Maximizing the latency fairness in UAV-assisted MEC system

Hassaan Hydher¹,² | Dushantha Nalin K. Jayakody¹,²,⁴,⁵ | Stefan Panic³

¹ Centre for Telecommunication Research, School of Engineering, Sri Lanka Technological Campus, Colombo, Sri Lanka
² School of Computer Science and Robotics, National Research Tomsk Polytechnic University, Tomsk, Russia
³ Faculty of Science, University of Pristina, Kosovska Mitrovica, Serbia
⁴ Universidade Autónoma de Lisboa, Portugal
⁵ Autonoma TechLab, Portugal

Correspondence
Stefan Panic, Faculty of Science, University of Pristina, Lole Ribara 29, 38220 Kosovska Mitrovica, Serbia.
Email: stefan.panic@pr.ac.rs

Abstract
Unmanned aerial vehicles (UAVs) assisted edge computing has risen as an assuring technique to accommodate ubiquitous edge computation for resource-limited devices. Thus, this paper proposes an approach to maximize the latency fairness in a UAV-assisted multi-access edge computing (MEC) system. To maximize latency fairness, the authors focus on minimizing the maximum latency experienced among the users. In here, multiple ground users (GUs) offload their tasks to MEC UAV in the absence or unavailability of ground servers due to a disaster or heavy traffic where an iterative algorithm is proposed to minimize the maximum latency among the users subject to minimum control link rate and total power constraints. Sequentially, the UAVs’ 3D location, offloading ratio, GUs’ transmit power and GUs’ computational capacity are optimized. The location of the UAV is optimized by using the novel approach, guided pattern search algorithm while the altitude of the UAV is optimized by analyzing the elevation angle dependant behaviour of the channel gain. A simple approach is utilized for optimizing the offloading ratio of the users by considering the problem as minimizing the point-wise maximum of two convex functions while the bisection method is used to optimize the power allocation. Numerical simulation results illustrate that the proposed approach outperforms other baseline approaches in convergence, minimizing the maximum latency and maximizing and maintaining the fairness among the GUs. Furthermore, it is proved that the guided pattern search algorithm converges at least 3.5 times better while the proposed combined optimization gives 400% fairness gain, in comparison with the baseline approach.

1 | INTRODUCTION

Unmanned aerial vehicles (UAVs) are a promising solution considering on-demand deployment in wireless networks thanks to their low implementation cost, mobility and adaptability. Moreover, the ability to make a strong line-of-sight (LoS) connection plays a prominent role in providing high-standard aerial services. Consequently, several research attempts have been carried out to develop a range of UAV-assisted wireless platforms, such as aerial base stations (ABSs) and relays[1–5]. In [2], a pattern search algorithm is proposed to find the optimal 3D position that maximizes the system profit. A disaster-resilient system is studied in [3] where the intelligent 3D placement of UAV ABS is investigated to maximize the total spectral efficiency of the system. In [4], UAV is deployed as an aerial base station. They have jointly optimized user scheduling, transmit power allocation and UAV trajectory such that they maximize the energy efficiency of the NOMA downlink communication. A UAV-assisted relay communication system is studied in [5] where a UAV is served as an aerial relay to maintain a communication link between a user and a remote destination. The transmission power and the UAV’s trajectory are jointly optimized such that the average outage probability is minimized.

Recently, the rapid development of smart devices has led to the exponential growth in computationally rigorous and low latency service requirements, such as remote surgery, tactile internet, augmented and virtual reality, pedestrian re-identification, mobile crowd-sensing etc. [6, 7, 9, 10]. The aforementioned applications are restricted by resource-constrained edge devices with lower computing capability and limited energy supply. Thus, Offloading computation tasks of such devices to a centralized server can conceivably reduce the computing delay and energy dissipation of edge devices [7, 9]. There are several challenges involved in enabling MEC systems such as...
privacy, prediction [11], time complexity in decision-making [12], burst load evacuation and scheduling [13], virtual and physical resource allocations [14] etc. These problems are studied in the current literature. However, it is possible that a centralized ground server or cloud network may result in network congestion or could experience malfunction [7, 8]. In such an event, UAVs can be deployed as an MEC server to assist the central server or temporarily to do the computation for time critical tasks until the recovery [7, 9]. In conclusion, the main application of UAV-assisted multi-access edge computing (MEC) networks can be categorized as improving user experiences in hot-spot areas, offer stable wireless access and edge computing services for edge devices and enabling computing capability in remote areas [8].

There are several benefits in UAV-assisted MEC networks compared to traditional ground MEC networks. It can cover a larger region thanks to mobility and the ability to create strong LoS propagation. Additionally, low deployment cost makes it feasible to enable MEC services in rural areas. Moreover, less infrastructure requirement makes it suitable for on-demand deployment. However, all these benefits come with ample challenges to overcome, such as intelligent planning for deployment and trajectory, efficient power allocation considering the limited power source and self-organizing network capability to enable on-demand deployment [1, 3].

