sec. IV-A when fixing the caching placement and user scheduling and association, the same approaches proposed in Sec. IV-A are directly used for solving them. Thus, in the following, we first discuss the solution approach for solving the user scheduling, association, and cache placement sub-problem. Then, the overall algorithm for solving the JCCT problem is provided.

1) SUB-PROBLEM 1: USER SCHEDULING, ASSOCIATION AND CACHE PLACEMENT SUB-PROBLEM

Based on (3) and (4), the user scheduling, association, and cache placement sub-problem can be formulated as:

$$\begin{aligned} \max_{\substack{\eta, \alpha_{u,m}[n], \beta_{u,k}[n], \\ c_{m,f}, \forall u,m,k,n,f}} \quad & \eta \\ \text{subject to } & R_u^{net} \geq \eta, \forall u, \\ & (1f), (1g), (1h), (1i), (1j), (3b), (3c). \end{aligned} \tag{19}$$

The difference between (19) and (5) lies only in that the cache placement constraints, i.e., $\sum_{f=1}^{F} c_{m,f} \leq C, \forall m$ and $c_{m,f} \in \{0, 1\}, \forall m, f$ are additional included and for the joint optimization with the user scheduling and association. From (19), we can see that this is a standard LIP, and thus the standard solver can be used for solving it. Interestingly, since the existing CVX toolbox support solving the mixed-integer linear problem, and thus we have come up with an alternative solution approach that can take advantage of such solver and obtain effective solutions. The alternative solution approach follows the similar approach for solving (5), in which $\alpha_{u,m}[n], \forall u, m, n$ and $\beta_{u,k}[n], \forall u, k, n$ are first relaxed to be within 0 and 1. Subsequently, the resulting mixed-integer linear problem is solved by using CVX. Finally, the obtained solution with fractional $\alpha_{u,m}[n], \forall u, m, n$ and $\beta_{u,k}[n], \forall u, k, n$ are rounded to 0 and 1 to attain the final solution. It turns out that such approach can provide almost the same performance as that of directly using LIP solver, but, surprisingly, be more efficient in terms of runtime for finding a solution in the Matlab platform in certain cases. The possible reason might be that treating $\alpha_{u,m}[n], \forall u, m, n$ and $\beta_{u,k}[n], \forall u, k, n$ as integers increases the worst-case complexity of the LIP solver, while this might not be the case for the algorithm adopted by the mixed-integer linear problem solver.

2) OVERALL ALGORITHM

With the proposed solution approaches for each sub-problem, we now present the overall algorithm for solving the formulated JCCT problem which is summarized in Alg. 2.

Algorithm 2 Proposed Algorithm for Solving Joint Caching, Communication, and Trajectory Optimization Problem

1: Given users request $r_{u,f}$
2: Set $i = 0$. Initialize $p^i_m[n], \forall m, n$, $q^i_m[n], \forall m, n$, $B^i_{\text{UAV}}[n], \forall n$, and $B^i_{\text{BS}_k}[n], \forall k, n$.
3: **repeat**
4: Given $p^i_m[n], \forall m, n$, $q^i_m[n], \forall m, n$, $B^i_{\text{UAV}}[n], \forall n$, and $B^i_{\text{BS}_k}[n], \forall k, n$, solve sub-problem 1 using (19) with the LIP solver (or using the proposed alternative approach). Then, update $c^i_{m,f}, \forall m, f$, $\alpha^{i+1}_{u,m}[n], \forall u, m, n$ and $\beta^{i+1}_{u,k}[n], \forall u, k, n$.
5: Solve the UAV power allocation, system bandwidth allocation, and UAVs trajectory optimization sub-problems following the same approaches provided in Alg. 1.
6: Update $i = i + 1$.
7: **until** The improvement is below a predefined threshold ϵ.

V. CONVERGENCE ANALYSIS OF THE PROPOSED SOLUTION APPROACHES

In this section, the convergences of the proposed solution approaches in Sec. IV are analyzed. We start our analysis with the proposed approach for solving the CJCT problem. Then, the analysis is directly extended to the proposed approach for solving the JCCT problem. Note that the fundamental idea of our analysis is to show that using our proposed approach to optimize each sub-problem can monotonically improve the objective function of the original problem with a better feasible solution.

By the descriptions in Sec. IV-A and Alg. 1, it is clear that our proposed approach is to iteratively solve sub-problems in (5), (7), (11), and (12). Therefore, we first provide the following lemma:

Lemma 1: When solving (2) by iteratively solving sub-problems in (5), (7), (11), and (12), such solution approach can monotonically improving the solution for (2) given that the obtained solutions for solving each sub-problem are also monotonically improving.

proof: The proof is straightforward. We first observe that when solving a sub-problem, the solution obtained from the previous sub-problem is also one of the feasible solutions of the current sub-problem. Thus, as long as we can obtain a solution better than the feasible solution provided by the previous sub-problem when solving the sub-problem at each iteration, we are improving the solution for (2) at each iteration. □

With Lemma 1, the remaining is to show that our proposed approaches for solving each sub-problem can always provide a better solution as compared to the solution of the previous sub-problem. To do this, we first provide proposition 3 for characterizing the solution approaches for sub-problems 1 and 3 in Alg. 1:

Proposition 2: Our solution approaches for solving sub-problem 1 in (5) and sub-problem 3 in (11) improve the solution.

