

# QTCAJOSA: Low-Complexity Joint Offloading and Subchannel Allocation for NTN-Enabled IoMT

Alejandro Flores C.; Konstantinos Ntontin; Ashok Bandi; and Symeon Chatzinotas

Interdisciplinary Centre for Security, Reliability and Trust (SnT), University of Luxembourg, Luxembourg.

e-mails:{alejandro.flores, kostantinos.ntontin, ashok.bandi, symeon.chatzinotas} @uni.lu

**Abstract**—In this work, we consider the resource allocation problem for task offloading from Internet of Medical Things (IoMT) devices, to a non-terrestrial network. The architecture considers clusters of IoMT devices that offload their tasks to a dedicated unmanned aerial vehicle (UAV) serving as a multi-access edge computing (MEC) server, which can compute the task or further offload it to an available high-altitude platform station (HAPS) or to a low-earth orbit (LEO) satellite for remote computing. We formulate a problem that has as objective the minimization of the weighted sum delay of the tasks. Given the non-convex nature of the problem, and acknowledging that the complexity of the optimization algorithms impact their performance, we derive a low-complexity joint subchannel allocation and offloading decision algorithm with dynamic computing resource initialization, developed as a greedy heuristic based on convex optimization criteria. Simulations show the gain obtained by including the different non-terrestrial nodes against architectures without them.

**Index Terms**—multi-access edge computing (MEC), non-terrestrial networks, resource allocation, task offloading, internet of medical things (IoMT).

## I. INTRODUCTION

As an enabler for the Healthcare 4.0 paradigm, the Internet of medical things (IoMT) allows for the acquisition and processing of biosignals recovered from patients, and shared with healthcare actors, such as hospitals and physicians. In the context of remote patient monitoring (RPM), some examples of IoMT devices include continuous glucose monitoring sensors [1] and portable electrocardiograms [2]. The continuous monitoring of biosignals of the patient allows for a more personalized healthcare by processing this data for detection or prediction of adverse patient conditions [3].

IoMT devices are usually resource- and energy-constrained, thus processing their generated tasks at remote multiaccess edge computing (MEC) servers allows for their timely execution. Nevertheless, less than 45% of the landmass of the Earth [4] and only 15% of its surface has mobile coverage [5], leaving many areas where IoMT devices cannot offload their tasks to the network. A cost-efficient solution to this problem is to make use of non-terrestrial networks (NTNs), which may comprise of unmanned aerial vehicles (UAVs), high altitude platform stations (HAPS) and satellites. Low-earth orbit (LEO) satellites are located 200-2000 km above Earth surface and have a beam footprint that can span up to a thousand kilometers, potentially having a large number of devices within service range. LEO satellites are also characterized by limited computing resources, high mobility, and non-negligible propagation delays, making the coordination of

a large number of devices challenging. A HAPS may offer a larger computing resource pool, while having a more limited coverage area. Moreover, UAVs can be deployed in an ad-hoc manner to gather IoMT-generated computing tasks, and computing them, or further offloading them for computation at a HAPS or a LEO.

The use of NTNs in task offloading has risen research interest in recent years. The problem of resource allocation in single-layer NTNs has been studied with UAVs [6], [7], HAPS [8], or LEO satellites [9], [10], as well as with multi-layer NTNs [11], [12]. The solutions in the existing literature do not consider the algorithm runtime delay into the optimization problems, which further constraints the maximum delay of the tasks. Moreover, the solutions to the offloading problem usually consider a fixed initialization of variables in iterative solutions, which has the potential to bias the offloading decisions, or to not properly advertise the resources available at remote MEC servers.

Given the large distances across air and space nodes, it is important for delay-sensitive computing tasks, to design efficient algorithms to solve the problem of resource allocation in such networks. With this in mind, and different from related literature, we present a novel, low-complexity, joint subchannel allocation and offloading decision algorithm, with dynamic resource initialization to better advertise the resources of remote nodes to the tasks generated by IoMT devices. We propose this over a 3-layer NTN comprised of UAVs, a resource-rich HAPS and a large-coverage LEO satellite. Adding to the methodology of the literature, we consider the algorithm runtime delay in the total delay of the tasks.

The remainder of this paper is structured as follows: Section II introduces the system model with the parameters of interest, whereas Section III presents the optimization problem to be solved. Section IV develops the resource allocation algorithm. In Section V simulation results are presented, and finally, in Section VI the conclusions are drawn.

## II. SYSTEM MODEL

Consider a set of ground IoMT devices  $\mathcal{I} = \{1, \dots, I\}$  separated in geographically distant clusters. A UAV, from set of UAVs  $\mathcal{U} = \{1, \dots, U\}$  is assigned to each cluster, where the set of devices associated with UAV  $u$  is  $\mathcal{I}_u \subseteq \mathcal{I}$  with  $|\mathcal{I}_u| = I_u$  and  $\sum_{u \in \mathcal{U}} I_u = I$ . In the system, there are a HAPS  $h$  and a LEO satellite  $s$  that act as MEC servers, with the HAPS acting as a coordinator across UAVs. The antenna elements at the HAPS and LEO satellite exhibit a smooth radiation pattern,

such as in [9], which affects the range of the scanning beams of the arrays, determining their coverage areas. The UAVs are assumed to receive a minimum level of service, which allows them to exchange signaling information with both the HAPS and LEO satellite through channels in the same band. Each node is located in Cartesian coordinates  $\mathbf{r}_k = [x_k, y_k, z_k]^T$ , with  $k \in \mathcal{I} \cup \mathcal{U} \cup \{h, s\}$ . A diagram of the system is shown in Fig. 1.