To overcome these challenges, considerable research efforts have been dedicated [9, 16–21]. In [9], a game theoretic based hierarchical system is proposed to improve the quality of data offloading. Stackelberg game-based model is utilized to correlate the relationship between the UAV and the ground users with simultaneous data gathering through a coalition formation game. Based on the above models, the position of the UAV and the formation of the ground nodes are optimized to increase the reward. In [16], a collaborative computation offloading design is proposed to maximize the long-term utility of the MEC enabled Internet of Things (IoT) network while combating interference reduction. The long-term utility is maximized by optimizing offloading decisions and resource management. In [17], a game-theoretic approach is proposed for effective computing resource allocation among MEC servers and ground users in mobile edge cloud computing networks. Although the MEC enabled system provides the platform for efficient computing, there is always a security concern in the offloading of the task. Thus a physical layer security-enabled system is proposed in [18]. In there, a joint optimization approach is proposed to optimize users’ transmission power, UAV’s position, offloading ratio, UAV computing power and the user association such that it maximizes the minimum secrecy capacity of the user constrained to total power and minimum offloading requirement.

By considering the limited power available in the edge devices, a wireless powered edge computing system is proposed in [19] where sum computation rate maximization is investigated for different computational offloading modes. In the binary offloading mode, it simply takes a binary decision whether to offload the task or not by comparing the best offloading computation and the local computation. In partial offloading mode, sum computation rate is maximized by optimizing processing capacity, user offloading time and the transmit power of the users constrained to energy harvesting causality and UAV’s speed. Another challenge that arises in congested MEC systems is the efficient spectrum utilization as spectrum requirement experiences exponential growth and it is not always feasible to extend the spectrum resources. Therefore, to relieve the spectrum demand in MEC systems, cognitive radio (CR) enabled UAV MEC system is proposed in [15]. Here, the energy efficiency maximization problem is studied by jointly optimizing the sensing time, offloading power and the association. To include the subjectivity of the users’ in the data offloading decision-making, user risk awareness is incorporated in [20] for the decision-making strategy. It is modelled based on the principles of prospect theory. The offloading decision is optimized such that it maximizes users’ overall expected prospect-theoretic utility. However, they have not considered the intelligent power utilization of the available power at the users for computing and data offloading such that it will help to reduce the overall latency and power consumption which will indirectly increase the utility. In [21], the most suitable MEC server identification by the end-user and the optimal way of setting prices by the MEC servers are studied.

There are some drawbacks in the current literature which fails to include some important characteristics of UAV-assisted MEC networks. One of the major advantages of UAVs in communication is that it helps to improve the channel quality by increasing the possibility of LoS propagation. However, it will not always guarantee pure LoS propagation where the contribution of LoS depends on the elevation angle in between the UAV and the user. In current literature, most of the research attempts have assumed pure LoS propagation regardless of the elevation angle [9], [16–20]. To reflect the actual behaviour, we have considered the probabilistic LoS model proposed by the International Telecommunication Union (ITU) which calculates the contribution of the LoS component based on the environmental parameters and the elevation angle. Another key advantage of UAV is the mobility which allows the UAV to adjust its position such that the system performance can be maximized. Therefore, positioning of the UAV plays a major role in UAV-assisted MEC systems which is not considered as a primary fact in the existing literature [9, 16–18, 20]. However, we have considered position optimization such that the expected performance is maximized. Furthermore, the end-user devices are limited in power which requires intelligent power utilization where most of the works assume fixed power allocation for computing and data offloading [9, 16–20]. In contrast, we have optimized the power allocation for computing and data offloading such that it will reduce the overall latency.

Notably, latency is an essential and crucial requirement in MEC systems considering raising interest in time-critical applications. Most of the existing works fail to characterize the latency in the MEC system [9, 16, 17]. There are only a few works that have considered processing latency in UAV-assisted MEC systems [18–20]. Surprisingly, no work can be found in the literature which focuses on maximizing the latency fairness among the users. Therefore, in this paper, we propose an
iterative combined optimization approach to maximize latency fairness among the users. To the best of the authors’ knowledge, this is the first work that proposed an approach to increase the latency fairness among the users in UAV-assisted MEC networks. The main contributions of the paper are as follows:

- 3D placement of MEC UAV is considered to maximize the offloading rate fairness. A guided pattern search is proposed for finding the 2D position and LoS map is utilized for finding the optimal altitude which guarantees a quicker convergence.
- Identified a simple approach of optimizing the offloading ratio of the users by considering the problem as minimizing the point-wise maximum of two convex functions where one is not strictly upper bound or lower bound of the other.
- Total available power of the user is jointly allocated for computation and transmission and it is optimized such that it increases the latency fairness.