proof: This is straightforward as their optimal solutions can be obtained by the standard solvers. The only thing to notice is that in Alg. 1, we are suggesting solving (6) with rounding instead of directly solving (5). However, by our empirical experience, the monotonicity is still maintained even if the relaxation-and-rounding is adopted for complexity reduction, though this indeed is not theoretically guaranteed. □

We then provide proposition 4 for characterizing the solution approach for sub-problem 2 in Alg. 1:

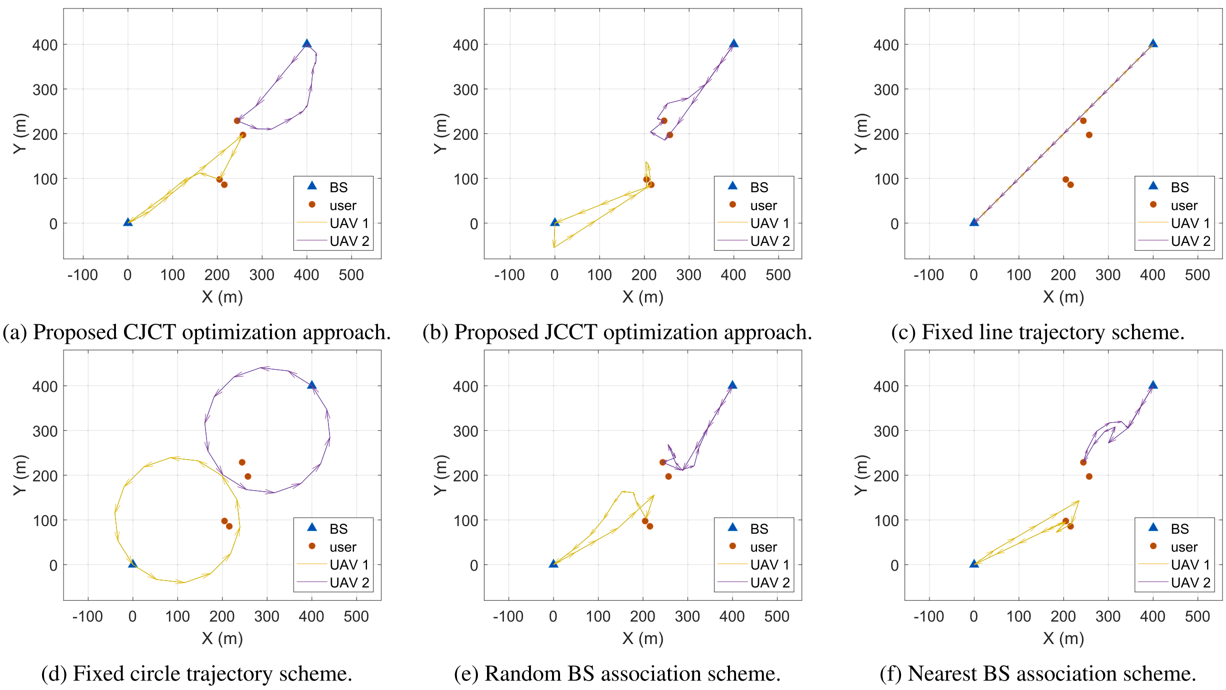
Proposition 3: Solving sub-problem 2 in (7) by iteratively solving (10) can improve the solution at each iteration.

proof: Recall that the difference between (7) and (10) is that the SCA is used to convexify the first constraint which gives the left-handed side function of the first constraint in (10) to always be a low bound of the left-handed side function of the first constraint in (7). Furthermore, if the lineraization is conducted on the solution of the previous iteration, the solution of the previous round is a feasible solution for both (7) and (10). We then denote $(\eta_{pre}, \mathbf{P}_{pre})$ as the solution of the previous iteration and $(\eta_{now}, \mathbf{P}_{now})$ as the solution after solving (10). We can then see that we have $\eta_{now} \geq \eta_{pre}$ as the optimal solution of (10) can be obtained by a convex solver. Subsequently, with a given solution $\mathbf{P}_{now}$, we denote $f^{Sub2}_{1,u}(\mathbf{P}_{now})$ and $f^{Sub2}_{2,u}(\mathbf{P}_{now})$ as the values of the left-handed side functions of the first constraints in (7) and (10), respectively. It follows that since $f^{Sub2}_{2,u}(\cdot)$ is a lower bound of $f^{Sub2}_{1,u}(\cdot)$, we have $f^{Sub2}_{1,u}(\mathbf{P}_{now}) \geq f^{Sub2}_{2,u}(\mathbf{P}_{now}) \geq \eta_{now} \geq \eta_{pre}, \forall u$. Therefore, the solution $\mathbf{P}_{now}$ obtained by solving (10) can provide an improved result for (7). □

Finally, we provide proposition 5 to characterize the solution approach for sub-problem 4 in Alg. 1:

Proposition 4: Solving sub-problem 4 in (12) by iteratively solving (18) can improve the solution at each iteration.

