Multi-hop Relaying with Mixed Half and Full Duplex Relays for Offloading to MEC

Pavel Mach, Zdenek Becvar, Mohammadsaleh Nikooroo

***Abstract*—In this paper, we focus on offloading a computing task from a user equipment (UE) to a multi-access edge computing (MEC) server via multi-hop relaying. We assume a general relaying case where relays are energy-constrained devices, such as other UEs, IoT devices, or unmanned aerial vehicles. To this end, we formulate the problem as a minimization of the sum energy consumed by the energy-constrained devices under the constraint on the maximum requested time of the task processing. Then, we propose a multi-hop relaying combining half and full duplexes at each individual relay involved in the offloading. We proof that the proposed multi-hop relaying is convex, thus it can be optimized by conventional convex optimization methods. We show our proposal outperforms existing multi-hop relaying schemes in terms of probability that tasks are processed within required time by up to 38% and, at the same time, decreases energy consumption by up to 28%.**

***Index Terms*—offloading, MEC, half/full duplex relaying.**

I. Introduction

The multi-access edge computing (MEC) introduces a concept of offloading computationally demanding tasks from the energy-constrained user equipment (UE) to the MEC server located at the edge of mobile network [1]. Hence, the task processing delay and/or energy consumption of the UE can be reduced [2].

Benefits facilitated by MEC can be further augmented by a relaying of the tasks from the UE to the MEC servers via intermediate relay(s). The exploitation of neighboring UEs as relays and, thus, capitalizing on device-to-device (D2D) relaying concept [3], helps to minimize the task processing delay [4] or increase the number of tasks completed within a required time [5]. Moreover, an adoption of unmanned aerial vehicles (UAVs) acting as the relays can improve quality of experience to the UEs [6] or minimize their energy consumption [7]. Besides, the use of vehicles as the relays is considered in [8] to ensure a reliable offloading from the vehicles in the area without coverage of the MEC servers.

All above-mentioned works assume only two-hop relaying, i.e., only one relay is used in the offloading process. To fully grab the potential of the relays, multi-hop relaying for the offloading purposes has recently drawn an attention from researches. The multi-hop relaying is addressed from a perspective of balancing the load among MEC servers [9], minimizing the processing delay of the tasks offloaded from the vehicles to the MEC servers [10]-[12] or to other computing vehicles [13][14], or to offload the tasks from one UE to other neighboring computing UEs [15].

This work was supported by the Czech Science Foundation under the grant no. 23-05646S. The authors are with the Department of Telecommunication Engineering, Faculty of Electrical Engineering, Czech Technical University in Prague, Prague, 166 27 Czech Republic (email:machp2@fel.cvut.cz; zdenek.becvar@fel.cvut.cz; nikoomoh@fel.cvut.cz).

The primary objective of all existing studies on the offloading with multi-hop relaying is to find a proper route between the offloading UE and the MEC server or other computing UE. All works but [15] assume only less efficient half-duplex (HD) mode adopted at each relay with the task subsequently offloaded over each hop in individual time intervals, thus, increasing communication delay. The paper [15] considers full-duplex (FD) mode, however, the paper fully disregards the problem of self-interference (SI) with which the FD is inevitably plagued [16]. Moreover, none of the existing works optimize multi-hop relaying in terms of radio resource management including *i*) allocation of time slots at each hop, *ii*) allocation of transmission power of the offloading UE as well as relays, and *iii*) allocation of bandwidth at each hop.

Motivated by the above-mentioned gaps, the objective of this paper is to optimize radio resource management aspects of multi-hop relaying for the task offloading. Since the offloading UE and relays are usually energy-constrained (e.g., smartphones, UAVs, or IoT devices) we formulate the problem as the minimization of the sum energy consumed by the energy-constrained UEs involved in the multi-hop relaying under the constraint on the maximum processing time of the computing tasks. First, we propose several unique relaying cases combining HD and FD at each relay involved in multi-hop relaying. Note that existing works always assume the same relaying mode at all relays. Second, we adapt the general problem for each multi-hop relaying case and we prove its convexity so that we can solve it in an optimal way. Finally, we demonstrate that the proposal increases the probability of the tasks being processed within required time by up to 38% and, at the same time, decreases energy consumption by up to 28% with respect to state-of-the-art works.

II. System Model

This section first describes the network model. Then, communication and computing models are introduced.

A. Network model

We contemplate a scenario with one powerful MEC server located, for example, at the BS. Further, we assume one UE generating highly computationally demanding tasks offloaded to the MEC servers via multi-hop relaying. We consider the end-to-end relaying link is already established while leaving the joint optimization of relays selection and multiple UEs scenario for future work. In our work, the relays can be any energy-constrained devices, such as smartphones, IoT devices, UAVs, or vehicles. Moreover, we envisage that different types of relays can be exploited at each hop, e.g., the smartphone can be used as the first relay while vehicle or UAV as other relay.