2 | SYSTEM MODEL

Figure 1 illustrates the proposed system model. There are ground users (GUs) with limited computing capabilities which are required to process a set of tasks. The non-offloading users are the ones whose processing power is adequate to compute their respective tasks while adhering to the QoS requirements. However, due to the limited computing capacity, some users are unable to completely compute their tasks locally. Those users should offload their task to a ground server (GS) to meet the latency requirements who are labelled as offloading users. As our analysis is focused on offloading users, from here onwards the GUs only refer to the offloading users. There are \( M \) GUs who are randomly located in the area of interest. The set of GU is defined as \( M \). Due to a natural disaster or any other technical malfunction, GS could not process the offloaded tasks. Therefore, a UAV can be deployed as an MEC server to process the offloading tasks until the rescue or reconstruction of GS. Simultaneously, the UAV-assisted MEC system will be controlled by the next nearby GS. As spatial flexibility is an essential requirement in these types of critical applications, rotary-wing UAV is considered for deployment. The GUs are equipped with a single antenna. Considering the multiple access technology, non-orthogonal multiple access (NOMA) is a promising solution that will help to increase the spectral efficiency by allowing multiple users to communicate through a single frequency-time resource block. However, it requires a complex transmitter and receiver processing to cancel out the intra-channel interference. Moreover, each user requires channel state information (CSI) of every other user which will create a huge overhead. In light of these facts, employing NOMA is not suitable for the proposed system as it focuses on a disaster-resilient application. Therefore, we assume all the ground users can offload their tasks to the MEC UAV concurrently through the dedicated channel. Thus the co-channel interference can be neglected.

2.1 | Communication model

The GU’s location is considered as fixed within the frame of analysis and the coordinates of \( m \)-th GU is denoted as \( s_m = [x_m, y_m] \), \( \forall m \in M \). UAVs’ location is denoted as \( s_u = [x_u, y_u] \). The location of the nearby GS is denoted as \( s_g = [x_g, y_g] \). Considering the control channel between UAV and the GS, it will be dominated by LoS due to the altitude of the UAV and the height of the GS tower. Therefore, the channel gain between the UAV and the controlling GS \( (h_{U-GS}) \) is modeled as

\[
b_{U-GS} = \frac{G_1}{H_u^2 + ||s_u - s_g||^2}, \tag{1}
\]

where \( G_1 = g_u g_s \left( \frac{\lambda}{4\pi d_0} \right)^2 \) is the channel gain at the reference distance \( (d_0 = 1 \ m) \). \( g_u \) and \( g_s \) are the antenna gain of the transmitter and the receiver, respectively, and \( \lambda \) denotes the wavelength. The data rate of the control link is given as

\[
R_{U-GS} = B_s \log_2 \left( 1 + \frac{p_G h_{U-GS}}{N_0} \right), \tag{2}
\]

where \( B_s \) is the bandwidth allocated for the control channel, \( p_G \) is the transmit power of the GS and \( N_0 \) is the noise power.

However, the channel from the GU to the UAV will be referred to as the ground-to-air (GTA) channel that could experience LoS propagation or NLoS propagation based on the altitude of the UAV and the environmental parameters. To precisely identify whether the GTA channel will experience LoS or NLoS propagation, we should have accurate and precise information related to the geometrical structure of that particular environment. As we do not have access to such information, we have to think of an effective alternative method to characterize the LoS and the NLoS effect propagation of the channel. There
are no widely accepted matured channel models to characterize the GTA channel. However, if we analyse the current literature, mostly there are four kinds of models that have been used to characterize the GTA channels. First, the geometrical channel models which take into account the exact geometrical structure of the environment and the reflective nature of the wavefront to characterize the channel [24–26]. Second, the channel models that are based on empirical measurements [27–29]. Third, the free-space path loss model with the assumption of pure LoS propagation [16–19, 30–32]. Finally, the channel model based on the probabilistic information of the LoS propagation [7, 8, 23, 33–35].

In our work, the geometrical channel model cannot be used as we do not have the access to the exact geometrical structure of the environment. It is also not possible to model it using empirical models as we do not have access to an empirical measurement database. Modelling the channel using free-space path loss model will not help to study the altitude diversity of the system which is one of the advantages of UAV-assisted systems. Also, in there, the optimal altitude of the UAV always would be the minimal possible altitude as the performance of the channel only depends on the distance. However, this is not the case in the practical scenario where one would be able to gain more advantage by enabling the LoS by adjusting the position and the altitude of the UAV. Moreover, the effect of multipath propagation also varies with the elevation angle of the channel. Although the channel model based on the probabilistic information is not practically accurate, it helps to capture the above-discussed characteristics. Thus, in the proposed system, we focused on a probabilistic model proposed by the ITU [22]. It is noteworthy that our proposed approach also can be accurately applied if we have the access to the exact LoS map of the environment. For tractability, we are adopting the approximated version proposed in [23] where the probability of having LoS propagation between the UAV and the $m^{th}$ GU can be given as