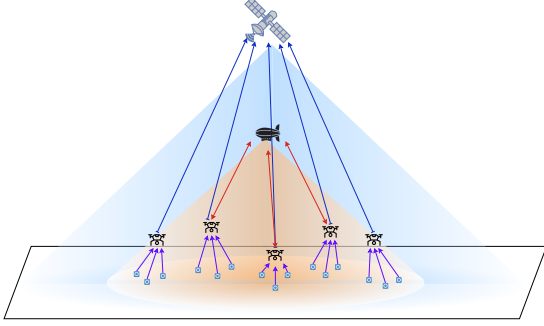


Fig. 1. System model

In a snapshot of the system, the IoMT devices generate tasks irregularly to be processed. Each task is represented as  $\psi_i = [d_i, c_i, \tau_i^{\max}]$ , where  $d_i$  is the number of bits in the task,  $c_i$  is its computational density (in CPU cycles per bit), and  $\tau_i^{\max}$  is its maximum tolerable delay (in seconds). Every IoMT device offloads its task to their local UAV, which acts as a MEC server and local coordinator for the cluster. The UAV can compute the task, or further offload it to the HAPS or to the LEO satellite for remote computation. The set of computing nodes is  $\mathcal{R} = \mathcal{U} \cup \{h, s\}$ .

A UAV communicates with its associated IoMT devices through orthogonal subchannels  $\mathcal{B} = \{1, \dots, B\}$ , of bandwidth  $B_u$ . In remote areas the spectrum is not used extensively, for which there is no need for high frequency reuse, allowing for orthogonal access. Thus, each IoMT device occupies one orthogonal subchannel, having no inter-IoMT device interference. Due to the directivity of the antenna arrays and the geographical separation between clusters, we assume no interference between clusters. For purposes of communication and signaling, the IoMT devices are single-antenna devices, while UAVs have a uniform linear arrays (ULAs) of  $N_u^{\text{ULA}}$  antenna elements. For communication between UAVs and the HAPS or LEO satellite, the UAVs have an upward-facing uniform planar array (UPA) of  $N_u^{\text{UPA}}$  antenna elements, while the HAPS and the LEO satellite have downward facing UPAs of  $N_h$  and  $N_s$  antenna elements respectively. They communicate through orthogonal subchannels, having no interference.

#### A. Communication model

1) *IoMT - UAV*: We consider that a given UAV applies a maximum-ratio combiner to the signal from IoMT device  $i$  communicating task  $\psi_i$ . Assuming a quasi-static channel in the S band, the instantaneous channel capacity of the link between IoMT device  $i$  to UAV  $u$  through subchannel  $b$  is given as

$$R_{i,u}^b = B_u \log_2 \left( 1 + \frac{\|\mathbf{h}_{i,u}^b\|^2 p_{i,u}^b}{B_u N_0} \right) \quad (1)$$

where  $p_{i,u}^b$  is the transmit power,  $N_0$  is the noise power per unit bandwidth, and  $\mathbf{h}_{i,u}^b = \sqrt{L_{i,u}^{-1}} \bar{\mathbf{h}}_{i,u}^b$  is the air-to-ground (A2G) channel response between  $i$  and  $u$  through subchannel  $b \in \mathcal{B}$ , where  $L_{i,u}$  is the pathloss between  $i$  and  $u$  given as in [13], and the small-scale fading component  $\bar{\mathbf{h}}_{i,u}^b$  is modeled as Rician fading as in [14], for which the expression for the corresponding Rician factor can be found in [15]. Let  $\rho_{i,u}^b$  be a binary variable indicating if subchannel  $b$  is allocated to device  $i$  by UAV  $u$  ( $\rho_{i,u}^b = 1$ ), or not ( $\rho_{i,u}^b = 0$ ). Then, the instantaneous channel capacity can be written as  $R_{i,u} = \sum_{b \in \mathcal{B}} \rho_{i,u}^b R_{i,u}^b$ .

2) *UAV - HAPS/LEO*: Having more simplified coordination between UAVs and the HAPS/LEO satellite, we assume that the corresponding links in the Ka band are orthogonal of bandwidth  $B_h$  and  $B_s$ , with channel matrices  $\mathbf{H}_{u,k} \in \mathbb{C}^{N_k \times N_u^{\text{UPA}}}$  with  $k \in \{h, s\}$ . Given the spatial sparsity of the links between the upward facing UPAs of the UAVs, and the downward facing UPA of the HAPS/LEO satellite, these channels are modeled as LoS channels undergoing free-space pathloss, as

$$\mathbf{H}_{u,k} = \left( \frac{\lambda_k g_k(\theta_{u,k}) \sqrt{G_{u,k}}}{4\pi \|\mathbf{r}_u - \mathbf{r}_k\|_2} \right) \mathbf{a}_{k,u}^{\text{AoA}} (\mathbf{a}_{u,k}^{\text{AoD}})^H, \quad (2)$$

where  $\mathbf{a}_{k,u}^{\text{AoA}}$  and  $\mathbf{a}_{u,k}^{\text{AoD}}$  are beamsteering vectors of the corresponding angle of arrival and angle of departure,  $G_{u,k}$  includes the atmospheric losses due to weather phenomena in the troposphere, and  $g_k(\theta_{u,k})$  models the normalized radiation pattern of the antenna elements of node  $k$ . For the LEO, we consider horn-type antenna elements as commercially available [16], and we choose a Bessel-based radiation pattern as in [17]. For the HAPS, we consider a patch antenna array as considered in the literature [18], and in commercial models [19] and consider a cosine radiation pattern.