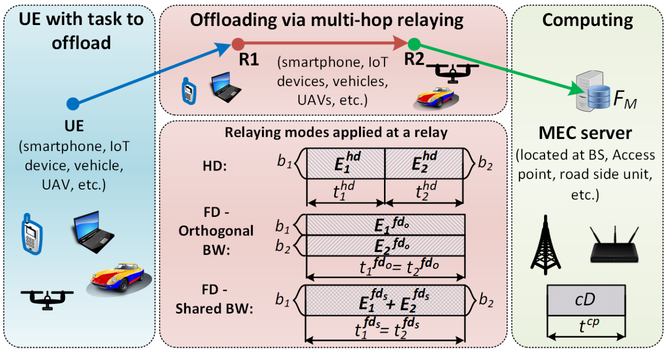


Fig. 1: System model with one UE having a task to offload via multi-hop relaying and computed at MEC server.

B. Communication and relaying models

We limit our scenario to two relays (labeled as R1 and R2 in Fig. 1), since this number is sufficient to illustrate benefits of the multi-hop relaying in the offloading while fitting page limit and avoiding cluttering of text and derivations. Still, all the math derivations and proposed relaying principles can be extended to more than two relays as well. We consider that each relay can adopt one of the three relaying modes: *i)* *HD*, *ii)* FD with orthogonal bandwidth at each hop (labeled as *FD–Orthogonal*), and *iii)* FD with the same bandwidth utilized at both hops (labeled as *FD–Shared*). All these modes are described in details in the following subsections.

1) HD: In case of HD relaying, the task is first sent over one hop and, then, the same task is relayed over the next hop (see Fig. 1). The capacity at the n-th hop is:

$$C_n^{hd} = b_n log_2\left(1 + \frac{p_n g_n}{b_n\left(\sigma + I_b\right)}\right), \qquad (1)$$

where b_n, p_n, g_n are the allocated bandwidth, transmission power, and channel gain at the n-th hop, respectively, σ is the noise spectral density, and I_b is the background interference from other UEs in the adjacent cells.

The communication delay in the HD is composed of the delays at individual hops (t_1^{hd}, t_2^{hd}), and is expressed as:

$$t^{hd} = t_1^{hd} + t_2^{hd} = \sum\nolimits_n t_n^{hd} = D\sum\nolimits_n \frac{1}{C_n^{hd}}, \qquad (2)$$

where D is the size of the task offloaded by the UE. Similarly as in many works (see, e.g., [9]), we neglect the delivery of the computing results back to the UE, as it is insignificant with respect to the whole communication delay.

The sum energy consumed to forward the task in HD is composed of the energies consumed at individual hops (E_n^{hd}):

$$E^{hd} = E_1^{hd} + E_2^{hd} = \sum\nolimits_n E_n^{hd} = \sum\nolimits_n t_n^{hd} p_n. \qquad (3)$$

2) FD–Orthogonal: In this mode, the relay receives and transmits the data simultaneously, but transmission at both hops are orthogonal in frequency domain (see Fig. 1) to avoid SI. Thus, we assume orthogonal bandwidth b_1 and b_2 at the first and second hops, respectively (we derive the optimal bandwidth allocation later in the paper). Then, the capacity $C_n^{fd_o}$ at the n-th hop is expressed as in (1).

Since the propagation delay and time for processing of communication at the relay (both jointly denoted as ϵ) are very short (scale of μs or ms) compared to the offloading time (hundreds of ms or seconds), these can be neglected without breaking a relaying causality and we can assume $t_1^{fd_o} = t_2^{fd_o} + \epsilon \approx t_1^{fd_o}$ for $\epsilon <<$ overall offloading time. In practice, the whole offloading task is transmitted in a series of many smaller transport blocks (in scale of ms in 5G) and each block can be forwarded right after its reception and processing by the relay, hence, fulfilling the relaying causality principle for the whole task. Thus, the communication delay is:

$$t^{fd_o} = t_1^{fd_o} = t_2^{fd_o} = \max(D/C_1^{fd_o}, D/C_2^{fd_o}), \qquad (4)$$

The energy consumption required to relay the task in FD with orthogonal bandwidth is calculated as:

$$E^{fd_o} = E_1^{fd_o} + E_2^{fd_o} = \sum\nolimits_n E_n^{fd_o} = \sum\nolimits_n t_n^{fd_o} p_n. \qquad (5)$$

3) FD–Shared: Similarly as in the previous FD case, the relays can receive and transmit data simultaneously (assuming $\epsilon <<$ offloading time as explained for *FD-Orthogonal bandwidth*). The fundamental difference is, however, that the transmissions at both hops share the same bandwidth. Then, the capacities at each hop are:

$$C_1^{fd_s} = b_1 log_2\left(1 + \frac{p_1 g_1}{b_1\left(\sigma + I_b\right) + p_2 g_{1,1}}\right), \qquad (6)$$

$$C_2^{fd_s} = b_2 log_2\left(1 + \frac{p_2 g_2}{b_2\left(\sigma + I_b\right) + p_1 g_{1,2}}\right), \qquad (7)$$

where $g_{1,1}$ is the channel gain between the transmitter and the receiver of the relay, thus, $p_2 g_{1,1}$ in (6) represents the SI in FD [16]; and $g_{1,2}$ is the channel gain between the transmitter at the first hop and the receiver at the second hop and $p_1 g_{1,2}$ representing interference in (7) from the former to the latter.