$$P(\text{LoS}, \theta_m) = \frac{1}{1 + a \exp(-b[\theta_m - \alpha])},$$

where $\theta_m$ is the elevation angle between the UAV and the $m^{th}$ GU, $a$ and $b$ are the parameters that depend on the propagation environment. Considering the LoS and NLoS propagation, GTA channel gain from the $m^{th}$ GU to the UAV is modeled as [16]

$$h_{U-GU}^k(m) = \frac{G_k}{\sqrt{H_U^2 + ||s_m - s_G||^2}}^{\alpha_k},$$

where $k \in \{L, N\}$ such that $L$ and $N$ denote the LoS and NLoS propagation, respectively, $G_k$ is the combination of average multipath propagation gain and antenna gains. $\alpha$ is the large-scale path loss exponent, where $\alpha_L < \alpha_N$. The effective channel gain of the off-loading link between the $m^{th}$ GU and the MEC UAV is given as

$$b(m) = P(\text{LoS}, \theta_m)h_{U-GU}^L(m) + P(\text{NLoS}, \theta_m)h_{U-GU}^N(m),$$

and the rate of the off-loading link between the $m^{th}$ GU and the MEC UAV is given as

$$R_{U-GU}(m) = B_s \log_2 \left(1 + \frac{\rho_{m,a}(m)}{N_0} \right),$$

where $B_s$ is the channel bandwidth allocated for the off-loading link and $\rho_{m,a}$ is the transmit power of the $m^{th}$ GU.

### 2.2 Computing model

We consider that the computation task can be partially executed both locally and externally at the same time. The power required for local computing for the $m^{th}$ GU ($P_m$) is given as [18]

$$P_m = \psi(f_m)^3,$$

where $f_m$ is the computing capacity of the $m^{th}$ GU. $\psi$ is the effective capacitation depending on the chip architecture. In local computing, we assume that we have the flexibility to adjust the computing capacity according to the computing power. Therefore, the total power available at the GU can be flexibly divided between offloading transmission and local computing. Therefore, the total power consumption at the $m^{th}$ GU ($P_m$) is given as

$$P_m = P_m + P_m.$$

Time taken to complete one bit of computing locally at the $m^{th}$ GU is given as,

$$T_m = \frac{1 - \eta_m}{f_m},$$

where $\eta_m$ is the task offloading ratio of the $m^{th}$ GU, $\rho$ is the average number of CPU cycles required to compute one bit of the task.

The latency related to the offloaded computation is the combination of data transmission latency, processing latency and downloading latency. However, the downloading latency can be neglected compared to others as the data sent back to the GU is much smaller compared to the data offloaded. Therefore, the downloading latency is neglected. The time taken to offload one bit of task from the $m^{th}$ GU ($T_m(m)$) is given by

$$T_m(m) = \frac{\eta_m}{R_{U-GU}(m)}.$$
where $f_{UAV}$ is the computing capacity allocated to the $m^{th}$ GU. The total time to complete the processing of one bit of offloaded task $T_{u}$ is

$$T_{u}(m) = T_{u}^{e}(m) + T_{u}^{c}(m), \quad (12)$$

As we are considering partial computation offloading mode, effective latency is bottle-necked by higher latency among $T_{u}^{e}$ and $T_{u}^{c}(m)$. Therefore, the effective latency $T_{u}^{e}$ is given by,

$$T_{u}^{e} = \max \{ T_{u}^{e}, T_{u}^{c}(m) \}. \quad (13)$$

### 2.3 Problem formulation

In this work, we aim to maximize the latency fairness among the offloading GUs. Therefore, the objective is to minimize the maximum latency per bit experienced by the GUs. Here, we seek to optimize four essential variables which influence the latency. Those are UAVs 2D location $s_{u}$ and the altitude $H_{u}$, offloading ratio $\eta = \{ \eta_{m}, \forall m \in \mathcal{M} \}$, GU's transmit power $P_{GU} = \{ p_{m}, \forall m \in \mathcal{M} \}$ and GU's computing capacity $P_{GU} = \{ f_{m}, \forall m \in \mathcal{M} \}$. The optimization problem is formulated as,

$$\begin{align*}
\text{minimize} & \quad T_{u}^{e} \quad \text{Max} \quad n \in \mathcal{M} \\
\text{subject to} & \quad p_{u} + p_{m} = P_{u}, \forall m \in \mathcal{M} \quad (14b) \\
R_{U-GS} \geq R_{th} \quad (14c) \\
H_{u} \geq b_{\text{min}} \quad (14d) \\
H_{u} \leq b_{\text{max}} \quad (14e) \\
\eta_{m} \in [0, 1], \forall m \in \mathcal{M}, \quad (14f)
\end{align*}$$

where $R_{th}$ is the minimum rate requirement of the control link, $b_{\text{min}}$ and $b_{\text{max}}$ minimum and maximum altitude allowed, respectively, to operate the UAV. (14b) is the total power constraint of the GU; (14c) is the rate constraint imposed in the control link that should be greater than the given threshold; (14d) is the offloading ratio constraint. Considering the control link between the UAV and the GS, it would have greater strength compared to GTA link as it experiences LoS propagation. Moreover, it will not demand higher data rate compared to the offloading links. Also, the required rate can be achieved by adjusting $P_{GU}$ as GS does not have crucial power constraints. Therefore, we are not considering the (14c) in the proposed optimization problem assuming it can be achieved by adjusting the transmit power of the GS. However, this constraint would be essential in a multi UAV deployment problem.

