Dynamic Transmission Power Allocation for Cache-enabled Multi-hop Networks

Emre Gures, Pavel Mach, and Zdenek Becvar
Faculty of Electrical Engineering Czech Technical University in Prague, Prague, Czech Republic
{guresemr, machp2, zdenek.becvar}@fel.cvut.cz

***Abstract*—In this paper, we focus on a cache-enabled multihop network, where the unmanned aerial vehicles (UAVs), the user equipment (UEs), or both can serve as relays to deliver contents to individual users from a ground base station (GBS). We formulate a power optimization problem with the objective to minimize the sum content delivery delay. We show the optimization problem is non-convex, thus, we propose a novel heuristic algorithm to allocate the power to individual contents at individual hops in multi-hop scenario. The proposed heuristic algorithm iteratively re-allocates the transmission power among the contents at the same transmitting node. To reduce the number of iteration, the proposed algorithm enables parallel power re-allocation for multiple content pairs and dynamic adaptation of power re-allocation steps, thereby enabling faster convergence of the algorithm. We demonstrate that our proposal reduces the average content delivery duration by up to 31.8% compared to state-of-the-art works. At the same time, proposal is suitable for real systems due to a very fast convergence.**

***Index Terms*—caching, content delivery duration, D2D relaying, UAV, power allocation**

I. Introduction

In recent years, content caching has emerged as one of the promising solutions to alleviate the backhaul load and enable low latency communications in mobile networks [1]. By caching frequently requested contents in proximity to end-users, such as at a ground base station (GBS), data transmissions across the core network can be avoided. The benefits of content caching can be amplified by using neighboring user equipments (UEs) as relays delivering the content by means of the device-to-device (D2D) relaying [2]. Similarly, integration of unmanned aerial vehicles (UAVs) acting as relays can improve the quality of experience for the UEs due to improved channel quality on individual communication hops.

The networks integrating caching and, at the same time, multi-hop communication via the UAV relays and/or via the relaying UEs (RUEs) face many challenges [3]. Thus, the authors in [4] optimize content placement in order to minimize content delivery duration. Furthermore, the authors in [5] jointly optimize the UAV deployment and content placement also with the objective to minimize a content delivery duration. However, neither [4] nor [5] optimize the transmission power allocation and transmission nodes (GBS or any relay) simply send the contents sequentially with the maximum transmission power. In addition, the problem of route selection, critical for multi-hop communication, is also not considered.

The problems of power allocation and route selection in cache-enabled networks with D2D and UAVs are tackled in several works targeting optimization of minimum secrecy rate among requesting UEs [6], maximization of the sum throughput [7], [8], maximization of content hit ratio [9], or maximization of energy efficiency [10], [11]. None of these works, however, targets optimization of the content delivery duration, which is one of the key parameters for caching from the UEs perspective. Besides, the aforementioned works are focused on direct communication [6], [7], [9], [10] or two-hop communication using just a single UAV relay [8], [11]. Hence, these works become of limited efficiency if the distance between the source and target nodes increases or the quality of communication channels deteriorates.

The minimization of the sum content delivery duration in multi-hop cache-enabled networks via the joint route selection and power allocation is addressed in [12]. The authors assume multiple contents are transmitted simultaneously by any transmitting node, since the simultaneous transmission reduces the content delivery duration with respect to sequential transmission (considered, e.g., in [4], [5]) of contents [13]. The reason is that the transmission duration is not linearly proportional to the allocated transmission power. Nevertheless, in [12], the entire transmission power budget of any transmitting node is *equally distributed* across the individual transmitted contents. While even such approach outperforms sequential transmission of the contents, the content delivery duration can still be further reduced for practical scenarios, where: i) each content has usually different sizes, hence, equal power allocation does not minimize the sum or average content delivery delay, and ii) the individual contents may be transmitted over different channels with distinct conditions, which can result in some contents being allocated insufficient power while others are allocated more power than necessary.

In this regard, the main goal of this paper is to optimize the transmission power allocated to multiple contents transmitted simultaneously by any transmitting node. The major contributions of this paper are summarized as follows:

- We formulate power optimization problem for cache-enabled multi-hop networks to minimize the overall content delivery duration and we demonstrate that the power allocation problem is non-convex.

This work was supported by Ministry of Education, Youth and Sport of the Czech Republic under Grant No. LTT20004, and by Czech Technical University in Prague under Grant no. SGS23/171/OHK3/3T/13.