The communication delay is analogous to FD with orthogonal bandwidth, i.e.,:

$$t^{fd_s} = t_1^{fd_s} = t_2^{fd_s} = \max(D/C_1^{fd_s}, D/C_2^{fd_s}). \qquad (8)$$

The energy consumption in this relaying mode is defined as:

$$E^{fd_s} = E_1^{fd_s} + E_2^{fd_s} = \sum\nolimits_n E_n^{fd_s} = \sum\nolimits_n t_n^{fd_s} p_n. \qquad (9)$$

C. Computing model

We focus on the offloading of tasks to the MEC server that is able to process F_M CPU cycles per second. Let c is the average number of CPU cycles to process one bit of the task [2]. Then, we express the computing delay as:

$$t^{cp} = cD/F_M. \qquad (10)$$

III. Problem Formulation

We formulate a resource allocation problem to minimize the sum energy consumption for the offloading of the task from the UE over N (in this paper $N = 3$) hops as both the UE and relays are assumed to be energy-constrained while the task processing time meets the maximum required processing time T_{max}. This is achieved by optimization of time slots

$\mathcal{T}$, transmission power $\mathcal{P}$, and bandwidth allocation $\mathcal{B}$ at individual hops. Hence, the problem is formulated as:

$$\begin{aligned} \mathcal{T},\mathcal{P},\mathcal{B} = \ & \underset{t_n,p_n,b_n}{\text{argmin}} \sum\nolimits_n E_n \\ \text{s.t.} \quad & \text{(a)} \ \sum\nolimits_n t_n \leq T_{max} - t^{cp} \\ & \text{(b)} \ t_n > 0, \forall n \\ & \text{(c)} \ p_n \leq P_{max}, \forall n \\ & \text{(d)} \ b_n \leq B_{max}, \forall n \end{aligned} \qquad (11)$$

where (11a) ensures that task is processed within T_{max}, (11b) ensures that each time slot is positive, (11c) limits the transmission power at each hop to P_{max}, and (11d) guarantees bandwidth at any does not exceed B_{max}.

IV. Optimization of multi-hop relaying

In this section, we present the proposed relaying and its optimization. We distinguish three multi-hop relaying cases: *i*) both relays uses HD (labeled as *HD+HD*), *ii*) one relay use *HD* while *FD–Orthogonal* is employed by the other relay (*HD+FD–Orthogonal*), and *iii*) *HD* is exploited by the first relay while *FD–Shared* is used at the second relay (*HD+FD–Shared*). Note that the relaying modes at R1 and R2 can be switched for *ii*) and *iii*) with no impact on derivations presented in the paper. We optimize *i*)-*iii*) in the following subsections.

A. HD+HD relaying case

If both relays employ HD, the task offloading is done during three consecutive time slots (see Fig. 2). The task is sent first by the UE to the R1 within t_1^{hd}, then relayed by the R1 to the R2 during t_2^{hd}, and finally delivered from the R2 to the MEC server in t_3^{hd}.

To optimize $\mathcal{P}$ in (11), we express p_n from (1) as a function of t_n^{hd} while assuming $C_n^{hd} = D/t_n^{hd}$ (see (2)), i.e.:

$$p_n = \frac{K_n}{g_n}\left(2^{\frac{D}{t_n^{hd} b_n}} - 1\right), \qquad (12)$$

where $K_n = b_n(\sigma + I_b)$. Then, the sum energy consumption over all hops is expressed as:

$$\sum_n E_n = \sum_n t_n^{hd} p_n = \sum_n \frac{t_n^{hd} K_n}{g_n}\left(2^{\frac{D}{t_n^{hd} b_n}} - 1\right). \qquad (13)$$

To optimize $\mathcal{B}$ in (11), since the first derivative of E_n^{hd} with respect to b_n is decreasing with increasing b_n, E_n^{hd} at any n-th hop is minimized if $b_n = B_{max}$. Thus, we can rewrite (11) for optimizing *HD+HD* case as:

$$\begin{aligned} \mathcal{T} = \ & \underset{t_n^{HD}}{\text{argmin}} \sum\nolimits_n \frac{K_n t_n^{hd}}{g_n}\left(2^{\frac{D}{t_n^{hd} b_n}} - 1\right) \\ \text{s.t.} \quad & \text{(11a)} - \text{(11c)} \\ & \text{(d)} \ b_n = B_{max}, \forall n \end{aligned} \qquad (14)$$

where (14d) ensures that whole bandwidth is used at all hops.