### 3 PROPOSED ALGORITHM FOR FAIRNESS MAXIMISATION IN UAV ASSISTED MEC SYSTEMS

In this section, we explain our proposed algorithm to minimize the maximum latency among the GUs which will increase the latency fairness between the GUs. The above optimization problem is non-convex since there exist non-linear couplings among the variables $s_{u}, \eta, P_{GU}, P_{GU}$ and the objective function is non-convex with respect to the variables. In order to solve it, the problem is decoupled into three subproblems. Namely, UAV positioning, offloading ratio and power allocation. UAV positioning subproblem is solved using guided pattern search optimization technique. Subsequently, the offloading ratio is optimized through minimizing point-wise maximum of two convex functions. Finally, power allocation for computation and transmission is optimized through bisection optimization.

#### 3.1 UAV positioning subproblem

Initially, the UAV positioning problem is given priority as it is independent of other subproblems. Moreover, it is related to the data offloading rate which is a bottleneck in a MEC system. Therefore, it is essential to maximize the minimum offloading rate among the GUs. Thus the utility of this subproblem is the minimum offloading rate among the GUs where it should be maximized. This problem is approached in two phases. Initially, the 2D position is adjusted through a guided pattern search optimization technique. Next, the altitude of the UAV is adjusted by considering the LoS propagation. The altitude adjusting phase would be highly effective if we have the exact LoS map of the environment. As we do not have the access to such databases, we are adapting the probabilistic LoS map of the environment suggested by ITU. Notably, our approach is also feasible with the exact LoS map of the environment. For a given $P_{GU}$ and $P_{GU}$, the utility of the subproblem can be replaced as channel gain of the respective link instead of offloading rate. Therefore the objective function for this subproblem $S_{0}$ at UAV’s position $s_{u}$ is given by

$$S_{0}(s_{u}) = \min_{m \in \mathcal{M}} b(m), \quad (15)$$

where

$$m_{0} = \arg \min_{m \in \mathcal{M}} b(m). \quad (16)$$

To obtain the 2D location, we propose a guided pattern search (PS) technique. Although PS is simple and easy to implement, the initial location should be wisely selected to achieve lower search time. Therefore, UAV’s initial 2D position is set as the centroid of GUs coordinates. The GU which has the minimum channel gain $m_{0}$ will be referred to as minimum gain user (MGU). From the centroid position, UAV will be moved towards MGU as long as it gives an ascending behaviour for
Subsequently, altitude will be adjusted by \( \Delta \) steps towards the ascending direction of the objective function. Update of the altitude can be given as

\[
H_u^{N} = H_u^{C} + \text{sign} \left[ \frac{\partial h(m)}{\partial m} \right] \Delta ,
\]

where \( H_u^{N} \) and \( H_u^{C} \) are next and current altitude of the UAV, respectively. \( \text{sign}[\cdot] \) will return the sign of the argument. The altitude will be updated as long as it gives ascending behaviour in \( s_0 \). As \( h(m) \) is quasi-concave with respect to \( \theta_m \), convergence altitude is the optimal altitude for the given 2D position. Then it will be followed by guided pattern search in the horizontal direction to find a new MGU. If it finds a new MGU then it will again follow altitude update as per (17). This process will continue until it could not find a new MGU which gives a higher \( s_0 \) than the current MGU. The whole process will run iteratively with the reduction of step size until the convergence. The convergence position will give the maximum \( s_0 \) possible. Thus it is the optimal 3D position of the UAV with the expected precision. The number of iteration for convergence \( \gamma \) will be assigned as per the precision requirement. These ideas are formally stated in Algorithm 1.

### 3.2 Offloading ratio subproblem

For any given UAV position, transmit power, and GUs computing capacity, the offloading ratio optimization is to minimize the point-wise maximum of two convex functions. This optimization can be cascaded into \( M \) parallel optimization each focusing on minimizing the maximum latency of each GU. The optimization subproblem is formulated as

\[
\begin{align*}
\text{minimize} & \quad \max \left\{ T^*_m, T^c_m(m) \right\}, \quad \forall m \in M \\
\text{subject to} & \quad \eta_m \in [0,1], \quad \forall m \in M ,
\end{align*}
\]

where \( T^*_m \) and \( T^c_m(m) \) are convex with respect to \( \eta_m \) where other parameters are fixed. Therefore, our objective is to minimize the point-wise maximum of two convex functions. The point-wise maximum of two convex functions itself is also a convex function. The minimum of such convex function comes at a point where both the functions are equal given one is not strictly the upper bound of the other. If there exist multiple such points, the point which gives the minimum value would be the optimal point. Thus for this subproblem, The optimal offloading ratio of the \( m \)th GU \( \eta^*_m \) is given as

\[
\eta^*_m = \frac{\rho R_{UA\rightarrow GU}(m) f^{UA\rightarrow GU}_m}{f^{UA\rightarrow GU}_m + R_{UA\rightarrow GU}(m)} + \rho R_{UA\rightarrow GU}(m) f^{UA\rightarrow GU}_m .
\]