- We propose a heuristic algorithm for power allocation to minimize the sum content delivery duration. To enable fast convergence of the proposed power allocation, the proposed algorithm leverages on: *i)* parallelization to simultaneously re-allocate the transmission power of contents in multiple pairs from the same node, and *ii)* dynamic setting of power adaptation step.
- We demonstrate that the proposal achieves up to a 31.8% reduction in average content delivery duration compared to the best performing state-of-the-art work. At the same time, we show the proposed power allocation approach is suitable for real networks as the number of iterations is very low (less than 100 iterations for 30 requesting UEs) and the number of iterations is increasing linearly with the number of requesting UEs.

The remainder of this paper is organized as follows. Section II presents the system model. In Section III, we formulate and discuss the problem addressed in this paper. The proposed power allocation algorithm is explained in Section IV. Section V and VI describe simulation description and results, respectively. Finally, Section VII concludes the paper.

II. System Model

This section describes the network model, the cache placement model, and model of the content delivery duration.

A. Network Model

We consider a multi-hop network with the UAVs and D2D communication, as illustrated in Fig. 1. We assume one GBS and U UAVs acting as flying relays. Further, we assume that N UEs, out of the total T UEs, request a specific content for delivery. The remaining $R = T - N$ UEs, who are currently not requesting any content, can act as the RUEs to facilitate the content dissemination by means of D2D relaying [14]. In summary, we define K as the total number of transmitting nodes in the network, including the GBS, UAVs, and RUEs (i.e., $K = 1 + U + R$). The index k is used to identify any individual transmitting node within this total.

The content can be delivered to the UEs directly by the GBS or relayed from the GBS via the UAV(s) and/or via the RUE(s). Both the UAVs and the RUEs operate in the half-duplex mode [15]. The consideration of full-duplex mode at the side of relays is relegated to the future work due to page limitations. To balance efficiency and complexity, we limit the number of relaying nodes used for the content delivery to two (resulting in up to 3-hop communication) [12]. Although, in theory, additional hops could further reduce the content delivery duration, it would come at the cost of significantly increased complexity [14].

B. Cache Placement Model

The GBS serves as the central repository for F contents and each UE can select any content from this repository. The size of the f-th content is denoted by S_f. In addition, each f-th content can be transmitted at any k-th transmission node over several transmission intervals while different power is allocated in each transmission interval (see Fig. 1, where "magenta" content is transmitted over five such transmission intervals at the GBS, etc.). Then, the part of the f-th content transmitted in the l-th interval send by the k-th node is further denoted as $S^k_{f,l}$.

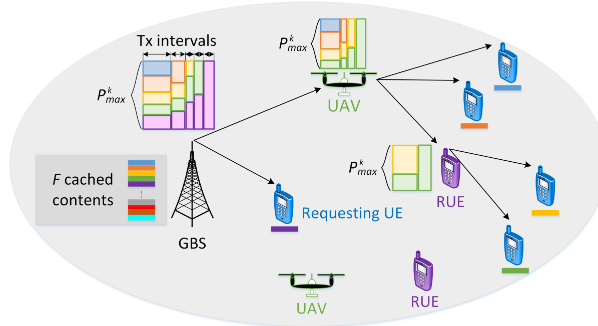

Fig. 1: System model encompassing multi-hop cache-enabled network. The cached contents (distinguished by different colors) are requested by different UEs and delivered to these either directly by GBS or relayed via UAVs and/or RUEs by means of D2D communication. Transmission power allocated to each content in individual transmission (Tx) intervals is also distinguished by colors.

A content popularity distribution, denoted by $\gamma = \{\gamma_1, \gamma_2, \ldots, \gamma_F\}$, captures the likelihood of the UEs requesting individual contents, where γ_f represents the probability of the UE requesting the content *f*. The sum of probabilities the UE requests the content is equal to 1 (i.e., $\sum_f \gamma_f = 1$), ensuring a single content selection per user. We model the content popularity with Zipf distribution, hence, $\gamma_f = f^{-\lambda} / \sum_{j=1}^{F} j^{-\lambda}$, where λ is the Zipf exponent and indicates the degree of skewness in the popularity.

C. Content Delivery Duration

We assume that the entire system bandwidth B is divided into N orthogonal channels, ensuring that the n-th requesting UEs is assigned with a dedicated channel with a bandwidth of B_n. Our proposed solution is applicable to any bandwidth allocation scheme. Thus, without loss of generality, we assume an equal bandwidth allocation among all requesting UEs (i.e., $B_n = B/N$).