proof: The proof of this proposition is similar to the proof of Proposition 3. Recall that the difference between (12) and (18) is that the SCA is used to convexify the first constraint of (12), which gives rise to the first and second constraints of (18). Then, as we can again see that the convexification is conducted on the solution of the previous iteration, the solution of the previous round is a feasible solution for both (12) and (18). We then denote $(\eta_{pre}, \mathbf{q}_{pre})$ as the solution of the previous iteration and $(\eta_{now}, \mathbf{q}_{now}, \mathbf{Q}_{now})$ as the solution after solving (18). Subsequently, with a given solution $\mathbf{q}_{now}$, we denote $f^{Sub4}_{1,u}(\mathbf{q}_{now})$ and $f^{Sub4}_{2,u}(\mathbf{q}_{now}, \mathbf{Q}_{now})$ as the values of the left-handed side functions of the first constraints in (12) and (18), respectively, where $\mathbf{Q}_{now}$ is a feasible solution of (18) corresponding to the provided $\mathbf{q}_{now}$. We then see that $f^{Sub4}_{2,u}(\mathbf{q}_{now}, \mathbf{Q}_{now})$ is always a lower bound for $f^{Sub4}_{1,u}(\mathbf{q}_{now})$, as this can be seen from that



(a) Proposed CJCT optimization approach. (b) Proposed JCCT optimization approach. (c) Fixed line trajectory scheme. (d) Fixed circle trajectory scheme. (e) Random BS association scheme. (f) Nearest BS association scheme.

FIGURE 1. **UAV trajectories comparison of different schemes.**

$\widetilde{R}_{u,m}[n] \geq \widetilde{R}^{\text{lb}}_{u,m}[n]$ by (15) and that

$$\left\|\mathrm{q}_m[n] - \mathrm{q}_{\text{US},u}[n]\right\|^2 \geq \left\|\mathrm{q}^i_m[n] - \mathrm{q}_{\text{US},u}[n]\right\|^2 + 2\left(\mathrm{q}^i_m[n] - \mathrm{q}_{\text{US},u}[n]\right)^T\left(\mathrm{q}_m[n] - \mathrm{q}^i_m[n]\right) \geq Q_{u,m}[n]$$

by (17) with (18). It then follows that since the linearization is conducted at the previous solution, we have $\eta_{now} \geq \eta_{pre}$ as the optimal solution of (18) can be obtained by the standard solver. This leads to that $f^{Sub2}_{1,u}(\mathbf{q}_{now}) \geq f^{Sub2}_{2,u}(\mathbf{q}_{now}, \mathbf{Q}_{now}) \geq \eta_{now} \geq \eta_{pre}, \forall u$. Therefore, the solution $\mathbf{q}_{now}$ obtained by solving (18) can provide an improved result for (12). □

By Lemma 1 and above propositions, we provide the concluding theorems below:

Theorem 5: The proposed solution approach in Alg. 1 can monotonically improve the solution of the proposed CJCT problem in (1) until convergence.

proof: Recall that (1) is equivalent to solving (2). Thus, by using Lemma 1 and above propositions, we see that the proposed Alg. 1 can monotonically improve the performance. Then, since there must exist an upper bound for the objective function, the algorithm must converge. □

Theorem 6: The proposed solution approach in Alg. 2 can monotonically improve the solution of the proposed JCCT problem in (3) until convergence.

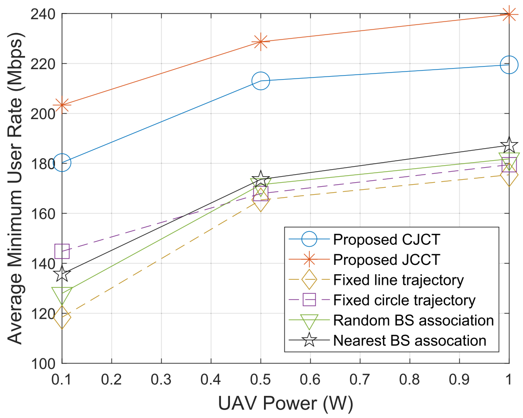
proof: This is straightforward by following the above analysis for obtaining Theorem 5 and using the fact that solving (19) improves the solution. □

VI. COMPUTER SIMULATIONS

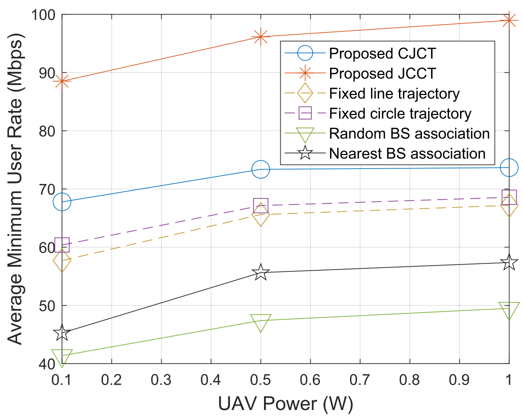
A. SIMULATION SETUP

In this section, the proposed approaches are evaluated in terms of the minimum user rate by using computer simulations. In the simulations, we consider that the BSs, UAVs and users are located within a square area of 420×420 m^2. In addition, we consider that the locations of BSs are fixed, where one BS is with coordinates $[0, 0]^T$ and the other is with coordinates $[400, 400]^T$. We consider that the initial locations of the UAVs are uniformly and randomly assigned to the locations of two BSs. We consider that UAVs fly at a fixed altitude of $H = 30$ m. In addition, for the parameters in the channel models, the noise power and the reference channel gain are given by $\sigma^2 = -110$ dBm, $\rho_0 = -60$ dB, respectively. Also, λ_c is the wavelength corresponding to the carrier frequency of 5 GHz. We consider that the maximum moving distance of a UAV for each time-slot is $d_{\max} = 75$ m and that the total bandwidth of the network is $B_{\text{total}} = 100$ MHz. We consider that the number of files is $F = 10$ and assume that each user can request three different files, where requests are generated uniformly at random. When the locations of users and the BSs are fixed, the shadowing effect for a user-BS link is invariant. Then, the exact values of the shadowing coefficient Ξ for each BS-user link are not critical, and thus we consider $\Xi = 0$ dB for simplicity.