The LoS-MIMO channel is rank-deficient, having only array gain, for which a singular-value decomposition is performed to obtain the precoder employed by the UAV and the combiner employed by the HAPS or LEO satellite. Then, the capacity of the channel is

$$R_{u,k} = B_k \log_2 \left( 1 + p_{u,k} \frac{\sigma_k^2}{B_k N_0} \right) \quad (3)$$

where  $\sigma_k$  is the non-zero singular value of the channel matrix. Let  $\beta_{u,k}^i$  be a binary variable indicating if task  $i$  is offloaded from UAV  $u$  to the node  $k$  ( $\beta_{u,k}^i = 1$ ) or not ( $\beta_{u,k}^i = 0$ ).

#### B. Delay model

We consider the delay by the transmission and computing of the task, and the delay by the signaling and algorithm runtime. The delay of reporting the results is typically not considered due to their small size compared to the size of the tasks.

1) *Transmission delay*: This delay is induced by the transmission of the tasks between nodes.

a) *IoMT - UAV*: The transmission delay experienced by offloading task  $\psi_i$  from device  $i$  to UAV  $u$  is given as

$$\tau_{i,u} = \frac{d_i}{R_{i,u}} = \frac{d_i}{\sum_{b \in \mathcal{B}} \rho_{i,u}^b R_{i,u}^b}. \quad (4)$$

b) *UAV - HAPS/LEO*: The transmission delay of offloading the required tasks from UAV  $u$  to node  $k \in \{h, s\}$ , is

$$\tau_{u,k} = \frac{\sum_{i \in \mathcal{I}_u} \beta_{u,h}^i d_i}{R_{u,h}} + 2 \frac{\|\mathbf{r}_k - \mathbf{r}_u\|_2}{c}. \quad (5)$$

The second term is the round-trip propagation delay, which is negligible for the UAV-HAPS link, but it is considered for the UAV-LEO link due to the large distances considered.

2) *Computing delay*: If node  $k \in \mathcal{R}$  computes task  $\psi_i$ , the corresponding computing delay is given as

$$\tau_i^k = \frac{d_i c_i}{f_i^k} \quad (6)$$

where  $f_i^k$  are the resources allocated for task  $\psi_i$  by node  $k$ .

3) *Algorithm Runtime Delay*: For real-time task offloading, the runtime of the resource allocation algorithm must be considered. To compute this, the number of basic operations in the algorithm (additions, multiplications and divisions<sup>1</sup>) is  $o_{op}$ , while the number of CPU cycles per operation is  $c_{op}$ , which depends on parameters such as the type of processor, number of cores and operation units [20]. Then, the algorithm runtime delay for centralized execution at node  $k$  is given as

$$\tau_{op} = \frac{o_{op} c_{op}}{F_k^{\max}} + \tau_k^{\text{RTT}}. \quad (7)$$

$F_k^{\max}$  is the total computing resources of node  $k$  and  $\tau_k^{\text{RTT}}$  is the maximum round trip time between ground nodes and  $k$ . The transmission times are not considered since the signaling overhead is assumed to be of much lower size.

### C. Weighted Sum Delay

The delay minimization may prioritize tasks with longer delay constraints, so to introduce fairness across tasks, a weighted sum delay is considered as

$$\hat{\tau} = \sum_{i \in \mathcal{I}} \frac{1}{\tau_i^{\max}} \left( \tau_{op} + \tau_{i,u} + \sum_{k \in \{h,s\}} \beta_{u,k}^i \tau_{u,k} + \left( 1 - \beta_{u,h}^i - \beta_{u,s}^i \right) \tau_i^u + \sum_{k \in \{h,s\}} \beta_{u,k}^i \tau_i^k \right) \quad (8)$$

We can write  $\hat{\tau} = \sum_{i \in \mathcal{I}} \frac{1}{\tau_i^{\max}} \tau_i$  with  $\tau_i = \tau_{op} + \bar{\tau}_{i,u}^u + \beta_{u,h}^i \bar{\tau}_i^h + \beta_{u,s}^i \bar{\tau}_i^s$ , where each term represents the delay of offloading and computing a task at a certain node. Here  $\bar{\tau}_i^u = \tau_{i,u} + \tau_i^u$ , and  $\bar{\tau}_i^k = \tau_{i,u} + \tau_{u,k} + \tau_i^k - \tau_i^u$  for  $k \in \{h, s\}$ .

### III. PROBLEM FORMULATION

Let  $\mathbf{P}_u = [\rho_1, \dots, \rho_I]$  with  $\rho_i = [\rho_{i,u}^1, \dots, \rho_{i,u}^B]^T$ ,  $\mathbf{B}_u = [\beta_{u,h}, \beta_{u,s}]$ , with  $\beta_{u,k} = [\beta_{u,k}^1, \dots, \beta_{u,k}^I]^T$  for  $k \in \{h, s\}$ , and  $\mathbf{F} = [\mathbf{f}^1, \dots, \mathbf{f}^U, \mathbf{f}^h, \mathbf{f}^s]$ , with  $\mathbf{f}^k = [f_1^k, \dots, f_I^k]^T$  for  $k \in \{u, h, s\}$ . Defining  $\mathcal{S} = \{\{\mathbf{P}_u\}_{u \in \mathcal{U}}, \{\mathbf{B}_u\}_{u \in \mathcal{U}}, \mathbf{F}\}$ , the optimization problem to be solved is formulated as follows:

$$\mathcal{P} : \min_{\mathcal{S}} \quad \hat{\tau} \quad (9a)$$

$$\text{s.t. } C_1 : \tau_i \leq \tau_i^{\max}, \quad \forall i \in \mathcal{I} \quad (9b)$$

<sup>1</sup>Usually, divisions are considered with a different number of cycles, but for simplicity we will include them with the basic operations.