Lemma 1. *The optimization problem in (14) and all its constraints are convex with respect to $\mathcal{T}$.*

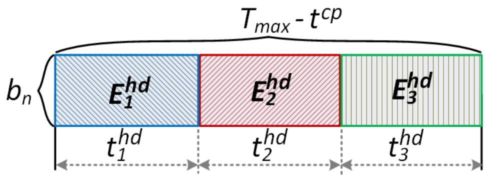


Fig. 2: Optimization of HD+HD by setting of each t_n^{hd}.

Proof. The Hessian matrix H corresponding to the objective function in (14) is:

$$H = \begin{bmatrix} L\frac{2^{\frac{D}{t_1^{hd} B_{max}}}}{t_1^{hd^3} g_1} & 0 & 0 \\ 0 & L\frac{2^{\frac{D}{t_2^{hd} B_{max}}}}{t_2^{hd^3} g_2} & 0 \\ 0 & 0 & L\frac{2^{\frac{D}{t_3^{hd} B_{max}}}}{t_3^{hd^3} g_3} \end{bmatrix} \qquad (15)$$

where $L = (\sigma + I_b)D^2 \ln^2 2/B_{max}$. The entries on the main diagonal of H are positive for $t_1^{hd} > 0$, $t_2^{hd} > 0$, and $t_3^{hd} > 0$. Since the diagonal matrix H is positive definite, the objective function in (14) is convex.

Further, the constraints (11a), (11b), and (14d) are linear, thus, also convex. Last, using (1) while considering $t_n^{hd} = D/C_n^{hd}$ (see (2)), any p_n in (11c) can be rewritten as:

$$t_n^{hd} \geq \frac{D}{B_{max}\log_2(1 + \frac{g_n P_{max}}{K_n})}, \qquad (16)$$

which is convex (linear) with respect to any $t_n^{hd} > 0$. ■

Since the optimization problem in (14) and all its constraints are convex, any convex optimization method can be used to solve it optimally. We have adopted CVX [17].

B. HD+FD–Orthogonal relaying case

The second case is the combination of *HD* (used by R1) and *FD–Orthogonal* (used by R2), see Fig. 3. Thus, the task is first sent to R1 during t_1^{hd} using b_1. Then, the task is simultaneously sent from R1 to R2 and from R2 to MEC server during $t_2^{fd_o} = t_3^{fd_o}$ (neglecting ϵ as explained in Section II.B) using b_2 and b_3, respectively. Note that relaying modes can be switched at R1 and R2.

Like for *HD*, we express p_n as in (12) for all hops as the function of time to solve $\mathcal{P}$ in (11). Then, the sum energy consumption is expressed as:

$$\sum_n E_n = E_1^{hd} + E_2^{fd_o} + E_3^{fd_o} = \frac{t_1^{hd} K_1}{g_1}\left(2^{\frac{D}{t_1^{hd} b_1}} - 1\right) + \\ + \frac{t_2^{fd_o} K_2}{g_2}\left(2^{\frac{D}{t_2^{fd_o} b_2}} - 1\right) + \frac{t_3^{fd_o} K_3}{g_3}\left(2^{\frac{D}{t_3^{fd_o} b_3}} - 1\right). \qquad (17)$$

Further, like for *HD*, we can assume $b_1 = B_{max}$, since this

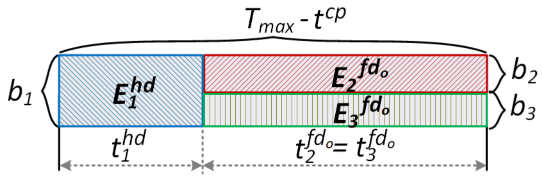


Fig. 3: Optimization of HD+FD - Orthogonal bandwidth by jointly setting t_1^{hd}, $t_2^{fd_o}$, b_2, and b_3.

minimize E_1^{hd}. Then, we can reformulate (11) as:

$$\begin{aligned} \boldsymbol{\mathcal{T}}, \boldsymbol{\mathcal{B}} = \; & \underset{t_1^{hd}, t_2^{fd_o}, b_2, b_3}{\operatorname{argmin}} \left(E_1^{hd} + E_2^{fd_o} + E_3^{fd_o} \right) \\ \text{s.t.} \quad & \text{(a) } t_1^{hd} + t_2^{fd_o} \leq T_{max} - t^{cp} \\ & \text{(b) } t_1^{hd} > 0, t_2^{fd_o} > 0 \\ & \text{(c) } b_2 + b_3 \leq B_{max} \\ & \text{(d) } b_2 > 0, b_3 > 0 \\ & \text{(e) } p_n \leq P_{max}, \forall n \end{aligned} \tag{18}$$

where (18a) and (18b) ensure that T_{max} is not violated and the duration of each time slot is positive, respectively, (18c) assures the sum bandwidth at the second and third hops is at most B_{max}, (18d) guarantees bandwidth of b_2 and b_3 is positive, and (18e) is the same as (11c).