### 3.3 Power allocation subproblem

In this subproblem, our objective is to allocate the available power of GUs for offloading transmission and local computing such that it minimizes the effective processing latency per bit. For any given \( s_m, \eta \) power allocation subproblem can be formulated as,

\[
\begin{align*}
\text{minimize} & \quad \max \left\{ T^*_m, T^c_m(m) \right\}, \quad \forall m \in M \\
\text{subject to} & \quad p^c_m + p^u_m = P_m, \quad \forall m \in M. 
\end{align*}
\]

Substituting (19) to \( T^*_m \) and \( T^c_m(m) \) will give

\[ T^*_m = T^c_m(m). \]
Thus the point-wise maximum of (20a) can be removed an replaced with $T_m^y$ or $T_m(x)$.

\[
\begin{align*}
\text{minimize} & \quad T_m^y, \quad \forall m \in \mathcal{M} \\
\text{subject to} & \quad \sum_{m \in \mathcal{M}} (\rho_m + \rho_m') = P_m, \quad \forall m \in \mathcal{M}.
\end{align*}
\]

Given the other parameters, both the functions $T_m^y$ and $T_m^z$ are convex with respect to $\rho_m$ and $\rho_m'$, respectively, within the constraint set (22b). Therefore, power allocation subproblem is a convex optimization problem. Thus, it is solved using bisection method.

In UAV positioning algorithm, each user will access one time within a single iteration. Therefore, the time complexity of each iteration of the UAV positioning algorithm is $O(M)$ where the time required for positioning will linearly increase with the number of users. The time complexity involved in the other two subproblems is $O(1)$ where it does not depend on the number of users. If all the algorithms are implemented in a centralized manner, the overall time complexity of the algorithm is $O(M)$. Implementing a centralized server will add a huge overhead to transfer all the relevant information to the centralized server. Therefore, the UAV positioning algorithm will be implemented in the UAV and the other two subproblems will be solved in the GU device. Therefore, the UAV processing will have the time complexity of $O(M)$ and the GU device processing will have the time complexity of $O(1)$. As the UAV has higher computing power than the GU device, it is acceptable to distribute in this manner. At the same time, this setup will immensely reduce the overhead compared to the centralized system. The only information to be shared is the location of the GUs and the UAV.

<table>
<thead>
<tr>
<th>TABLE 1 Simulation settings</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M$</td>
</tr>
<tr>
<td>$h_{\min}$</td>
</tr>
<tr>
<td>$h_{\max}$</td>
</tr>
<tr>
<td>$B_x$</td>
</tr>
<tr>
<td>$P_m$</td>
</tr>
<tr>
<td>$f_{L,x}$</td>
</tr>
<tr>
<td>$a$</td>
</tr>
<tr>
<td>$b$</td>
</tr>
<tr>
<td>$b$</td>
</tr>
<tr>
<td>$b$</td>
</tr>
</tbody>
</table>

4 | SIMULATION RESULTS

In this section, the numerical simulation results are presented to validate our analysis and indicate the effectiveness of our approach. All the experiments are implemented in MATLAB R2017a on a desktop computer with an Intel Core-i7-8700 processor running at 3.19 GHz using 16 GB of RAM, running Windows version 10. We consider that $M = 8$ GUs are randomly and uniformly distributed inside a 300 m $\times$ 300 m square area ($A$). $h_{\min} = 80$ m and $h_{\max} = 500$ m are the minimum and maximum legally allowed altitude to operate the UAV. The average gain values are set as $G_L = 3.5$, $G_N = 1.5$ as the average gain due to multipath is higher for LoS compared to NLoS, the pathloss coefficients are set as $\delta_x = 2.5$, $\gamma_N = 3$ [26] as the NLoS propagation experience quicker loss compared to LoS propagation and the noise power is set to $N_0 = -110$ dBm. Environmental and other related parameters used in the simulation setting are summarized in Table 1.

Figure 2 illustrates the movement of the MEC UAV towards the optimal 3D position which maximizes the minimum offloading rate among the GUs. Offloading rate of the link mainly depends on two parameters. One is the euclidean distance of the link which accounts for the pathloss and the elevation angle of the link which accounts for the probability of LoS. It is observable that the UAV converges to the position which is approximately equidistant from all the GUs. At the same time, the hovering altitude of the UAV is adjusted such that the rate fairness is maintained by manipulating the strength of LoS. The red colour points denote the movement related to the guided pattern search technique in the 2D plane which moves the UAV approximately equidistant from all the GUs and the green colour point indicates the altitude adjustment that is done through (17) which controls the elevation angle to further decrease the objective function (i.e. (15)) to move towards the optimal solution of UAV positioning subproblem.