Then, a delivery duration of f-th content requested by the n-th UE and transmitted by any k-th transmission node included in the communication route (i.e., the GBS, the UAVs, or the RUEs) is defined as:

$$t^k_{n,f} = \sum_l \frac{S^k_{f,l}}{B_n \log_2 \left(1 + \frac{p^k_{f,l} g^k_{n,f}}{B_n(\sigma_0 + I_b)}\right)}, \tag{1}$$

where $p^k_{f,l}$ denotes the transmission power of the k-th node to deliver the f-th content over l-th transmission interval, $g^k_{n,f}$ is the channel gain over which the k-th node is transmitting the f-th content on behalf of the n-th requesting UE, σ_0 is the noise power, and I_b represents the interference from other transmitting nodes.

III. Problem Formulation

The objective of this paper is to minimize the content delivery duration for all contents requested by the UEs by optimizing the power allocation among contents transmitted by the GBS directly or via multi-hop communication. The problem is formulated as:

$$\min_{p} \quad \sum_{k}\sum_{f}\sum_{l} \frac{S^{k}_{f,l}}{B_n \log_2\left(1+\frac{p^{k}_{f,l} g^{k}_{n,f}}{B_n(\sigma_0+I_b)}\right)} \tag{2a}$$

$$\text{s.t.} \quad \sum_{f} S^{k}_{f,l} = S_f, \ \forall f \tag{2b}$$

$$\sum_{f} p^{k}_{f,l} = P^{k}_{max}, \ \forall l,k \tag{2c}$$

$$0 \leq S^{k}_{f,l} \leq S_f, \ \forall f,l \tag{2d}$$

$$0 \leq p^{k}_{f,l} \leq P^{k}_{max}, \ \forall f,l,k \tag{2e}$$

where constraint (2b) ensures that the f-th content is fully delivered with all its parts by each k-th transmission node, (2c) guarantees the total power allocated by the k-th transmission node to all currently transmitted contents does not exceeds its maximum power budget P^{k}_{max} at any l-th transmission interval l, (2d) guarantees that the size of the k-th content transmitted during any l-th interval is non-negative and does not exceed the total size of content S_f, and (2e) ensures the power allocated to the f-th content during the interval l is non-negative and does not exceed the maximum power budget P^{k}_{max}.

The power allocation optimization problem introduced in (2a) is non-convex due to the coupled constraints involving $p^{k}_{f,l}$ and $S^{k}_{f,l}$. Even for single transmission node, the relationship between $p^{k}_{f,l}$ and $S^{k}_{f,l}$ is interdependent, as altering one affects the other. More specifically, setting $p^{k}_{f,l}$ to any arbitrary value impacts not only $S^{k}_{f,l}$ in the current transmission interval but also in all subsequent intervals due to constraint (2b). Moreover, changing $p^{k}_{f,l}$ affects the amount of data transmitted for other contents across all intervals, where those contents are being delivered. If multiple nodes are involved in the transmission of various contents across several hops, as considered in (2a), the power allocation becomes even more complex. The reason is that one must accounts also for the interactions between individual transmission nodes and the cumulative effects of power allocation across different transmission intervals. Consequently, in the following section, we introduce proposed heuristic algorithm for optimizing power allocation in multi-hop networks.

IV. Proposed Power Allocation

In this section, we first provide a high-level overview of the proposed concept of power allocation and, then, we introduce a novel heuristic algorithm to manage the power allocation.

A. High-level Description of the Proposed Power Allocation

In this paper, we employ dynamic re-allocation of the transmission power among simultaneously transmitted contents to minimize the content delivery duration for the requesting UEs. For clarity of presentation, we describe the proposed power allocation on a single k-th node transmitting multiple contents over a single hop in Fig. 2. Initially, the transmitting node allocates its power budget to four different contents. Once transmission of *content 1* is completed (i.e., when the whole *content 1* with a total size $S^{k}_{1,1}$ is transmitted), the power previously dedicated to the transmission of *content 1* becomes vacant and is redistributed to the transmission of the remaining parts of the contents still not completely transmitted, i.e., *content 2*, *content 3*, and *content 4*. After *content 2* is fully transmitted (i.e., $S^{k}_{2,2}$ is also transmitted), the power re-allocation process is repeated until all contents are fully transmitted.

The general power (re-)allocation for one content influences the power available to other contents and, consequently, also the transmission duration for all other contents. For instance, allocating more power to *content 1* reduces its transmission duration t_1 of this content; however, this adjustment decreases the power available for other contents, thus, potentially prolonging their transmission durations t_2-t_4. Nonetheless, because the transmission of *content 1* is completed sooner if more power is allocated to *content 1*, the vacant power after t_1 becomes available for other contents earlier, which may, in contrast, ultimately reduce their overall transmission durations of t_2, t_3, and t_4. Due to this dynamic interplay in power re-allocation, the power adjustment is done in advance before any content is being transmitted.