$$C_2 : \sum_{i \in \mathcal{I}} f_i^k \leq F_k^{\max}, \quad \forall k \in \mathcal{R} \quad (9c)$$

$$C_3 : \sum_{b \in \mathcal{B}} \rho_{i,u}^b \leq 1, \quad \forall u \in \mathcal{U}, \forall i \in \mathcal{I}_u \quad (9d)$$

$$C_4 : \sum_{i \in \mathcal{I}} \rho_{i,u}^b \leq 1, \quad \forall u \in \mathcal{U}, \forall b \in \mathcal{B} \quad (9e)$$

$$C_5 : \beta_{u,s}^i + \beta_{u,h}^i \leq \sum_{b \in \mathcal{B}} \rho_{i,u}^b, \quad \forall u \in \mathcal{U}, \forall i \in \mathcal{I}_u \quad (9f)$$

$$C_6 : \beta_{u,h}^i, \beta_{u,s}^i \in \{0, 1\}, \quad \forall u \in \mathcal{U}, \forall i \in \mathcal{I} \quad (9g)$$

$$C_7 : \rho_{i,u}^b \in \{0, 1\}, \quad \forall u \in \mathcal{U}, \forall i \in \mathcal{I}_u, \forall b \in \mathcal{B} \quad (9h)$$

$$C_8 : f_i^k \geq 0, \quad \forall i \in \mathcal{I}, k \in \mathcal{U} \cup \{h, s\} \quad (9i)$$

where  $C_1$  constraints the maximum delay of every task,  $C_2$  constraints the maximum computing resources at the remote nodes,  $C_3$  constraints IoMT devices to access their local UAVs through at most one subchannel,  $C_4$  constraints the IoMT-UAV subchannels to be occupied by at most one IoMT device,  $C_5$  indicates a UAV can offload a task to the HAPS/LEO satellite only if the task has been offloaded to it, and  $C_6$  and  $C_7$  constraints the decision variables to be binary.

### IV. PROPOSED RESOURCE ALLOCATION ALGORITHM

To solve  $\mathcal{P}$  we propose a two-stage block-coordinate descent algorithm that performs the offloading decision of every task, followed by the optimal resource allocation at every computing node. The offloading decision is designed with dynamic resource initialization, based on optimality conditions. For this reason, the computing resource optimization is presented first.

1) *Remote Computing Resources*: We can express  $\mathcal{P}$  in terms of the remote computing resources, keeping the offloading variables fixed. This problem can be decoupled per remote node, and the maximum delay constraint can be transformed into a minimum computing resources constraint. Defining  $\mathcal{I}^k$  as the set of tasks computed in remote node  $k$ , each subproblem can be written as

$$\mathcal{P}_k^{\text{R-F}} : \min_{\{f_i^k\}_{i \in \mathcal{I}^k}} \sum_{i \in \mathcal{I}^k} \frac{d_i c_i}{\tau_i^{\max} f_i^k} \quad (10a)$$

$$\text{s.t.} \quad \sum_{i \in \mathcal{I}^k} f_i^k \leq F_k^{\max} \quad (10b)$$

$$f_i^k \geq \frac{d_i c_i}{\tau_i^{\max} - \tau_{i \rightarrow k}} = f_i^{k, \min} \quad \forall i \in \mathcal{I}^k \quad (10c)$$

$$f_i^k \geq 0 \quad \forall i \in \mathcal{I}^k \quad (10d)$$

where  $\tau_{i \rightarrow k}$  is the delay of task  $\psi_i$  to arrive to node  $k$ , given as  $\tau_{i \rightarrow u} = \tau_{i,u}$  and  $\tau_{i \rightarrow k} = \tau_{i,u} + \tau_{u,k}$  for  $k \in \{h, s\}$ , where  $\mathcal{I}_u^k$  are the set of tasks offloaded from UAV  $u$  to  $k \in \{h, s\}$ .

The solution to problem  $\mathcal{P}_k^{\text{R-F}}$  without considering the minimum-resources constraint, is given in closed-form as

$$f_i^{k,*} = \frac{\sqrt{\frac{d_i c_i}{\tau_i^{\max}}}}{\sum_{i \in \mathcal{I}^k} \sqrt{\frac{d_i c_i}{\tau_i^{\max}}}} F_k^{\max} \quad \forall i \in \mathcal{I}^k \quad (11)$$

The details are omitted due to page constraints. If a task is not feasible at the optimal point, the algorithm will allocate their minimum resources, if possible. If not, the task is computed at the local UAV. If the task can not be feasibly computed at the

UAV, it allocates  $f_i^{u,\min}$  to every task in ascending order, until it is not possible for a given task. The remaining resources are given to the remaining tasks equally in a best-effort solution.

The remote computation resource allocation sets computation resources for tasks not allocated to the remote node to zero, which stops a task from changing their computing node. To avoid this, the remote computing resource allocation problem is solved after the offloading decision problem.

2) *Quadratic Transform-enabled Computing-Aware Joint Offloading and Subchannel Allocation (QTCAJOSA)*: We write  $\mathcal{P}$  in terms of the subchannel allocation and offloading decision variables, while keeping the other variables fixed, as

$$\mathcal{P}^{\text{B-D}}: \min_{\substack{\{\mathbf{B}_u\}_{u \in \mathcal{U}} \\ \{\mathbf{P}_u\}_{u \in \mathcal{U}}}} \sum_{i \in \mathcal{I}} \left( \sum_{b \in \mathcal{B}} \rho_{i,u}^b \frac{v_{u,b}^i}{\tau_i^{\max}} \right. \\ \left. + \sum_{k \in \{h,s\}} \beta_{u,k}^i \frac{1}{\tau_i^{\max}} \left( \frac{\sum_{j \in \mathcal{I}} \beta_{u,k}^j d_j}{R_{u,k}} + v_k^i \right) \right) \quad (12a)$$