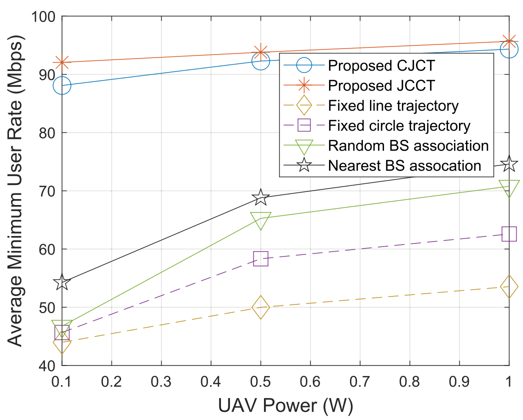
We consider both uniform and non-uniform user distributions. When with the uniform user distribution, user are uniformly distributed within the considered area. On the other hand, when with the non-uniform user distribution, users are only uniformly distributed within two specific areas $\mathcal{A}_1$ and $\mathcal{A}_2$, where the ranges for the x-axis and y-axis of $\mathcal{A}_1$ are $[-20, 20]$ and $[380, 420]$, respectively, and the ranges for the x-axis and y-axis of $\mathcal{A}_2$ are $[380, 420]$ and $[20, -20]$. In the simulations, unless otherwise indicated, three setups are considered for evaluation, where the first setup considers the number of UAVs, the number of users, and the number of



(a) First setup with $C = 3$, $M = 2$, $U = 4$, and $N = 15$.

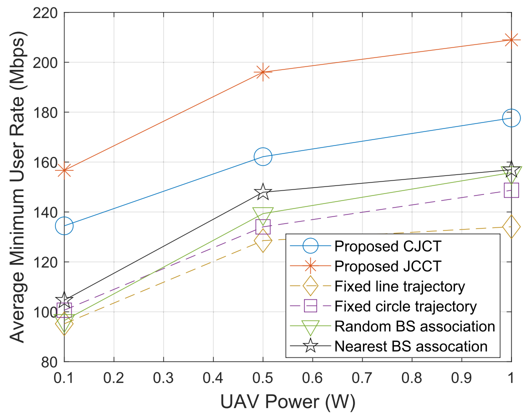


(b) Second setup with $C = 3$, $M = 2$, $U = 8$, and $N = 15$.

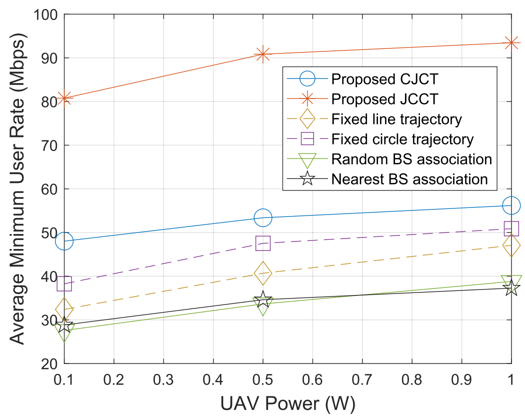


(c) Third setup with $C = 3$, $M = 4$, $U = 8$, and $N = 20$.

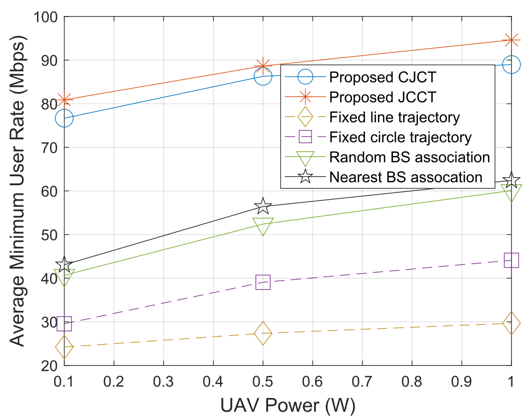
FIGURE 2. Performance evaluation with uniform user distribution as a function of the maximum transmit power.



(a) First setup with $C = 3$, $M = 2$, $U = 4$, and $N = 15$.



(b) Second setup with $C = 3$, $M = 2$, $U = 8$, and $N = 15$.



(c) Third setup with $C = 3$, $M = 4$, $U = 8$, and $N = 20$.

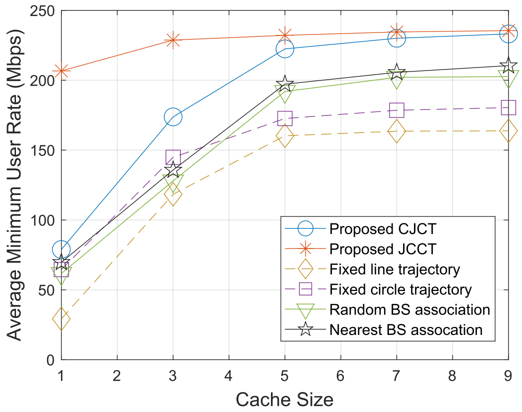
FIGURE 3. Performance evaluation with non-uniform user distribution as a function of the maximum transmit power.

time-slots to be $M = 2$, $U = 4$, and $N = 15$, respectively; the second setup considers $M = 2$, $U = 8$, and $N = 15$; and the third setup considers $M = 4$, $U = 8$, and $N = 20$.