The convergence speed of our proposed approach Algorithm 1 is compared with the PS approach proposed in [2]. It is illustrated in Figure 3 where it shows the convergence of Algorithm 1 (guided PS) and typical PS algorithm [2] for urban and dense urban environments. Since the UAV positioning algorithm is implemented in the UAV, we assume that the computing power of the UAV is fully utilized for the execution of the algorithm. With that assumption, the execution time is calculated.
FIGURE 3 Max–min offloading rate versus execution time, illustrating the convergence of guided pattern search (proposed approach) and typical pattern search algorithm.

The graph shows that the proposed algorithm quickly converges compared to the typical PS algorithm. This is because in our problem the decent direction of our objective function is predictable with the nature of our objective. The decent direction can be predictable with the current position of the user which has the minimum offloading rate. Guided PS based Algorithm 1 utilizes the predicted knowledge to make a quicker convergence. In contrast, typical PS does not utilize the knowledge of decent direction which results in a slower convergence. It is observable that guided PS based Algorithm 1 gives 3.8 fold and 3.6 fold reduction in the convergence time compared to typical PS for urban and dense urban environments, respectively.

The time complexity of each iteration of guided PS algorithm is $O(M)$. Similarly, the time complexity involved in each iteration of a typical PS algorithm is also $O(M)$. However, guided PS takes less number of iteration compared to typical PS. Thus, guided PS gives quicker convergence in time compared to the typical PS. Furthermore, we can observe that a dense urban environment gives a lower max–min offloading rate compared to the urban environment. This is because it is difficult to make LoS connection in the dense environment; thus it is needed to climb a higher altitude to enable the LoS which will increase the pathloss of the links. Therefore, as it is discussed earlier, the algorithm balances the trade-off between the euclidean distance and the probability of LoS to achieve the optimal solution for the UAV positioning subproblem.

The number of required CPU cycles $^1$ is calculated as per the execution time obtained through our simulation servers. As the UAV position algorithm is implemented in UAV, it fully utilizes its processing capacity of 8 GHz to execute it. Offloading ratio and power allocation subproblems are solved in GU device utilizing its full capacity of 1 MHz. As per that, the respective execution times are calculated. Table 2 shows the execution time of each subproblem. The UAV will take 16.25 $\mu$s to determine the position. Similarly, the GU device will take 10 $\mu$s to decide the offloading ratio and the power allocation. Both the durations can be neglected compared to the processing latency which is in the order of milliseconds.

The offloading ratio and power allocation subproblem’s execution time does not depend on the number of users. However, the UAV positioning sub problem’s execution time depends on the number of users. Figure 4 illustrates the behaviour of execution time of the UAV positioning algorithm along with the number of users. The graph shows that given the threshold of real-time execution, our proposed algorithm is able to serve 4 times more users compared to the baseline approach. For example, if we set the threshold for the real-time execution as 100 $\mu$s, our approach can handle up to 60 users where the baseline approach is only able to handle 14 users in real time.

UAVs are energy-constrained device as it is powered by limited battery power. Therefore, energy consumption is a crucial factor when it comes to UAV related applications. In MEC UAV, energy consumption happens due to three operations, namely mechanical movement, computational processing and data transmission. Considering the data transmission from the

---

$^1$ Number of required CPU cycles can slightly vary with the architecture of the processors.
Our approach
Approach proposed in [2]

FIGURE 5 Energy consumption for mechanical movement in various environment. Our proposed approach is compared with the approach in [2]. We assumed energy required for per meter movement with the given constant speed as 2 J/m

UAV, it will only send back the processed information to the user which is negligible compared to the size of the task. Thus the energy consumption for data transmission can be neglected. As these UAVs are dedicated to edge computing, we assume there is a dedicated power to perform the computational task which can provide the guaranteed computational capacity for a limited time period. However, compared with the overall energy consumption, the energy utilization for mechanical movement plays the dominant role. Thus the other two facts considered are negligible. Therefore, the energy consumption for mechanical movement is compared between our proposed approach and the approach proposed in [2]. It is illustrated in Figure 5.

Figure 5 illustrates that our proposed approach consumes less energy compared to the scheme proposed in [2]. This is because our approach always tends to follow the shortest path by moving towards the minimum gain user. In contrast to that, the approach proposed in [2] tend to move to the best utility around it regardless of the position of the minimum gain user. Therefore, most of the time it results in a longer path.

Figure 6 plots the min–max latency for different sizes of computational task. Our proposed approach is compared with fixed offloading ratio and fixed computing capacity of the GU

FIGURE 6 Min–max latency for different sizes of computational task. Our proposed approach is compared with fixed offloading ratio and fixed computing capacity of the GU

FIGURE 7 Deviation of maximum latency from the average threshold value in various environments. The fairness gain achieved through our proposed approach is indicated

optimization compared to offloading ratio optimization. Moreover, the performance of our approach increases with the size of the task. Therefore, the proposed approach would be more effective in applications which have massive tasks to compute.