The proposed power (re-)allocation considers two distinct features to ensure fast convergence. To shed light on the first feature, let's first assume that at any k-th transmission node there is a set $\Omega_k(l)$ representing all possible combinations how contents, whose power re-allocation is not yet completed at the l-th transmission interval, can be paired together (e.g., in Fig. 2, there are six content pairs that can be created; *content 1* paired with *content 2*, *content 1* with *content 3*, and so on). Then, for a subset $L_k(l) \subseteq \Omega_k(l)$ of content pairs, the power re-allocation is done in parallel to speed up the re-allocation process (e.g., in Fig. 2, we can adjust in parallel transmission power of two content pairs, one pair can containing *content 1* and *2* while second pair is created from *content 3* and *4*). The second feature lies in the adaptive and dynamic setting of power allocation step, denoted for any content pair ρ as Δp_ρ. In particular, let's consider that each content pair ρ

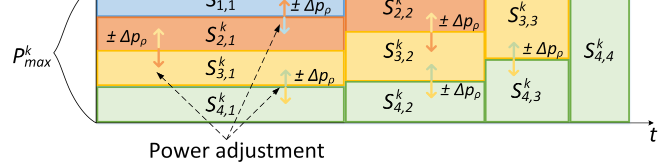

Fig. 2: The high-level principle of proposed power allocation at one transmitting node.

is composed of the contents C_i and C_j to which at any l-th transmission interval transmission power $p_{i,l}^k$ and $p_{j,l}^k$ is allocated, respectively. Then, Δp_ρ is dynamically updated to ensure fast convergence.

In the next section, we describe in detail the proposed power allocation algorithm and delve deeper into the principles of both above-mentioned features assuring quick convergence.

B. Proposed Heuristic Power Allocation Algorithm

This section introduce proposed power allocation principle summarized in Algorithm 1. Initially, Algorithm 1 allocates by default the transmission power at each k-th node equally among the currently transmitted contents as in [12] and calculates sum transmission time over all contents t_{sum} (line 1 in Algorithm 1). Then, the following steps are repeated until the convergence is reached. First, the algorithm randomly selects a subset $L_k(l)$ of content pairs from $\Omega_k(l)$ to perform the parallel power re-allocation (line 4). Subsequently, the algorithm evaluates different power allocations for content pair ρ in $L_k(l)$ with the objective to minimize the sum content delivery duration. In particular, the algorithm evaluates two cases for the re-allocation:

- **Case 1:** The power allocated to C_i is increased by Δp_ρ, while the power allocated to C_j is decreased by the same amount, resulting in temporary power allocations $p_{i,l}^{k,1} = p_{i,l}^k + \Delta p_\rho$ and $p_{j,l}^{k,1} = p_{j,l}^k - \Delta p_\rho$ (line 7).
- **Case 2:** The power allocated to C_i is decreased by Δp_ρ, while the power allocated to C_j is increased by same amount, resulting in the temporary power allocations $p_{i,l}^{k,2} = p_{i,l}^k - \Delta p_\rho$ and $p_{j,l}^{k,2} = p_{j,l}^k + \Delta p_\rho$ (line 8).

The algorithm initially sets $\Delta p_\rho = \Delta p_{\rho,1} = \min(p_{i,l}^k, p_{j,l}^k)/2$ to half of the power of the smaller allocated power out of the contents in the pair. This initial setting is chosen to promote rapid convergence of the power re-allocation process, thereby reducing the number of iterations required. Additionally, this setting prevents overly aggressive allocations that could destabilize converge.