Lemma 2. *The optimization problem in (18) and all its constraints are jointly convex with respect to $\boldsymbol{\mathcal{T}}$ and $\boldsymbol{\mathcal{B}}$.*

Proof. Similar as in the proof to Lemma 1, the first term in (17) is convex with respect to t_1^{hd}. To show that the second term in (17) is also convex, we prove that for any non-zero $v_1, v_2 \in R$ we have:

$$\begin{bmatrix} v_1 & v_2 \end{bmatrix} \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} > 0, \tag{19}$$

where $\begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix}$ is the Hessian matrix for $E_2^{fd_o}$ in (17) with respect to included variables $t_2^{fd_o}$ and b_2. To this end, the left-hand side in (19) is first expanded and rewritten as:

$$\begin{aligned} & \frac{\sigma + I_b}{g_2} ({b_2}^2 v_1^2 D^2 \ln^2(2)\alpha + {t_2^{fd_o}}^2 v_2^2 D^2 \ln^2(2)\alpha + \\ & 2 v_1 v_2 \big((t_2^{fd_o} b_2)^3 (\alpha - 1) - t_2^{fd_o} b_2 \alpha D \ln(2) \times \\ & (t_2^{fd_o} b_2 - D\ln(2)) \big)) > 0, \end{aligned} \tag{20}$$

where $\alpha = 2^{\frac{D}{t_2^{fd_o} b_2}}$. To prove (20) for any non-zero v_1, v_2, according to the Cauchy-Schwarz inequality, it is sufficient to prove that:

$$\begin{aligned} & (t_2^{fd_o} b_2)^3 (2^{\frac{D}{t_2^{fd_o} b_2}} - 1) - t_2^{fd_o} b_2 2^{\frac{D}{t_2^{fd_o} b_2}} \ln(2) \times \\ & (t_2^{fd_o} b_2 - \ln(2)) \geq -t_2^{fd_o} b_2 \ln^2(2) 2^{\frac{D}{t_2^{fd_o} b_2}}, \end{aligned} \tag{21}$$

or equivalently:

$$2^{\frac{D}{t_2^{fd_o} b_2}} ((t_2^{fd_o} b_2)^2 + 2\ln^2(2) - t_2^{fd_o} b_2 \ln(2)) \geq (t_2^{fd_o} b_2)^2. \tag{22}$$

The inequality in (22) always holds since $2^{\frac{D}{t_2^{fd_o} b_2}} > 1$ and $2\ln^2(2) - t_2^{fd_o} b_2 \ln(2) > 0$. Hence, the Hessian matrix is positive definite. Similarly, the Hessian matrix for $E_3^{fd_o}$ in (17) is positive definite with respect to the variables $t_3^{fd_o}$ and b_3. Last, the constraints (18a)–(18e) are all convex (linear) with respect to the optimization variables. ■

Due to convexity of (18) and all its constraints, we can again use CVX as in Section IV.A.

C. HD+FD–Shared relaying case

The last case is the one combining *HD* and *FD–Shared* (see Fig. 4). The offloading follows the same principle as in *HD+FD–Orthogonal*, but the transmissions at the second and third hops overlap also in frequency. Thus, the energy consumption at the second hop is affected by SI (see (6)) while the energy consumption at the third hop (at the MEC server) is impacted by interference from R1 (see (7)). Hence, to optimize $\boldsymbol{\mathcal{P}}$ in (11), the transmission power of R1 (i.e., p_2) is expressed from (6) while substituting $\{1, 2\} \Rightarrow \{2, 3\}$ as:

$$p_2 = \frac{K_2 + p_3 g_{2,2}}{g_2} \left(2^{\frac{D}{t_2^{fd_s} b_2}} - 1 \right). \tag{23}$$

Similarly, the transmission power of R2 (i.e., p_3) is calculated from (7) assuming $t_2^{fd_s} = t_3^{fd_s}$ and $K_2 = K_3$ (since $b_2 = b_3$, see Fig. 4), as:

$$p_3 = \frac{K_2 + p_2 g_{2,3}}{g_3} \left(2^{\frac{D}{t_2^{fd_s} b_2}} - 1 \right). \tag{24}$$

Next, we solve the system of equations formed by (23) and (24) in order to express p_2 and p_3 independently from each other and only in terms of the other parameters as follows:

$$\begin{aligned} p_2 &= \frac{K_2 \Gamma}{g_2 (1 - \beta \Gamma^2)} \left(1 + \frac{g_{2,2} \Gamma}{g_3} \right), \\ p_3 &= \frac{K_2 \Gamma}{g_3 (1 - \beta \Gamma^2)} \left(1 + \frac{g_{2,3} \Gamma}{g_2} \right), \end{aligned} \tag{25}$$

where $\beta = \frac{g_{2,3} g_{2,2}}{g_2 g_3}$, $\Gamma = 2^{\frac{D}{t_2^{fd_s} b_2}} - 1$.