Figure 7 illustrates the deviation of maximum latency from the average threshold latency in the proposed approach and fixed offloading ratio based approach for all general propagation environments. Fixed offloading ratio based approach is chosen for the fairness comparison as it performs better than fixed computing capacity based solution. Also, it illustrates the fairness gain achieved through our proposed approach compared to the fixed offloading ratio-based approach. As it outperforms the fixed offloading variable based approach, it will outperform the fixed computing capacity based solutions with a larger margin as Figure 6 illustrates. Also, it is observable that our proposed approach keeps the latency deviation almost constant regardless of the type of environment. This is because it
adopts the offloading ratios and power allocation to maintain fairness among all the GUs.

5 | CONCLUSION

In this work, we have studied the latency performance of UAV enabled MEC systems with randomly distributed ground users and an MEC UAV. Considering the vital role of latency fairness in such systems, we proposed a combined optimization approach to minimize the maximum latency among the GUs by performing adaptive adjustment in UAVs’ 3D position, offloading ratio, GUs’ transmit and computing power. We have proposed guided pattern search optimization approach for 3D positioning of the UAV; an approach to minimize the point-wise maximum of two convex functions is utilized to find the optimal offloading ratio and the bisection method is used to find the optimal transmit and computation power allocation. Numerical results illustrate that our proposed approach outperforms other baseline approaches in convergence, minimizing the maximum latency and maximizing and maintaining fairness among the GUs.

CONFLICT OF INTEREST
The authors declare no conflict of interest.

ACKNOWLEDGEMENTS
This work was funded in part by the framework of Competitiveness Enhancement Program of the Tomsk Polytechnic University, Russia.

DATA AVAILABILITY STATEMENT
Data openly available in a public repository that issues datasets with DOIs.

NOMENCLATURE: Indices and parameters

| \( \mathcal{M} \) | Set of ground users (GU) with the offloading requirement |
| \( s_m \) | 2D location coordinates of the \( m \)th GU |
| \( x_m \) | \( X \) coordinate of the \( m \)th GU |
| \( y_m \) | \( Y \) coordinate of the \( m \)th GU |
| \( s_G \) | 3D location coordinates of the MEC UAV |
| \( x_G \) | \( X \) coordinate of the MEC UAV |
| \( y_G \) | \( Y \) coordinate of the MEC UAV |
| \( h_{\text{U-GS}} \) | channel gain between the UAV and the GS |
| \( d_0 \) | Reference distance |
| \( g_r \) | Transmitter and receiver antenna gain, respectively |
| \( p_{\text{GS}} \) | Transmit power of the ground station |
| \( R \) | Data rate of the control link |
| \( B_r \) | Bandwidth allocated for the control link |
| \( N_0 \) | Noise variance |
| \( P(\text{LOS}, \theta_m) \) | Probability of line of sight for the \( m \)th GU |
| \( \theta_m \) | Elevation angle between the UAV and the \( m \)th GU |
| \( a \) & \( b \) | Environment dependent S-curve parameters |

| \( h^k_{u-GU}(m) \) | Channel gain from the \( m \)th GU to the UAV |
| \( \alpha_k \) | Path loss coefficient |
| \( G_k \) | Average multipath propagation gain and antenna gain |
| \( b(m) \) | Effective channel gain of the offloading link |
| \( R_b \) | Bandwidth allocated for the offloading link |
| \( p^m \) | Transmit power of the \( m \)th GU |
| \( \psi \) | Effective capacitance of the processing chip |
| \( \eta \) | Task offloading ratio of the \( m \)th GU |
| \( T^f \) | Time taken to offload one bit of task from the \( m \)th GU |
| \( T^c \) | Time taken to complete one bit of computing locally at the \( m \)th GU |
| \( P^c \) | Computing capacity allocated to the \( m \)th GU by the UAV |
| \( R_{th} \) | Minimum QoS rate threshold |
| \( b_{\text{min}} \) | Minimum altitude allowed to hover the UAV |
| \( b_{\text{max}} \) | Maximum altitude allowed to hover the UAV |
| \( \rho \) | Average number of CPU cycles requires to compute one bit of the task |

Decision variables

\( s^b \) | 2D ground location coordinates of the MEC UAV |
\( H_u \) | Hovering altitude of the MEC UAV |
\( \eta \) | Offloading ratio of the GUs |
\( P^G \) | Transmit power of the GUs |

ORCID
Hassaan Hydher \( \text{https://orcid.org/0000-0002-1685-9122} \)
Dushantha Nalin K. Jayakody \( \text{https://orcid.org/0000-0002-7004-2930} \)
Stefan Panic \( \text{https://orcid.org/0000-0002-5868-1764} \)

REFERENCES