Next, the algorithm updates the sum content delivery durations of the requesting UEs for each case, denoted as t_{sum}^1 and t_{sum}^2 for the first and second case, respectively (line 9). If t_{sum}^1 yields a lower sum content delivery duration than the t_{sum} and t_{sum}^2 (i.e., $t_{sum}^1 < t_{sum}$ and $t_{sum}^1 < t_{sum}^2$), the allocated powers are updated to $p_{i,l}^k = p_{i,l}^{k,1}$ and $p_{j,l}^k = p_{j,l}^{k,1}$, with t_{sum} being updated to t_{sum}^1 (line 11). Similarly, if t_{sum}^2 is lower than t_{sum} and t_{sum}^1, the allocated powers are updated to $p_{i,l}^k = p_{i,l}^{k,2}$ and $p_{j,l}^k = p_{j,l}^{k,2}$ while t_{sum} is set to t_{sum}^2 (line 14). If either of the two conditions ($t_{sum}^1 < t_{sum}$ and $t_{sum}^1 < t_{sum}^2$) or ($t_{sum}^2 < t_{sum}$ and $t_{sum}^2 < t_{sum}^1$) is met, Δp_ρ is updated employing the same formula used in the initial setting of Δp_ρ (i.e., $\Delta p_\rho = \Delta p_{\rho,1}$) (see line 12 and 15, respectively). The purpose of using this setting is to evaluate feasibility of a power allocation for the pair ρ that further minimizes the sum content delivery duration by leveraging its rapid converge capabilities. The algorithm iteratively updates Δp_ρ for the pair ρ using the initial setting until the iteration in which no power

Algorithm 1 Proposed heuristic algorithm for power allocation

1: Allocate $p_{f,l}^k$ equally $\forall f, l, k$ and calculate t_{sum}
2: **for** each currently transmitting node $k \in K$ **do**
3:   **repeat**
4:     Select $L_k(l) \subseteq \Omega_k(l)$
5:     **while** $\Omega_k(l) = \emptyset$ (parallel process for pairs in $L_k(l)$) **do**
6:       **for** each pair ρ with (C_i, C_j) in $L_k(l)$ **do**
7:         **Case 1:** $p_{i,l}^{k,1} = p_{i,l} + \Delta p_\rho$, $p_{j,l}^{k,1} = p_{j,l}^k - \Delta p_\rho$
8:         **Case 2:** $p_{i,l}^{k,2} = p_{i,l} - \Delta p_\rho$, $p_{j,l}^{k,2} = p_{j,l}^k + \Delta p_\rho$
9:         calculate t_{sum}^1 and t_{sum}^2
10:         **if** $t_{sum}^1 < t_{sum}$ and $t_{sum}^1 < t_{sum}^2$ **then**
11:           $p_{i,l}^k = p_{i,l}^{k,1}$, $p_{j,l}^k = p_{j,l}^{k,1}$, $t_{sum} = t_{sum}^1$
12:           Set $\Delta p_\rho = \Delta p_{\rho,1}$ or $\Delta p_\rho = \Delta p_{\rho,2}$
13:         **else if** $t_{sum}^2 < t_{sum}$ and $t_{sum}^2 < t_{sum}^1$ **then**
14:           $p_{i,l}^k = p_{i,l}^{k,2}$, $p_{j,l}^k = p_{j,l}^{k,2}$, $t_{sum} = t_{sum}^2$
15:           Set $\Delta p_\rho = \Delta p_{\rho,1}$ or $\Delta p_\rho = \Delta p_{\rho,2}$
16:         **else**
17:           Set $\Delta p_\rho = \Delta p_{\rho,2}$
18:         **end if**
19:       **end for**
20:       **if** $\Delta p_\rho < \Delta p_{\text{thr}}$ **then**
21:         $\Omega_k(l) = \Omega_k(l) \setminus \{\rho\}$
22:         update $L_k(l)$
23:       **end if**
24:     **end while**
25:   **until** convergence
26: **end for**

re-allocation case results in a shorter sum content delivery duration than the current minimum (i.e., $t_{sum} < t_{sum}^1$ and $t_{sum} < t_{sum}^2$).

In case that $t_{sum} < t_{sum}^1$ and $t_{sum} < t_{sum}^2$, the bounds of power re-allocation for the pair ρ are established. These bounds define the limits within which the power allocation for the pair ρ do not lead to any further reduction in the objective function. In other words, any power allocation outside these bounds results in the sum content delivery duration that fails to minimize the current minimum t_{sum}. Once the bounds of power re-allocation for the pair ρ are established, the gap between these bounds is iteratively narrowed until the power allocation that minimizes the objective function is found. Therefore, the algorithm updates Δp_ρ based on the current power allocation and the power re-allocations bounds to evaluate intermediate power allocations in order to minimize the objective function (i.e., $\Delta p_\rho = |(p_{i,l}^{k,1} - p_{i,l}^k)/2|$ or $\Delta p_\rho = |(p_{i,l}^{k,2} - p_{i,l}^k)/2|$), as the distances between the current power allocation and the both bounds are equivalent (line 17). Note that once during any iteration it is true that $t_{sum} < t_{sum}^1$ and $t_{sum} < t_{sum}^2$ for the content pair ρ, the algorithm sets Δp_ρ in subsequent iterations using $\Delta p_\rho = \Delta p_{\rho,2} = |(p_{i,l}^{k,1} - p_{i,l}^k)/2|$ instead of $\Delta p_{\rho,1}$ (see line 12 and 15, respectively).