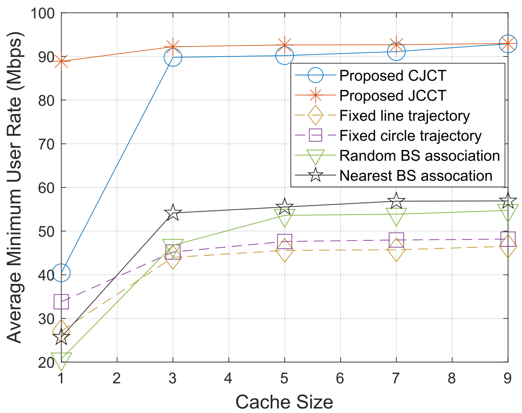
To evaluate our proposed CJCT and JCCT optimization approaches, we compare them with the following reference schemes:

(i) Fixed line trajectory scheme: In this scheme, the cached files of UAVs are given to be the same as the CJCT approach. and UAVs have fixed straight line trajectories that connect between two BSs. Then, the remaining arguments for optimization, i.e., the user association, power allocation, and bandwidth allocation, are optimized by iteratively using our proposed solution approaches for solving the sub-problems.

(ii) Fixed circle trajectory scheme: In this scheme, the cached files of UAVs are given and circle trajectories

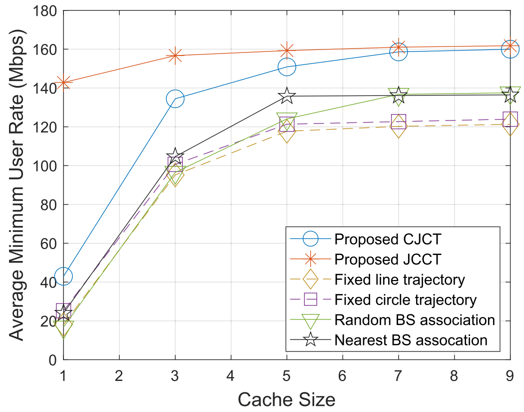


(a) First setup with $P_{\max} = 0.1$ W, $M = 2$, $U = 4$, and $N = 15$.

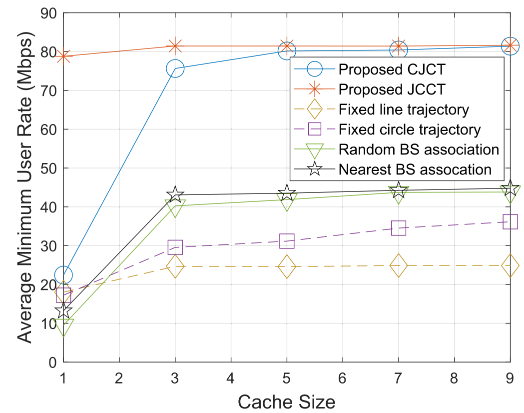


(b) Third setup with $P_{\max} = 0.1$ W, $M = 4$, $U = 8$, and $N = 20$.

FIGURE 4. Performance evaluation with uniform user distribution as a function of the cache capacity.



(a) First setup with $P_{\max} = 0.1$ W, $M = 2$, $U = 4$, and $N = 15$.



(b) Third setup with $P_{\max} = 0.1$ W, $M = 4$, $U = 8$, and $N = 20$.

FIGURE 5. Performance evaluation with non-uniform user distribution as a function of the cache capacity.

of UAVs are considered. The remaining arguments are again optimized by our proposed approaches.

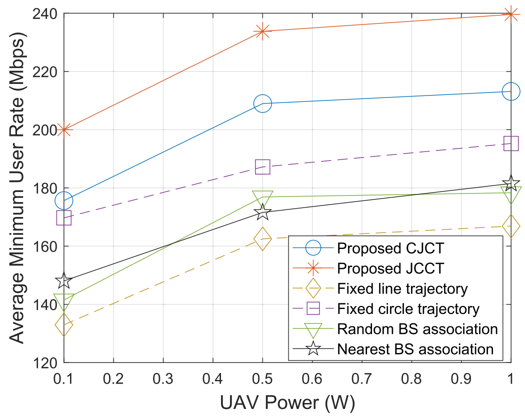
(iii) Random BS association scheme: In this scheme, users are first randomly associated to the BSs and served with round-robin scheduling. Then, given the user-BS associations and considering the UAVs are exclusively belong to the BSs that have the same locations as their initial locations, we can treat two BSs and their corresponding UAVs and users as independent groups, and then optimize the user-UAV scheduling and association, transmit powers, cache, and trajectories by using the approach proposed in [28]; the remaining system bandwidth allocation is optimized by our proposed approach.

(iv) Nearest BS association scheme: In this scheme, users are first associated with their nearest BSs. Then, the remaining follows the same approach as that of the random BS association scheme.

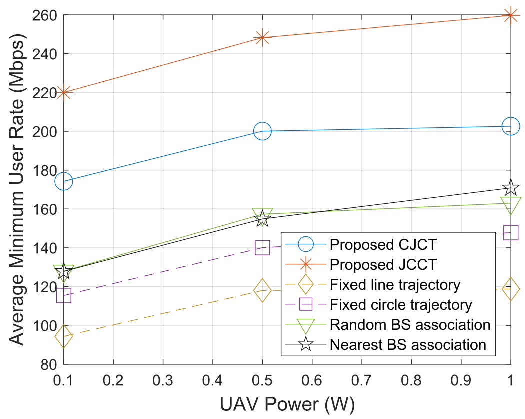
Fig. 1 shows the trajectories of two UAVs for different schemes with an illustrative realization. We can see that as compared to the reference schemes, our proposed approaches can provide (intuitively) more reasonable trajectories as they fly directly to the users they need to serve. We note that although the random and nearest BS association schemes also let the UAVs fly to the users, they indeed are less cooperative as compared to our proposed approaches, leading to worse performance as will be shown below.