Then, the energy consumption of this proposed relaying case is expressed as:

$$\begin{aligned} \sum_n E_n = E_1^{hd} + E_2^{fd_s} + E_3^{fd_s} = \frac{t_1^{hd} K_1}{g_1} \left(2^{\frac{D}{t_1^{hd} b_1}} - 1 \right) + \\ + \frac{t_2^{fd_s} K_2 \Gamma}{g_2 (1 - \beta \Gamma^2)} \left(1 + \frac{g_{2,2} \Gamma}{g_3} \right) + \frac{t_2^{fd_s} K_2 \Gamma}{g_3 (1 - \beta \Gamma^2)} \left(1 + \frac{g_{2,3} \Gamma}{g_2} \right). \end{aligned} \tag{26}$$

Since the whole B_{max} is used at all hops, as this minimizes the energy consumption (as explained in Section IV.A), the optimization problem in (11) can be formulated as follows:

$$\begin{aligned} \boldsymbol{\mathcal{T}} = \; & \underset{t_1^{hd}, t_2^{fd_s}}{\operatorname{argmin}} \left(E_1^{hd} + E_2^{fd_s} + E_3^{fd_s} \right) \\ \text{s.t.} \quad & \text{(a) } t_1^{hd} + t_2^{fd_s} \leq T_{max} - t^{cp} \\ & \text{(b) } t_1^{hd} > 0, t_2^{fd_s} > 0 \\ & \text{(11c), (14d)} \end{aligned} \tag{27}$$

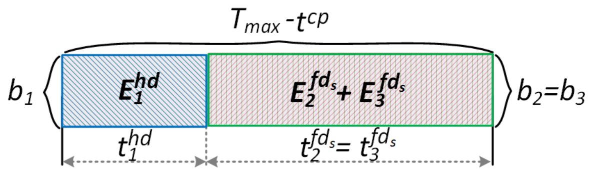


Fig. 4: Optimization of HD+FD - Shared bandwidth by setting t_1^{hd} and $t_2^{fd_s}$.

where the constraints are analogous to those in (14) and (18).

Lemma 3. *The optimization problem in (27) and all its constraints are convex with respect to* $\mathcal{T}$.

Proof. First, E_1^{hd} in (27) is convex with respect to t_1^{hd}, which is the only variable from $\mathcal{T}$ factoring in E_1^{hd}. This can be proved similarly as shown in the proof to Lemma 1.

Next, we show the convexity of $E_2^{fd_s}$ in (27). Since the Hessian matrix of $E_2^{fd_s}$ is too complex, we first decompose $E_2^{fd_s}$ into simpler factors and use the fact that, the multiplication of any positive, strictly decreasing, and convex functions is also convex. This fact can be verified via the equation $(fg)'' = f''g + fg'' + 2f'g'$ for positive, strictly decreasing, and convex arbitrary functions f and g. Now by considering the factors $t_2^{fd_s} K_2 \Gamma$ and $\frac{1}{g_2(1-\beta\Gamma^2)}$ in $E_2^{fd_s}$ taken from (26), both factors are positive, strictly decreasing, and convex with respect to $t_2^{fd_s}$. Hence, their multiplication, which yields the term $\frac{t_2^{fd_s} K_2 \Gamma}{g_2(1-\beta\Gamma^2)}$ in (27), is convex. In addition, the multiplication is also positive and strictly decreasing. Next, the term Γ and hence $\left(1+\frac{g_{2,2}\Gamma}{g_3}\right)$ in (26) is strictly decreasing and convex with respect to $t_2^{fd_s}$. Thus, its multiplication with the term $\frac{t_2^{fd_s} K_2 \Gamma}{g_2(1-\beta\Gamma^2)}$, which yields $E_2^{fd_s}$ in (27), is convex.

The third term $E_3^{fd_s}$ in (27) is also convex with respect to $t_2^{fd_s}$, as can be proven analogously to the convexity of $E_2^{fd_s}$.

Last, the constraints in (27) are also convex (linear) with respect to the optimization variables. ■

Since the optimization problem in (27) and all its constraints are convex, we solve it optimally by CVX, analogously as we solve (14) and (18) in Section IV.A and Section IV.B, respectively.

V. Performance Evaluation

In this section, we first describe the simulation models and parameters and then analyze the performance of individual relaying schemes.

A. Simulation models and parameters

For the performance evaluation, we assume the scenario with one UE offloading tasks via two relays to the MEC server (see Fig. 5). To get statistically valid results, we average out the results over 200 000 drops. Within each drop, we randomly generate: *i*) tasks parameters D and c, *ii*) distance between the UE and the MEC server between 25 and 150 m, and *iii*) positions of the relays within the areas shown in Fig. 5. This random generation of the relays' positions substitutes the relay selection process and each drop represents a case with relay at random positions. We adopt general modified COST 231 Hata path loss model at 2 GHz. All important simulation parameters are summarized in Table I.

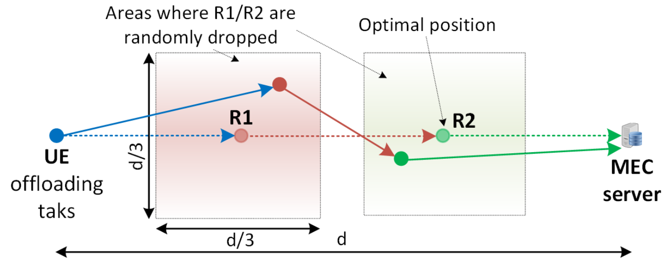


Fig. 5: Simulation scenario.