The algorithm updates the current power allocations with the established bounds at each iteration until Δp_ρ falls below a specified threshold for the pair ρ (i.e., $\Delta p_\rho < \Delta p_{thr}$). Note that Δp_{thr} is set by trial and error to obtain reasonable trade-off between delay and number of iterations. Once the power re-allocation process for any pair ρ in $L_k(l)$ is completed, it is removed from $\Omega_k(l)$ (21). Further, the algorithm updates

$L_k(l)$ and new pair(s) from $\Omega_k(l)$ for whom the power re-allocation is done (i.e., lines 6-19 are repeated). Moreover, to ensure that the power re-allocation process is independent of the random pair selection, the process is repeated with a newly selected subset of content pairs $L_k(l)$ once the allocation for the content pairs at the k-th transmission node is completed (i.e., $\Omega_k(l) = \varnothing$) (lines 4-24). This repetition continues until power allocation converges across different subsets $L_k(l)$. Then, once the content is transmitted to the receiving node with the converged power allocation during the transmission interval l, the algorithm updates the size of not yet delivered part of the content for all contents currently being transmitted in the subsequent transmission interval.

The computational complexity of Algorithm 1 is influenced by two factors; *i)* number of the users requesting content (affecting number of contents to be transmitted by each node and number of transmission intervals) and *ii)* number of potential transmission nodes distributing individual contents to requesting users. We evaluate the number of iterations that need to be performed in the next section.

V. SIMULATION DESCRIPTION

In this section, we first outline the simulation scenario and settings and, then, we describe the competitive algorithms.

A. Simulation Scenario

The simulations are conducted using MATLAB. We consider a 500×500 m^2 reference cell representing an urban environment with multiple buildings, whose heights are randomly generated between 20 and 29 m [16]. The GBS is positioned at the building closest to the center of the simulated area, in line with [17]. For aerial coverage, we deploy 4 UAVs using the k-means clustering algorithm [18]. The UAVs are assumed to operate at an altitude of 100 m, thereby ensuring line-of-sight (LoS) communication with the GBS [19].

We consider a total of 100 users, with up to 30 users potentially requesting specific content, while the remaining users function as the RUEs to assist in content dissemination. Each content has a distinct size, varying between 1 and 10 Mbit. The channel model between any two nodes (UEs, UAVs and GBS) is based on a well-established model for UAV communication in urban environments as introduced in [20]. This model considers the presence of buildings in the communication path. Specifically, we determine whether there are buildings obstructing the LoS between a transmitter and a receiver by considering their positions. Each building obstructing the LoS path adds 20 dB of attenuation to the signal. The simulation parameters as summarized in Table I.

B. Compared algorithms

The proposed power allocation (labeled as "Proposal") is managed jointly with route selection introduced in [12]. Despite the fact that existing route selection is used, as we do not target specifically the route selection in this work, route selection is affected by proposed power allocation and different routes may be selected when compared to solution

TABLE I: Parameters and settings for simulations.

Parameter	**Value**
Area size	500×500 m^2
Carrier frequency	2 GHz
Bandwidth (B)	20 MHz
Max. Tx power of GBS ($P_{\max}^{\text{GBS}}$); UAV ($P_{\max}^{\text{UAV}}$); UE ($P_{\max}^{\text{UE}}$)	30; 30; 23 dBm
Noise (σ)	-174 dBm/Hz
Background interference (I_b)	-120 dBm/Hz
No. of contents (F)	50 files
Content size (S_f)	1-10 Mbits
No. of UAVs (U); UEs (T); UEs req. a content (N)	4; 100; 1-30
Zipf exponent (λ)	0.5

introduced in [12]. We compare the proposal with its two variants:

- *Proposal: w/o parallelization + fixed step size* - Power allocation of the content pairs executed sequentially with a fixed power allocation step size (Δp_ρ) instead of an adaptive step size.
- *Proposal: w/o parallelization* - Power allocation of the content pairs is executed sequentially rather than in parallel while Δp_ρ is updated during the power allocation process.

Further, the proposal is confronted with the following state-of-the-art algorithms:

- *Equal power allocation* - The greedy algorithm that jointly considers route selection and power allocation while employing an equal power split [12].
- *Iterative* - The route is first selected using a heuristic algorithm, followed by power allocation through an iterative evaluation process [8].