$$\text{s.t.} \quad C_1, C_3 - C_7 \quad (12b)$$

where  $v_{u,b}^i = \tau_i^u + \frac{d_i}{R_{i,u}^b}$ ,  $v_h^i = \tau_i^h - \tau_i^u$  and  $v_s^i = \tau_i^s - \tau_i^u + \frac{2\|\mathbf{r}_s - \mathbf{r}_u\|_2}{c}$ . To decouple the objective function, we apply the following algebraic transformation to the quadratic cross-term

$$\frac{\beta_{u,k}^i \sum_{j \neq i} d_j \beta_{u,k}^j}{R_{u,k}} = \frac{\left( \sum_{j \neq i} d_j \beta_{u,k}^j + \beta_{u,k}^i \right)^2 - \left( \sum_{j \neq i} d_j \beta_{u,k}^j - \beta_{u,k}^i \right)^2}{4R_{u,k}} \quad (13)$$

Then, we apply the quadratic transform (QT) [21] to each of the terms as

$$\frac{\left( \sum_{j \neq i} d_j \beta_{u,k}^j \pm \beta_{u,k}^i \right)^2}{R_{u,k}} \Rightarrow 2z_{\pm,i} \left| \sum_{j \neq i} d_j \beta_{u,k}^j \pm \beta_{u,k}^i \right| - z_{\pm,i}^2 R_{u,k} \quad (14)$$

where  $z_{\pm,i} = \frac{1}{R_{u,k}^{(n)}} \left| \sum_{j \neq i} d_j \beta_{u,k}^{j,(n)} \pm \beta_{u,k}^{i,(n)} \right|$  is computed from the results of the previous iteration. Assuming that previously allocated tasks cannot be removed from their allocated nodes, we can remove the absolute value, and it becomes

$$\frac{\beta_{u,k}^i \sum_{j \neq i} d_j \beta_{u,k}^j}{R_{u,k}} = \left( \frac{\beta_{u,k}^{i,(n)}}{R_{u,k}^{(n)}} \right) \sum_{j \neq i} d_j \beta_{u,k}^j + \left( \frac{\sum_{j \neq i} d_j \beta_{u,k}^{j,(n)}}{R_{u,k}^{(n)}} \right) \beta_{u,k}^i \\ - \left( \frac{\beta_{u,k}^{i,(n)} \sum_{j \neq i} d_j \beta_{u,k}^{j,(n)}}{(R_{u,k}^{(n)})^2} \right) R_{u,k} \quad (15)$$

Thus, the objective function is decoupled per-task such that  $\beta_{u,k}^i$  is weighted by the cost on the task  $i$  and on the impact on all of the already allocated tasks. The optimization problem over the offloading decision variables per UAV  $u$  is given as

$$\mathcal{P}_u^{\text{B-D}}: \min_{\mathbf{B}_u, \mathbf{P}_u} \sum_{i \in \mathcal{I}_u} \left( \sum_{b \in \mathcal{B}} \rho_{i,u}^b v_{u,b}^{i,(n)} + \beta_{u,h}^i v_h^{i,(n)} + \beta_{u,s}^i v_s^{i,(n)} \right) \quad (16a)$$

$$\text{s.t.} \quad \hat{C}_1, \hat{C}_3 - \hat{C}_7 \quad (16b)$$

where  $\hat{C}$  is constraint  $C$  only for the nodes  $i \in \mathcal{I}_u$ , and

$$v_k^{i,(n)} = \frac{1}{\tau_i^{\max}} \left( v_k^i + \frac{d_i}{R_{u,k}} + \frac{\sum_{j \neq i} d_j \beta_{u,k}^{j,(n)}}{R_{u,k}} \right) + \frac{d_i \sum_{j \neq i} \frac{\beta_{u,k}^{j,(n)}}{\tau_j^{\max}}}{R_{u,k}} \quad (17)$$

for  $k \in \{h, s\}$ . With the problem expressed in this form, we define a greedy algorithm that allocates tasks one by one, checking the best allocation at each step. The solution to this problem depends on the initializations of the  $f_i^k$  variables. If  $f_i^k < f_i^u$  for a given  $i$ , the task will not be offloaded. This happens even if the total pool of resources that node  $k$  can provide to task  $i$  is much larger. For a large number of tasks in the system, this can be true for every task, and no offloading takes place. To address this issue, we propose two measures:

a) *Non-orthogonal Computing Resource Initializations*:

Let  $F_{k,u}^{\max}$  be the total resources available at remote node  $k$  allocated to UAV  $u$ . We consider,  $F_{k,u}^{\max} = F_k^{\max}$ , thus  $f_i^k$  is initialized considering that the serving UAV of  $i$  can make use of all the resources of node  $k$ . This results in overlapping resources across UAVs, thus after each allocation, the best task across the ones from each UAV is the one allocated.

b) *Dynamic Computing Resource Initialization*: After each allocation, the computing resources at each remote node  $k \in \mathcal{R}$  are initialized only for the non-allocated tasks. Let  $\hat{\mathcal{I}} \subseteq \mathcal{I}$  be the set of tasks not yet allocated. At the start of each round the computation resources  $f_i^k$  for  $i \in \hat{\mathcal{I}}$  is initialized as (11), for  $k \in \mathcal{R}$ . This is implemented for the UAV and LEO satellite and not the HAPS, given that the HAPS has a larger pool of resources that could deplete early with this method.

Then, considering checks for maximum delay feasibility, we present the quadratic transform-enabled computing-aware joint offloading and subchannel allocation (QTCAJOSA) algorithm in Algorithm 1. The total number of operations for QTCAJOSA is  $I(6(I^2) + 6IB + 100I + 3B + 100)$ , and for the remote computing resources allocation algorithm is  $(U + 2)(2(I_u^2) + 24I_u + 15)^2$ .