TABLE I: Parameters and settings for simulations

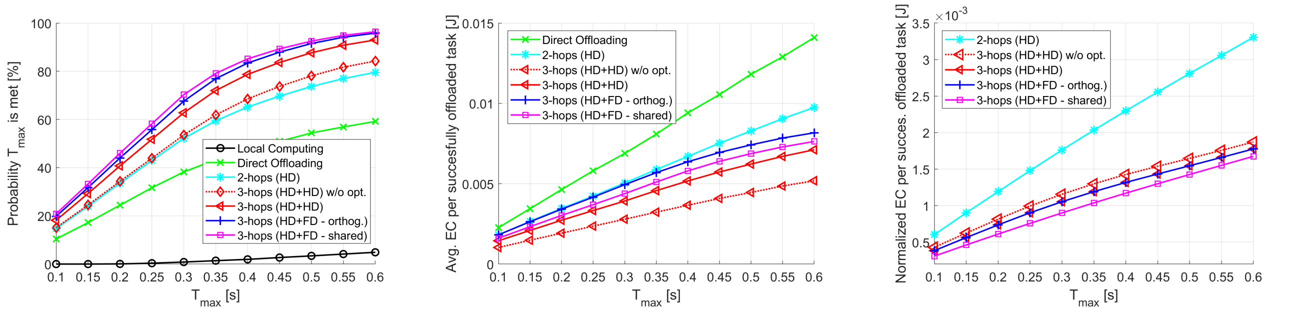
Parameter	Value	Parameter	Value
distance (d)	25-150 m	I_b	-150 dBm/Hz
Carrier freq.	2 GHz	D	[0.5 2] Mbits [2]
B_{max}	20 MHz	c	[1.5 2]x10^3 cyc./bit [2]
P_{max}	100 mW	F_u	[0.5 2]x10^9 cycles/s [2]
σ	-174 dBm/Hz	F_M	40x10^9 cycles/s [2]

We compare the results of multi-hop offloading with: *i*) local computing at the UE with computing power (F_u) randomly generated in each drop, *ii*) direct offloading to the MEC server without relays, *iii*) offloading via only relay working in HD usually considered in the related state-of-the-art works (denoted as 2-hops (HD)), *iv*) offloading via multi-hop (i.e., two relays) with HD at both relays while relaying is not optimized, as considered in [9]-[15] (denoted as 3-hops (HD+HD) w/o opt.)

B. Results

In Fig. 6a, we investigate the probability that the tasks are successfully processed within T_{max}. Following the intuition, the probability that the tasks are successfully processed increases with T_{max} for all investigated schemes, as there is more available time to process the tasks. Since we target tasks with relatively high requirements on computation, the local computing is not efficient with only up to 5% probability that tasks are processed within T_{max}. The direct offloading of tasks to MEC server significantly increases the probability that T_{max} is met (up to 60%). The introduction of relaying notably improves the performance of the offloading so that 2-hops relaying and 3-hops relaying without optimization leads to the probability of successful task processing within T_{max} up to 79.5% and 84.2%, respectively, if $T_{max} = 0.6$ s.

Now, let's discuss the probability of successful processing within T_{max} of the proposed optimized multi-hop relaying cases described in Section IV. The superior performance is provided by the multi-hop relaying combining *HD* and *FD–shared*, outperforming the direct offloading, 2-hops relaying, and conventional 3-hop relaying without optimization in terms of probability of T_{max} being met by up to 99.7%, 41.3%, and 38%, respectively. Among the proposed multi-relaying



(a) (b) (c)

Fig. 6: Effect of T_{max} on: a) probability that T_{max} is met, b) average energy consumption per successfully offloaded task, c) normalized energy consumption when only tasks successfully offloaded within T_{max} by direct offloading are considered.

approaches, the worst performance is observed for *HD+HD*, as it is less spectrum efficient. Still even this scheme outperforms direct computing, 2-hops relaying, and conventional 3-hop relaying without optimization in terms of probability T_{max} is met by up to 74.3%, 23.3%, and 20.4%, respectively.

In Fig. 6b, we analyze the average energy consumed per the task successfully processed within T_{max}. We demonstrate that the average energy consumption increases with T_{max} since, generally, offloading can take longer, thus consuming more energy. As expected, the highest energy consumption is spent by the direct offloading. The energy consumption is decreased by more than 30% if 2-hops relaying is introduced. Further significant decrease in the energy consumption (nearly 3 times compared to the direct offloading) is observed for the multi-hop relaying. If the multi-hop relaying is not optimized, it can even consume less energy than the proposed multi-hop relaying cases. This is due to the fact that the optimization of multi-hop relaying allows to accommodate also more demanding tasks within T_{max} that cannot be processed successfully without the proposed optimization. However, these more demanding tasks cost more energy during offloading. As a result, the average energy consumption per successfully offloaded task is increased.