B. PERFORMANCE EVALUATION RESULTS

In this subsection, the performance of our proposed solution approaches is evaluated and compared with the reference schemes. In Fig. 2, considering $C = 3$, the performance of our proposed approaches are evaluated with uniform user distribution as a function of the maximum transmit power. We observe that our proposed approaches can outperform the reference schemes. In addition, we see that as the caching is further optimized by the proposed JCCT approach, the JCCT approach is better than the proposed CJCT approach. We see that for all the approaches, the performance improves when the maximum transmit power increases as expected. In addition, we see that the approaches with fixed trajectories in general have poor performance because their trajectories

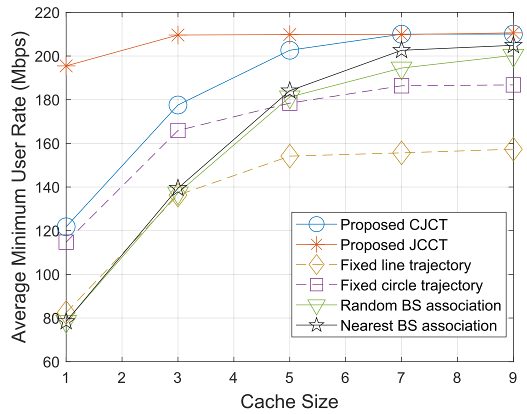


(a) Uniform user distribution.

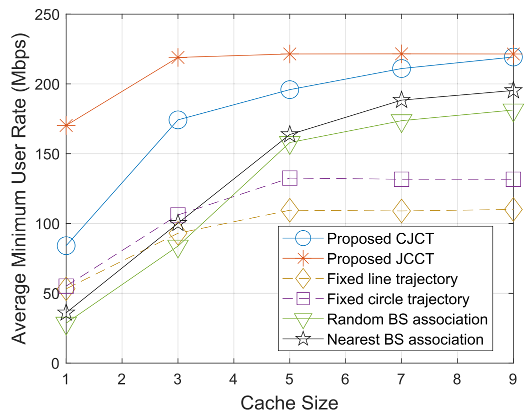


(b) Non-uniform user distribution.

FIGURE 6. Performance evaluation with a different BS placement under the first setup with $C = 3$ as a function of the maximum transmit power.



(a) Uniform user distribution.



(b) Non-uniform user distribution.

FIGURE 7. Performance evaluation with a different BS placement under the first setup with $P_{max} = 0.1$ W as a function of the cache capacity.

are not optimized. Furthermore, by comparing between Figs. 2(a) and 2(b), we see that when the number of users increase, the loading balance and interference issues become critical, and thus appropriate scheduling and association become highly influential. Fig. 3 evaluates the performance as a function of the transmit power with non-uniform distribution and $C = 3$. We again can see that the proposed approaches can outperform the reference schemes and that the approaches with fixed trajectories in general have poor performance. Finally, by comparing between the second and third setups in Figs. 2 and 3, it can be observed that the improvement of the proposed approaches as compared to the reference schemes seem to be more significant when the third setup is adopted, i.e., when the adopted setup is more complicated. This implies that our proposed approach can more effectively optimize the performance than the reference schemes in more complicated situations.

Considering the first and third setups, Fig. 4 shows the performance with uniform user distribution and $P_{\max} = 0.1$ W, as a function of the cache capacity of UAVs. We again see that our proposed approaches can significantly outperform the reference schemes. In addition, we see that when the cache capacity increases, the performance of all approaches improves as expected. Furthermore, it can be observed that the performance gap between the proposed JCCT and CJCT approaches diminishes as the cache capacity increases. This is because when the cache capacity is larger, the benefits of optimizing what to be cached in UAVs become less significant. Finally, we see that the approaches with fixed trajectories again have the worst performance. Fig. 5 shows the performance with non-uniform user distribution and $P_{\max} = 0.1$ W as a function of the cache capacity of UAVs. We observe that the proposed approaches significantly outperform the reference schemes as expected. In addition, all other observations found in Fig. 4 can again be found in Fig. 5. Furthermore, from both Figs. 4 and 5, it can be seen that although all approaches saturate when the cache capacity increases, the saturated performance levels of our proposed approaches are much larger than the reference schemes. This implies that even if determining what to be cached becomes less critical, our proposed approaches can provide better performance by optimizing other factors, e.g., trajectories, scheduling, and powers of UAVs.

Finally, in Figs. 6 and 7, we consider the first setup with a different BS placement for evaluation, where the locations of the BSs are now changed to be with coordinates $[120, 200]^T$

and $[280, 200]^T$, respectively. We observe that the proposed approaches still work effectively and the aforementioned observations are still valid under a different placement of BS locations, showing the ability of our approaches to be adopted in different scenarios. We note that simulation results with other setups show the behaviors similar to the above figures. Thus, they are omitted for brevity.