VI. SIMULATION RESULTS

Average content delivery duration of the proposed and competitive algorithms is investigated in Fig. 3a over the number of requesting UEs. Following this intuition, the average content delivery duration increases with the rising number of requesting UEs for all investigated algorithms. This is due to two main factors: *i)* the bandwidth B is divided among a larger number of requesting UEs, and *ii)* fewer relays remain available, as there are $100 - N$ UEs available for relaying. The proposal outperforms both the Equal power allocation and Iterative algorithms by up to 14.64% and 31.13%, respectively. Notably, the significant improvement over the iterative algorithm is attributed to our dynamic power re-allocation reflecting various channel quality over individual hops.

Fig. 3b shows the average content delivery duration versus the content size for the proposed and competitive algorithms for the different values of requesting UEs (i.e., $N = 15$ and $N = 30$). An increase in content size leads to a rise in the average content delivery duration for all algorithms across different values of requesting UEs. According to Fig. 3b, the proposed algorithm reduces the average content delivery duration by 12.7% to 17.6% and by 29.9% to 39.7% compared to the Equal power allocation and Iterative algorithms, respectively, for $N = 15$. Moreover, for $N = 30$, the proposed algorithm

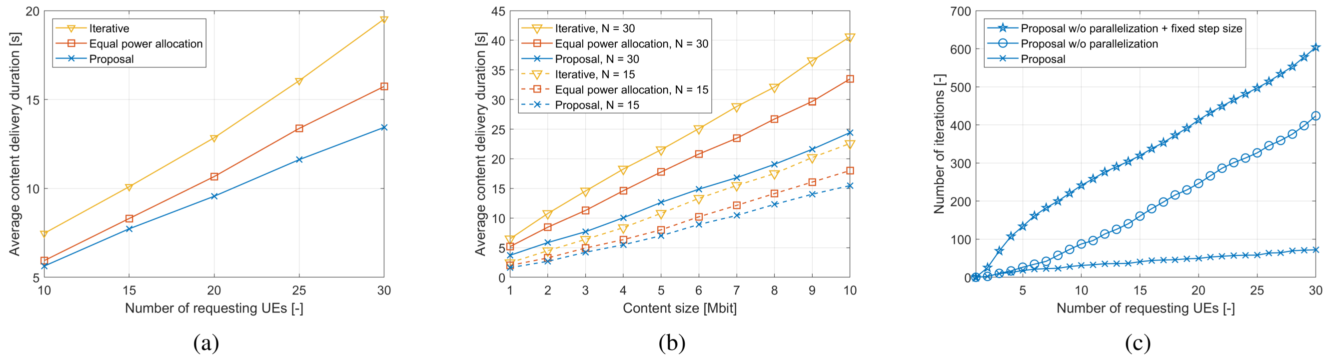

Fig. 3: Performance of the proposal and competitive state-of-the-art works in terms of (a) average content delivery duration over number of requesting UEs, (b) average content delivery duration over content size. Further, number of iterations of the proposal is investigated in (c).

reduces the average content delivery duration by 27.0% to 31.8% and by 39.8% to 47.1%, compared to the Equal power allocation and Iterative algorithms, respectively.

Last, in Fig. 3c, we analyze the number of iterations of the proposed algorithm as the number of requesting UEs increases. As the number of requesting UEs increases, power allocation across all algorithms necessitates more iterations due to the need for reallocating power among a greater number of content pairs and the additional iterations required to achieve convergence. The results confirm that parallelization is the primary factor in reducing the number of iterations required, while the incorporation of adaptive step size also contributes to this improvement. Specifically, the proposed algorithm reduces the number of required iterations by up to 83.04% and 88.13% compared to Proposal: w/o parallelization and Proposal: w/o parallelization + fixed step size, respectively. The results are encouraging in promoting the proposed algorithm in real networks as even with relatively high number of UEs simultaneously requesting the contents the number of iterations is very low and increasing linearly with number of requesting UEs.

VII. Conclusion

In this paper, we have proposed a novel heuristic algorithm for power allocation in cache-enabled multi-hop networks, aimed at minimizing the sum content delivery duration. In this regard, we have proposed iterative algorithm including parallelization and adaptive setting of transmission adjustment step between any content pairs. We have shown that the proposal reduces the average content delivery duration by more than 30% compared to related works while it exhibits a fast convergence.