To make the comparison of energy consumed per task fair, in Fig. 6c, we show the normalized energy consumption considering only the tasks successfully offloaded even by the direct offloading. Note that the direct offloading is omitted in Fig. 6c to keep a reasonable scale of the y-axis and also since the energy consumption for the direct offloading is, in fact, the same as in Fig. 6b. Fig. 6c demonstrates that the best performance is yielded by proposed multi-hop relaying combining *HD* with *FD–shared* as it decreases the energy consumption when compared to 2-hop and 3-hop relayings without optimization by up to 51.2% and 28%, respectively.

VI. Conclusions

In this paper, we have focused on offloading of highly computationally demanding tasks to MEC via multi-hop relaying. We have introduced and optimized several relaying cases combining half and full duplex relaying at individual relays. We have demonstrated multi-hop relaying improves the offloading experience while decreases the energy consumption of energy-constrained devices involved in the relaying. This paper is an initial work and a joint optimization of multi-hop relaying and relay selection should be carried out in the future.

References

[1] P. Mach, Z. Becvar, "Mobile Edge Computing: A Survey on Architecture and Computation Offloading," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 3., 3rd Quarter, pp. 1628-1656, 2017.

[2] T. Fang, *et al.*, "Joint Task Offloading, D2D Pairing, and Resource Allocation in Device-Enhanced MEC: A Potential Game Approach," *IEEE Internet Things J.*, vol. 9, no. 5, pp. 3226-3237, 2022.

[3] P. Mach and Z. Becvar, "Device-to-Device Relaying: Optimization, Performance Perspectives, and Open Challenges towards 6G Networks," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 3., 3rd Quarter, 2017.

[4] M. Li, *et al.*, "Multi-Relay Assisted Computation Offloading for Multi-Access Edge Computing Systems With Energy Harvesting," *IEEE Trans. Vehicular Tech.*, vol. 70, no. 10, Oct. 2021.

[5] J. Peng, *et al.*, "D2D-Assisted Multi-User Cooperative Partial Offloading, Transmission Scheduling and Computation Allocating for MEC," *IEEE Trans. Wireless Commun.*, vol. 20, no. 8, Aug. 2021.

[6] S. Zheng, Z. Ren, X. Hou, and H. Zhang, "Optimal Communication-Computing-Caching for Maximizing Revenue in UAV-Aided Mobile Edge Computing," In Proc. *GLOBECOM*, Taipei, Taiwan, 2020.

[7] X. Diao, *et al.*, "UAV-Relaying-Assisted Multi-Access Edge Computing With Multi-Antenna Base Station: Offloading and Scheduling Optimization," *IEEE Trans. Veh. Technol.*, vol. 70, no. 9, pp. 9495-9509, 2021.

[8] S. Zhang, *et al.*, "RAOA: A Relay-assisted Offloading Approach in Edge Computing Environment," *IEEE 24th CSCWD*, 2021.

[9] Y. Deng, *et al.*, "Task Offloading in Multi-Hop Relay-Aided Multi-Access Edge Computing," *IEEE Trans. Veh. Technol.*, 72(1), 2023.

[10] C.-M. Huanga, S.-Y. Linb, and Z.-Y. Wua, "The k-hop-limited V2V2I VANET data offloading using the Mobile Edge Computing (MEC) mechanism," *Vehicular Communications*, vol. 26, 2020.

[11] Z. Deng, *et al.*, "A Multi-Hop VANETs-Assisted Offloading Strategy in Vehicular Mobile Edge Computing," *IEEE Access*, vol. 8, 2020.

[12] Y. Deng, *et al.*, "Spectrum-aware Multi-hop Task Routing in Vehicle-assisted Collaborative Edge Computing," arxiv, Apr. 2023.

[13] C. Chen, Y. Zeng, H. Li, Y. Liu, and S. Wan, "A Multihop Task Offloading Decision Model in MEC-Enabled Internet of Vehicles," *IEEE Internet Things J.*, vol. 10, no. 4, Feb. 2023.

[14] L. Liu, *at al.*, "Mobility-Aware Multi-Hop Task Offloading for Autonomous Driving in Vehicular Edge Computing and Networks," *IEEE Trans. Intell. Transport. Syst.*, vol. 24, no. 2, Feb. 2023.

[15] J. Xie, *et al.*, "Dynamic D2D Multihop Offloading in Multi-Access Edge Computing From the Perspective of Learning Theory in Games," *IEEE Trans. Netw. Service Manag.*, vol. 20, no. 1, March 2023.

[16] H. Wang, *et al.*, "Spectrum sharing planning for full-duplex UAV relaying systems with underlaid D2D communications," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 9, pp. 1986–1999, Sep. 2018.

[17] M. Grant and S. Boyd. "CVX: MATLAB Software for Disciplined Convex Programming, Version 2.2." Jan. 2020. [Online]. Available: http://cvxr.com/cvx.