VII. CONCLUSION

In this paper, we considered the cache-aided integrated air-ground networks where both UAVs and BSs can serve the users and UAVs can fly across the BSs. To optimize the network performance, two maximum minimum user rate optimization problems were formulated. Then, approaches that can jointly optimize the UAV caching placements, UAV trajectories, user scheduling and association, UAV transmit powers, and system bandwidth allocation were proposed by first decomposing the formulated problems into sub-problems, and then solving the sub-problems iteratively until convergence. Solution approaches corresponding to each sub-problem were provided. In addition, convergences of the proposed approaches were analyzed. Computer simulations were conducted to evaluate our proposed approaches. Results showed that our proposed approaches work well and outperform the reference schemes.

REFERENCES

[1] F. Khan, "Multi-comm-core architecture for terabit-per-second wireless," *IEEE Commun. Mag.*, vol. 54, no. 4, pp. 124–129, Apr. 2016.

[2] H. Tataria, M. Shafi, A. F. Molisch, M. Dohler, H. Sjöland, and F. Tufvesson, "6G wireless systems: Vision, requirements, challenges, insights, and opportunities," *Proc. IEEE*, vol. 109, no. 7, pp. 1166–1199, Jul. 2021.

[3] Y. Zeng, Q. Wu, and R. Zhang, "Accessing from the sky: A tutorial on UAV communications for 5G and beyond," *Proc. IEEE*, vol. 107, no. 12, pp. 2327–2375, Dec. 2019.

[4] R. Ding, F. Gao, and X. S. Shen, "3D UAV trajectory design and frequency band allocation for energy-efficient and fair communication: A deep reinforcement learning approach," *IEEE Trans. Wireless Commun.*, vol. 19, no. 12, pp. 7796–7809, Dec. 2020.

[5] S. Jeong, O. Simeone, and J. Kang, "Mobile edge computing via a UAV-mounted cloudlet: Optimization of bit allocation and path planning," *IEEE Trans. Veh. Technol.*, vol. 67, no. 3, pp. 2049–2063, Mar. 2018.

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

[7] N. Tafintsev, D. Moltchanov, M. Gerasimenko, M. Gapeyenko, J. Zhu, S.-P. Yeh, N. Himayat, S. Andreev, Y. Koucheryavy, and M. Valkama, "Aerial access and backhaul in mmWave B5G systems: Performance dynamics and optimization," *IEEE Commun. Mag.*, vol. 58, no. 2, pp. 93–99, Feb. 2020.

[8] X. Wang, M. Chen, T. Taleb, A. Ksentini, and V. C. M. Leung, "Cache in the air: Exploiting content caching and delivery techniques for 5G systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 131–139, Feb. 2014.

[9] M. Ji, G. Caire, and A. F. Molisch, "Wireless device-to-device caching networks: Basic principles and system performance," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 1, pp. 176–189, Jan. 2016.

[10] L. Li, G. Zhao, and R. S. Blum, "A survey of caching techniques in cellular networks: Research issues and challenges in content placement and delivery strategies," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 3, pp. 1710–1732, 3rd Quart., 2018.

[11] D. Liu, B. Chen, C. Yang, and A. F. Molisch, "Caching at the wireless edge: Design aspects, challenges, and future directions," *IEEE Commun. Mag.*, vol. 54, no. 9, pp. 22–28, Sep. 2016.

[12] T.-Y. Kuo, M.-C. Lee, J.-H. Kim, and T.-S. Lee, "Quality-aware joint caching, computing and communication optimization for video delivery in vehicular networks," *IEEE Trans. Veh. Technol.*, vol. 72, no. 4, pp. 5240–5256, Apr. 2023.

[13] M.-C. Lee and A. F. Molisch, "Optimal delay-outage analysis for noise-limited wireless networks with caching, computing, and communications," *IEEE Trans. Wireless Commun.*, vol. 22, no. 2, pp. 1417–1431, Feb. 2023.

[14] T. Q. Duong, K. J. Kim, Z. Kaleem, M.-P. Bui, and N.-S. Vo, "UAV caching in 6G networks: A survey on models, techniques, and applications," *Phys. Commun.*, vol. 51, Apr. 2022, Art. no. 101532.

[15] M. Chen, M. Mozaffari, W. Saad, C. Yin, M. Debbah, and C. S. Hong, "Caching in the sky: Proactive deployment of cache-enabled unmanned aerial vehicles for optimized quality-of-experience," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 5, pp. 1046–1061, May 2017.

[16] T. Zhang, Y. Wang, Y. Liu, W. Xu, and A. Nallanathan, "Cache-enabling UAV communications: Network deployment and resource allocation," *IEEE Trans. Wireless Commun.*, vol. 19, no. 11, pp. 7470–7483, Nov. 2020.

[17] X. Xu, Y. Zeng, Y. L. Guan, and R. Zhang, "Overcoming endurance issue: UAV-enabled communications with proactive caching," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 6, pp. 1231–1244, Jun. 2018.

[18] B. Jiang, J. Yang, H. Xu, H. Song, and G. Zheng, "Multimedia data throughput maximization in Internet-of-Things system based on optimization of cache-enabled UAV," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 3525–3532, Apr. 2019.

[19] N. Zhao, F. Cheng, F. R. Yu, J. Tang, Y. Chen, G. Gui, and H. Sari, "Caching UAV assisted secure transmission in hyper-dense networks based on interference alignment," *IEEE Trans. Commun.*, vol. 66, no. 5, pp. 2281–2294, May 2018.

[20] M. Chen, W. Saad, and C. Yin, "Liquid state machine learning for resource and cache management in LTE-U unmanned aerial vehicle (UAV) networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 3, pp. 1504–1517, Mar. 2019.