## V. RESULTS

To evaluate the proposed algorithm and architecture, we develop simulations comparing our results with benchmarks where there is no LEO satellite present (No LEO), no HAPS present (No HAPS), and where the offloading decision problem is not adaptive (Non-adaptive), having the computing resources initialized in a non-overlapping manner at the start and do not change throughout. Following 3GPP recommendations [22], we consider the carrier frequency for the UAV-LEO/HAPS link in the Ka band, as 28GHz with bandwidths  $B_s = 200\text{MHz}$  and  $B_h = 100\text{MHz}$ . For the IoMT-UAV links we assume LTE-M enabled IoMT devices working in LTE channel [23], thus we choose  $B_u = 1.4\text{MHz}$  and  $B = 14$

<sup>2</sup>We assume that additions, multiplications and divisions count as elementary operations and that the number of elementary operations in other familiar functions is 10, to not consider a specific methodology or CPU architecture.

---

**Algorithm 1: QTCAJOSA**


---

```

1 foreach  $u \in \mathcal{U}$  do
2   Initialize  $\beta_u^h = \mathbf{0}_{I_u \times 1}$ ,  $\beta_u^s = \mathbf{0}_{I_u \times 1}$  and  $\rho_u = \mathbf{0}_{I_u \times B_u}$ ;
3   Define  $\mathbf{v}_u^h = \mathbf{v}_u^s = \mathbf{v}_u^u = \mathbf{1}_{I_u \times 1}$  and  $\mathbf{V}_u = \mathbf{1}_{I_u \times B_u}$ ;
4   Initialize  $\hat{\mathcal{I}}_u = \mathcal{I}_u$ ;
5 repeat
6   foreach  $u \in \mathcal{U}$  do
7     Initialize  $f_i^u$  as in (11)  $\forall i \in \hat{\mathcal{I}}_u$ ;
8     Compute  $\mathbf{U}_u$  where  $\mathbf{U}_u(i, b) = v_{u,b}^i$ ;
9     Compute  $\bar{\mathbf{U}}_u := \mathbf{U}_u \odot \mathbf{V}_u \odot \mathbf{v}_u^u$  (column-wise);
10    foreach  $k \in \{h, s\}$  do
11      Initialize  $f_i^k$  as in (11)  $\forall i \in \hat{\mathcal{I}}_u$ ;
12      Compute  $\bar{\mathbf{u}}_u^k(i) = [v_{i,k}^1, \dots, v_{i,k}^{I_u}]^T$ ;
13      Compute  $\bar{\mathbf{u}}_u^k := \bar{\mathbf{u}}_u^k \odot \mathbf{v}_u^k$ ;
14      Compute  $\bar{\mathbf{U}}_u^k = \mathbf{U}_u + \bar{\mathbf{u}}_u^k$  (column-wise);
15    if  $\min\{\bar{\mathbf{U}}_u^k\}_{k \in \mathcal{U} \cup \{h, s\}} = +\infty$ , then return;
16    Choose  $(i^*, u^*, b^*, k^*) = \operatorname{argmin}\{\bar{\mathbf{U}}_u^k(i, b)\}_{k \in \mathcal{U} \cup \{h, s\}}$ ;
17    if  $\tau_{i^*} > \tau_{i^*}^{\max}$  then
18      Set  $f_{i^*}^* = F_{u^*, k^*}^{\max}$ ;
19      if  $\tau_{i^*} > \tau_{i^*}^{\max}$  then
20        Set  $\mathbf{v}_{u^*}^{k^*}(i^*) = \infty$ ;
21        Go back to 2;
22    Set  $\mathbf{V}_{u^*}(i^*, :)$  and  $\mathbf{V}_{u^*}(:, b^*)$  to  $+\infty$ ;
23    Set  $\mathbf{v}_{u^*}^{k^*}(i^*)$  to  $+\infty$ ;
24    Set  $\hat{\mathcal{I}}_{u^*} := \hat{\mathcal{I}}_{u^*} \setminus \{i^*\}$ ;
25    Set  $\rho_{i^*, u^*}^b = 1$ ;
26    if  $k^* \in \{h, s\}$ , then Set  $\beta_{u^*, k^*}^{i^*} = 1$ ;
27    Update  $F_{u^*}^{\max} := F_{u^*}^{\max} - f_{i^*}^*$ ;

```

---

subchannels over the carrier at 2.1GHz. The altitude of the UAVs is 120m, according to the maximum altitude set by the European Union Aviation Safety Agency [24]. The altitude of the LEO satellite and the HAPS are 500 km and 20 km, both at horizontal coordinates  $[0, 0]$  at the time of algorithm execution. We consider clear-sky conditions such that  $G_{u,h} = G_{u,s} = 1$ . For the tasks, we consider  $\tau_i^{\max} = 30s$ ,  $c_i = 100$  and  $d_i$  to follow an uniform distribution  $U(100Kb, 10Mb)$ , where larger sizes represent tasks gathered from IoMT devices such as wearable ultrasound devices for image processing, or continuous high-frequency time-series processing from wearable ECGs, and smaller sizes represent tasks such as time-series processing from IoMT devices such as wireless stethoscopes or wearable pulse oximeters.