References

[1] L. Li *et al.*, “A Survey of Caching Techniques in Cellular Networks: Research Issues and Challenges in Content Placement and Delivery Strategies,” *IEEE Commun. Surv. Tutor.*, vol. 20, no. 3, pp. 1710-1732, 2018.

[2] F. S. Shaikh and R. Wismüller, “Routing in multi-hop cellular device-to-device (D2D) networks: A survey,” *IEEE Commun. Surv. Tutor.*, vol. 20, no. 4, pp. 1710-1732, 2018.

[3] J. Yao *et al.*, “On mobile edge caching,” *IEEE Commun. Surv. Tutor.*, vol. 21, no. 3, pp. 2525-2553, 2019.

[4] D. Wang *et al.*, ”Deep Reinforcement Learning for Caching in D2D-Enabled UAV-Relaying Networks,” *IEEE ICCC*, 2021, pp. 635-640.

[5] L. Luo, R. Sun, R. Chai, and Q. Chen, ”Cost-Efficient UAV Deployment and Content Placement for Cellular Systems With D2D Communications,” *IEEE Systems Journal*, 2023.

[6] J. Ji *et al.*, “Joint trajectory design and resource allocation for secure transmission in cache-enabled UAV-relaying networks with D2D communications,” *IEEE Internet Things J.*, vol. 8, no. 3, pp. 1557-1571, 2020.

[7] M. S. Al-Abiad *et al.*, ”Throughput Maximization of Network-Coded and Multi-Level Cache-Enabled Heterogeneous Network,” *IEEE Trans. Veh. Technol.*, vol. 70, no. 10, pp. 11039-11043, 2021.

[8] A. Douik *et al.*, ”Mode selection and power allocation in multi-level cache-enabled networks,” *IEEE Commun. Lett.*, vol. 24, no. 8, pp. 1789-1793, 2020.

[9] B. Tian *et al.*, ”UAV-assisted wireless cooperative communication and coded caching: a multiagent two-timescale DRL approach,” *IEEE Trans. Mob. Comput.*, 2023.

[10] W. Wang *et al.*, ”Joint cooperative caching and power control for UAV-assisted internet of vehicles,” *Scientific Reports*, vol. 14, no. 1, pp. 9341, 2024.

[11] X. Qi, M. Yuan, Q. Zhang, and Z. Yang, ”Joint power-trajectory-scheduling optimization in a mobile UAV-enabled network via alternating iteration,” *China Communications*, vol. 19, no. 1, pp. 136-152, 2022.

[12] E. Gures and P. Mach, “Joint Route Selection and Power Allocation in Multi-Hop Cache-Enabled Networks,” *IEEE WCNC*, 2024, pp. 1-6.

[13] Y. Fu *et al.*, ”Dynamic power control for NOMA transmissions in wireless caching networks,” *IEEE Wireless Commun. Lett.*, vol. 8, no. 5, pp. 1485-1488, 2019.

[14] P. Mach, and Z. Becvar, ”Device-to-device relaying: Optimization, performance perspectives, and open challenges towards 6G networks,” *IEEE Commun. Surv. Tutor.*, vol. 24, no. 3, pp. 1336-1393, 2022.

[15] M. Najla, Z. Becvar, P. Mach, and D. Gesbert, ”Positioning and association rules for transparent flying relay stations,” *IEEE Wireless Commun. Lett.*, vol. 10, no. 6, pp. 1276-1280, 2021.

[16] P. Mach, Z. Becvar, and M. Najla, ”Power allocation, channel reuse, and positioning of flying base stations with realistic backhaul,” *IEEE Internet Things J.*, vol. 9, no. 3, pp. 1790-1805, 2021.

[17] P. Mach, Z. Becvar, and M. Nikooroo, ”Offloading of Tasks with Tight Delay Requirements via Combined Half and Full Duplex UAV Relays,” *IEEE Trans. Veh. Technol.*, 2024.

[18] T. Ren *et al.*, ”Enabling efficient scheduling in large-scale UAV-assisted mobile-edge computing via hierarchical reinforcement learning,” *IEEE Internet Things J.*, vol. 9, no. 10, pp. 7095-7109, 2021.

[19] M. Nikooroo, and Z. Becvar, ”Optimization of total power consumed by flying base station serving mobile users,” *IEEE Trans. Netw. Sci. Eng.*, vol. 9, no. 4, pp. 2815-2832, 2022.

[20] A. Al-Hourani *et al.*, ”Optimal LAP altitude for maximum coverage,” *IEEE Wireless Commun. Lett.*, vol. 3, pp. 569-572, 2014.