[21] F. Cheng, G. Gui, N. Zhao, Y. Chen, J. Tang, and H. Sari, "UAV-relaying-assisted secure transmission with caching," *IEEE Trans. Commun.*, vol. 67, no. 5, pp. 3140–3153, May 2019.

[22] H. Wu, F. Lyu, C. Zhou, J. Chen, L. Wang, and X. Shen, "Optimal UAV caching and trajectory in aerial-assisted vehicular networks: A learning-based approach," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 12, pp. 2783–2797, Dec. 2020.

[23] T. Zhang, Z. Wang, Y. Liu, W. Xu, and A. Nallanathan, "Caching placement and resource allocation for cache-enabling UAV NOMA networks," *IEEE Trans. Veh. Technol.*, vol. 69, no. 11, pp. 12897–12911, Nov. 2020.

[24] T. Zhang, Z. Wang, Y. Liu, W. Xu, and A. Nallanathan, "Joint resource, deployment, and caching optimization for AR applications in dynamic UAV NOMA networks," *IEEE Trans. Wireless Commun.*, vol. 21, no. 5, pp. 3409–3422, May 2022.

[25] Y. Chen, K. Liao, M. Ku, F. Tso, and G. Chen, "Mobility-aware probabilistic caching in UAV-assisted wireless D2D networks," in *Proc. IEEE Global Commun.*, Dec. 2019, pp. 1–6.

[26] J. Ji, K. Zhu, D. Niyato, and R. Wang, "Probabilistic cache placement in UAV-assisted networks with D2D connections: Performance analysis and trajectory optimization," *IEEE Trans. Commun.*, vol. 68, no. 10, pp. 6331–6345, Oct. 2020.

[27] T. Zhang, Y. Wang, W. Yi, Y. Liu, and A. Nallanathan, "Joint optimization of caching placement and trajectory for UAV-D2D networks," *IEEE Trans. Commun.*, vol. 70, no. 8, pp. 5514–5527, Aug. 2022.

[28] J. Ji, K. Zhu, D. Niyato, and R. Wang, "Joint cache placement, flight trajectory, and transmission power optimization for multi-UAV assisted wireless networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 8, pp. 5389–5403, Aug. 2020.

[29] J. Ji, K. Zhu, and L. Cai, "Trajectory and communication design for cache-enabled UAVs in cellular networks: A deep reinforcement learning approach," *IEEE Trans. Mobile Comput.*, vol. 22, pp. 6190–6204, Oct. 2023.

[30] P. Tseng, "Convergence of a block coordinate descent method for nondifferentiable minimization," *J. Optim. Theory Appl.*, vol. 109, no. 3, pp. 475–494, Jun. 2001.

[31] M. Razaviyayn, "Successive convex approximation: Analysis and applications," Ph.D. dissertation, Dept. Elect. Comput. Eng., Univ. Minnesota, Minneapolis, MN, USA, May 2014.

[32] Y. Zeng, R. Zhang, and T. J. Lim, "Wireless communications with unmanned aerial vehicles: Opportunities and challenges," *IEEE Commun. Mag.*, vol. 54, no. 5, pp. 36–42, May 2016.
[33] A. F. Molisch, *Wireless Communications*, vol. 34. Hoboken, NJ, USA: Wiley, 2012.
[34] W. Yu and J. M. Cioffi, "FDMA capacity of Gaussian multiple-access channels with ISI," *IEEE Trans. Commun.*, vol. 50, no. 1, pp. 102–111, Jan. 2002.
[35] A. E. Garcia, M. Ozger, A. Baltaci, S. Hofmann, D. Gera, M. Nilson, C. Cavdar, and D. Schupke, "Direct air to ground communications for flying vehicles: Measurement and scaling study for 5G," in *Proc. IEEE 2nd 5G World Forum (5GWF)*, Sep. 2019, pp. 310–315.

YI-MO LIN received the M.S. degree from the Institute of Communications Engineering, National Yang Ming Chiao Tung University, Hsinchu, Taiwan, in 2023.

SHIN-PING HUANG received the B.S. degree in communications engineering from National Central University, Taiwan, in 2022. He is currently pursuing the M.S. degree with the Institute of Communications Engineering, National Yang Ming Chiao Tung University, Hsinchu, Taiwan.

MING-CHUN LEE (Member, IEEE) received the B.S. and M.S. degrees in electrical and computer engineering from National Chiao Tung University, Hsinchu, Taiwan, in 2012 and 2014, respectively, and the Ph.D. degree from the Ming Hsieh Department of Electrical Engineering, University of Southern California, in 2020. From 2014 to 2016, he was a Research Assistant with the Wireless Communications Laboratory, Research Center for Information Technology Innovation, Academia Sinica, Taiwan. He is currently an Associate Professor with the Institute of Communications Engineering, National Yang Ming Chiao Tung University. He received USC Annenberg Fellowship, from 2016 to 2020. He was awarded an Exemplary Reviewer of IEEE TRANSACTIONS ON COMMUNICATIONS, in 2019. His research interests include signal processing, design, modeling, and analysis in wireless systems and networks. He is especially working on topics relevant to caching, computing, and communication in wireless networks and integrated sensing and communication systems in recent years.

YUE-RONG HUANG (Student Member, IEEE) received the M.S. degree from the Institute of Communications Engineering, National Yang Ming Chiao Tung University, Hsinchu, Taiwan, in 2023. His research interests include the joint optimization of caching, computing, and communication for wireless networks.