Fig. 2 shows the total delay for varying maximum ground distance of an IoMT device to the center of coordinates  $d_{\max}$ , where curves are presented for different values of  $F_s$ . The delay obtained by using the HAPS and the LEO satellite as MEC servers is smaller than the one obtained by not using either. There is a single curve for the no-LEO case, since only the HAPS is present and not affected by the resources at the LEO, as well as for the non-adaptive case because the initialized resources at the LEO are so small that it is never considered for offloading. The no-LEO curve is lower at smaller values of  $d_{\max}$ , since the nodes are within coverage of the HAPS, but also, it is closer to the all-nodes case when there are fewer resources at the LEO. As  $d_{\max}$  increases, the

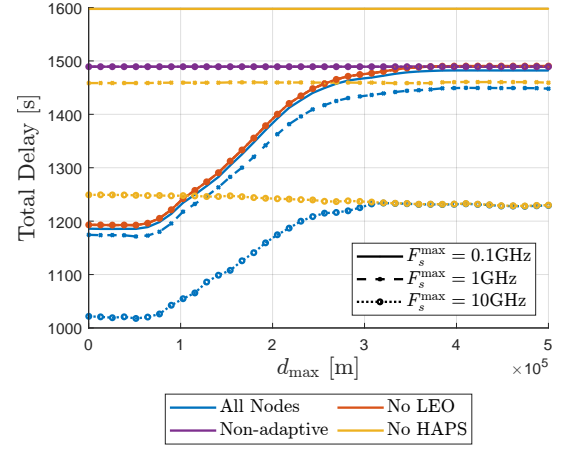


Fig. 2. Total delay across tasks for varying ground maximum distance. UAVs fall outside coverage of the HAPS, and the transmission delay increases. As the resources available at the LEO satellite increase, the gap between the all-nodes case and the no-HAPS case decrease, since the resources at the LEO satellite are further utilized to compute the tasks not done so at the HAPS.

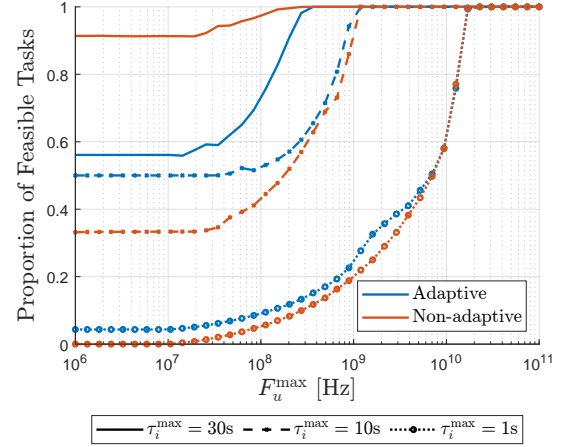


Fig. 3. Proportion of feasible tasks for varying UAV available resources.

Fig. 3 presents the proportion of feasible tasks in the system for varying  $F_u^{\max}$ , with curves for the proposed adaptive algorithm, and the non-adaptive version of the algorithm. The amount of feasible tasks in the system increases with the available resources at the UAVs, since they can compute more tasks. When the delay constraints are tighter, the proposed adaptive algorithm enables the fulfillment of more tasks feasibly due to its dynamic initialization and broadcasting of resources. If the delay constraints are looser, most of the tasks can be feasibly computed at the HAPS with the first initialization of resources, making the non-adaptive algorithm achieve higher number of feasible tasks than the adaptive algorithm, which may unevenly allocate resources across tasks.

Fig. 4 presents the proportion of tasks in the system offloaded to the HAPS or the LEO satellite for varying  $c_i$ , with three different values of  $F_s^{\max}$ . As the computing density increases, so does the offloading of the tasks to remote nodes because the computing density penalizes the computing delay at the UAV, but not the transmission delay of the task to the



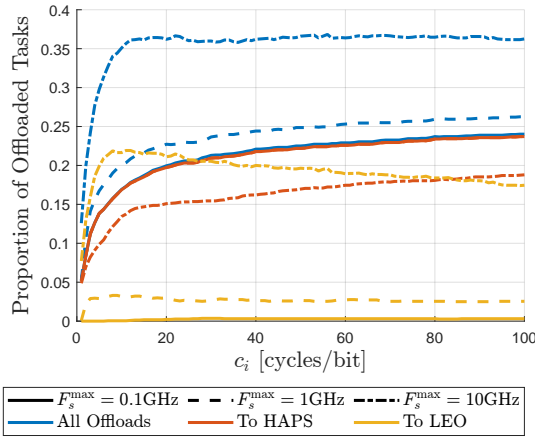


Fig. 4. Proportion of offloaded tasks per node for varying computation density.

remote nodes. As the resources at the LEO satellite increase, more tasks are offloaded to it. Nevertheless, for the case of  $F_s^{\max} = 10\text{GHz}$ , many more tasks are offloaded to the LEO rather than the HAPS, despite having the same amount of resources. This occurs because the HAPS initializes its resources at every step for all of the original tasks to avoid early depletion of resources, whereas for the LEO it offloads only to non-allocated tasks.

## VI. CONCLUSIONS

This work studied the problem of resource allocation for task offloading of IoMT-generated tasks to a NTN consisting of dedicated UAVs per cluster of ground devices, a common HAPS and a common LEO satellite acting as MEC servers. The solution was given as a low-complexity algorithm that optimizes the joint subchannel allocation and offloading decision, while having a dynamic computing resource initialization. It was shown how the HAPS offers more resources within its coverage, the LEO provides wide coverage with shared resources, and the UAV serves its specific IoMT devices. Adaptive computing resource initialization in offloading decisions reduces delay by better using smaller resource pools, while non-adaptive initialization may suit very large shared pools. Future work could explore optimizing other system resources and alternative algorithm distribution methods.

## ACKNOWLEDGEMENT

This work received funding from Horizon Europe under the Marie Skłodowska Curie actions: ELIXIRION (GA 101120135).

## REFERENCES

- [1] B. M. Ahmed, M. E. Ali, M. M. Masud, and M. Naznin, "Recent trends and techniques of blood glucose level prediction for diabetes control," *Smart Health*, vol. 32, p. 100457, 2024.
- [2] W. Qi and H. Su, "A cybertwin based multimodal network for ecg patterns monitoring using deep learning," *IEEE Trans. Ind. Informat.*, vol. 18, no. 10, pp. 6663–6670, 2022.
- [3] J. Medina, M. Espinilla, A. García-Fernández, and L. Martínez, "Intelligent multi-dose medication controller for fever: From wearable devices to remote dispensers," *Computers & Electrical Engineering*, vol. 65, pp. 400–412, 2018.
- [4] GSMA Intelligence, "The mobile economy 2024," 2024, accessed: 2024-10-16. [Online]. Available: <https://www.gsma.com/mobileeconomy/>

- [5] World Economic Forum, "What is a satellite-connected smartphone – and how can it help bridge the digital divide?" October 2022, accessed: 2024-10-16. [Online]. Available: <https://www.weforum.org/agenda/2022/10/satellite-connected-smartphones-digital-divide/>
- [6] J. Zhang, H. Luo, X. Chen, H. Shen, and L. Guo, "Minimizing response delay in uav-assisted mobile edge computing by joint uav deployment and computation offloading," *IEEE Trans. Cloud Comput.*, vol. 12, no. 4, pp. 1372–1386, 2024.
- [7] M. Hevesli, A. M. Seid, A. Erbad, and M. Abdallah, "Task offloading optimization in digital twin assisted mec-enabled air-ground iiot 6g networks," *IEEE Trans. Veh. Technol.*, vol. 73, no. 11, pp. 17 527–17 542, 2024.
- [8] Q. Ren, O. Abbasi, G. K. Kurt, H. Yanikomeroglu, and J. Chen, "Caching and computation offloading in high altitude platform station (haps) assisted intelligent transportation systems," *IEEE Trans. Wireless Commun.*, vol. 21, no. 11, pp. 9010–9024, 2022.
- [9] C. Ding, J.-B. Wang, H. Zhang, M. Lin, and G. Y. Li, "Joint optimization of transmission and computation resources for satellite and high altitude platform assisted edge computing," *IEEE Trans. Wireless Commun.*, vol. 21, no. 2, pp. 1362–1377, 2022.
- [10] Y. Lyu, Z. Liu, R. Fan, C. Zhan, H. Hu, and J. An, "Optimal computation offloading in collaborative leo-iot enabled mec: A multiagent deep reinforcement learning approach," *IEEE Trans. Green Commun. Netw.*, vol. 7, no. 2, pp. 996–1011, 2023.
- [11] N. Waqar, S. A. Hassan, A. Mahmood, K. Dev, D.-T. Do, and M. Gidlund, "Computation offloading and resource allocation in mec-enabled integrated aerial-terrestrial vehicular networks: A reinforcement learning approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 11, pp. 21 478–21 491, 2022.
- [12] Q. Chen, W. Meng, T. Q. S. Quek, and S. Chen, "Multi-tier hybrid offloading for computation-aware iot applications in civil aircraft-augmented sagin," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 2, pp. 399–417, 2023.
- [13] A. Al-Hourani, S. Kandeepan, and S. Lardner, "Optimal lap altitude for maximum coverage," *IEEE Wireless Commun. Lett.*, vol. 3, no. 6, pp. 569–572, 2014.
- [14] T. Cheng *et al.*, "IRS-enabled secure G2A communications for UAV system with aerial eavesdropping," *IEEE Sys. J.*, vol. 17, no. 3, pp. 3670–3681, 2023.
- [15] M. M. Azari, F. Rosas, K.-C. Chen, and S. Pollin, "Ultra reliable uav communication using altitude and cooperation diversity," *IEEE Trans. Commun.*, vol. 66, no. 1, pp. 330–344, 2018.
- [16] SwissTo12, "Rf products space. 3d printing space payloads: antennas and rf components," accessed: 2025-05-19. [Online]. Available: <https://swissto12.com/products/space/>
- [17] R. T. Schwarz, T. Delamotte, K.-U. Storek, and A. Knopp, "Mimo applications for multibeam satellites," *IEEE Trans. Broadcast.*, vol. 65, no. 4, pp. 664–681, 2019.
- [18] J. H. Kim and S. W. Choi, "A deep learning-based approach for radiation pattern synthesis of an array antenna," *IEEE Access*, vol. 8, pp. 226 059–226 063, 2020.
- [19] Cambridge Consultants, "The world's largest commercial airborne antenna for SPL," accessed: 2025-05-19. [Online]. Available: <https://www.cambridgeconsultants.com/project/the-worlds-largest-commercial-airborne-antenna-for-spl/>
- [20] T. M. Lovelley, "Comparative analysis of space-grade processors," Doctoral Dissertation, University of Florida, 2017.
- [21] K. Shen and W. Yu, "Fractional programming for communication systems—part i: Power control and beamforming," *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2616–2630, 2018.
- [22] 3GPP, "3rd generation partnership project; technical specification group radio access network; nr; user equipment (ue) radio transmission and reception; part 5: Satellite access radio frequency (rf) and performance requirements (release 18)," 3GPP, Tech. Rep. TS 38.101-5 V18.7.0 (2024-09), 2024.
- [23] 3GPP, "3rd generation partnership project; technical specification group services and system aspects; release 14 description; summary of rel-14 work items (release 14)," 3GPP, Tech. Rep. TR 21.914 V14.0.0 (2018-05), 2018.
- [24] European Union, "Commission implementing regulation (eu) 2019/947 of 24 may 2019 on the rules and procedures for the operation of unmanned aircraft," 2019, accessed: 2024-11-08. [Online]. Available: [https://eur-lex.europa.eu/eli/reg\\_impl/2019/947/oj](https://eur-lex.europa.eu/eli/reg_impl/2019/947/oj)