Joint User Association and Deployment Optimization for Delay-Minimized UAV-Aided MEC Networks

Zihao Han, Ting Zhou, *Member, IEEE*, Tianheng Xu, *Member, IEEE*, and Honglin Hu, *Senior Member, IEEE*

***Abstract*—The development of computation-intensive and delay-sensitive smart applications has put forward great challenge to the current cellular networks. To cope with this challenge, unmanned aerial vehicles (UAVs) providing additional on-demand communication and computing services has become a promising technology. This letter proposes a novel task offloading framework in UAV-aided mobile edge computing (MEC) networks. Specifically, the MEC queuing delay is considered in the formulated problem, where the average task delay is minimized via jointly optimizing user association and UAV deployment. The optimal transport theory is introduced to analyze the user association sub-problem, and the UAV deployment is optimized by the classical particle swarm optimization algorithm. Simulation results show that the delay performance is significantly improved by the proposed algorithm.**

***Index Terms*—UAV, MEC, task delay, user association, UAV deployment, optimal transport theory.**

I. Introduction

THE DEVELOPMENT of 5G technologies enables various applications such as intelligent transportation systems, VR (virtual reality) and AR (augmented reality), which need on-demand communication and computing services to meet the low delay requirement [1]. Unmanned aerial vehicle (UAV)-aided mobile edge computing (MEC), i.e., equipping UAVs with computing servers and communication devices to provide ground users with ubiquitous and flexible services, has attracted increasing attention in recent years [2].

Some works on UAV-aided MEC networks focus on minimizing the overall delay of task completion. In [3], UAVs are deployed to assist users with partial computing tasks to reduce the average delay. In [4], the UAV position, communication and computing resource allocation, and task splitting decisions are optimized to improve total service delay of users. In [5], a learning based algorithm is proposed to allocate multi-dimensional resource in the UAV-aided vehicular MEC networks. Some excellent works have taken queuing delay into consideration in UAV-aided MEC networks [6], [7], [8]. In [6], Zhang et al. give the closed-form solutions of the optimal response delay by using stochastic geometry and queuing theory. In [7], Zhu et al. present a multi-agent deep deterministic policy gradient based algorithm to optimize the UAV trajectory, association and offloading factor in an Internet of Things scenario. In [8], Traspadini et al. characterize the high altitude platforms aided MEC networks as a set of queues and optimize offloading factor to maximize the probability of real-time service. However, previous works mainly consider deploying UAVs separately to provide computing services, which fails to explore a more common heterogeneous network.

Motivated by the above observations, we aim to minimize the average task delay in UAV-aided MEC networks, where the terrestrial base stations (TBSs) and UAVs co-exist and have different capabilities. To make full use of the communication and computing resources of the network, as well as leverage the inherent mobility of UAVs [9], we try to find the optimal user association and UAV deployment. The contributions of this letter are listed as follows:

- Firstly, we consider heterogeneous UAV-aided MEC networks and formulate a problem to minimize the average task delay, where the MEC system is characterized more accurately with queuing theory. The problem is decomposed into two parts, i.e., user association sub-problem and UAV deployment sub-problem.
- Secondly, the user association sub-problem is modeled as a semi-discrete optimal transport problem. We prove the existence of the optimal solution by using optimal transport theory (OTT) and characterize the solution space. The UAV deployment is optimized by using particle swarm optimization (PSO) algorithm.
- Finally, compared with the benchmarks, the proposed algorithm, which jointly optimizes user association and UAV deployment, not only reduces the average task delay by up to 48.44%, but also ensures the fairness of user MEC services. Moreover, the convergence and complexity analysis of our algorithms are provided.

II. System Model and Problem Formulation

As shown in Fig. 1, we consider a geographical area $\mathcal{D} \subset \mathbb{R}^2$, where N users are located following a given distribution $f(x, y)$ over the two-dimensional plane. In addition to K_b TBSs, K_u UAVs equipping with communication devices and

Manuscript received 12 April 2023; revised 25 June 2023; accepted 7 July 2023. Date of publication 12 July 2023; date of current version 9 October 2023. This work was supported in part by the National Key Research and Development Program of China under Grant 2020YFB1806606; in part by the Science and Technology Commission Foundation of Shanghai under Grant 22511100600; and in part by the Young Elite Scientists Sponsorship Program by CIC under Grant 2021QNRC001. The associate editor coordinating the review of this article and approving it for publication was A. Guerra. *(Corresponding author: Honglin Hu.)*

Zihao Han is with the Shanghai Advanced Research Institute, Chinese Academy of Sciences, Shanghai 201210, China, and also with the School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100049, China.

Ting Zhou is with the School of Microelectronics, Shanghai University, Shanghai 200444, China, and also with Shanghai Frontier Innovation Research Institute, Shanghai 201100, China.

Tianheng Xu is with the Shanghai Advanced Research Institute, Chinese Academy of Sciences, Shanghai 201210, China, and also with Shanghai Frontier Innovation Research Institute, Shanghai 201100, China.

Honglin Hu is with the Shanghai Advanced Research Institute, Chinese Academy of Sciences, Shanghai 201210, China (e-mail: hlhu@ieee.org).

Digital Object Identifier 10.1109/LWC.2023.3294749

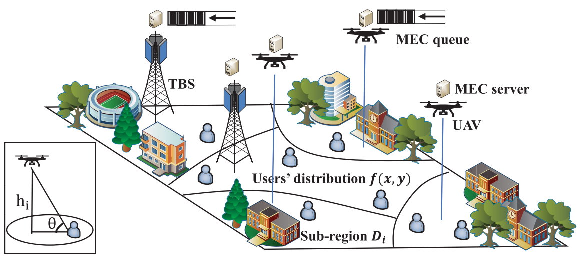


Fig. 1. System model.

computing servers are deployed to enhance the network capacity. Hence, a total of $K = K_b + K_u$ base stations (BSs) can provide computing services for the ground users, forming K disjoint serving sub-regions correspondingly. Assume that the size of each task is M. As the users' computation capability is limited, users' computation tasks are uploaded to the MEC servers equipped on TBSs or UAVs. We denote the three-dimensional (3D) coordinate of the TBS or UAV as $l_k = (x_k, y_k, h_k), k \in \mathcal{K}$ and the corresponding serving region as D_k. In the sub-region $D_k, k \in \mathcal{K}$, users first upload the computation tasks to the associated TBS or UAV, then wait for the MEC server to calculate these tasks, and finally the results are returned to the users.

A. Transmission Model

We consider the UAV-aided MEC networks in the urban environment. Therefore, we adopt probabilistic channel model given in [10], which includes the line-of-sight (LoS) and non-line-of-sight (NLoS) transmission models. The probability of LoS links depends on the elevation angle between the user and UAV, as well as the density and height of the obstacles. Thus, we express the path loss between UAV i and a user located at (x, y) as

$$PL_i^{uav}(x, y) = K_o d_i^2(x, y)\left[P_i^{\mathrm{LoS}}\mu_{\mathrm{LoS}} + P_i^{\mathrm{NLoS}}\mu_{\mathrm{NLoS}}\right], \quad (1)$$

where $K_o = (\frac{4\pi f_c}{c})^2$, c is the speed of light, f_c is the carrier frequency, $d_i(x, y) = \sqrt{(x - x_i)^2 + (y - y_i)^2 + h_i^2}$ is the distance between the user at (x, y) and UAV i, μ_{LoS} and μ_{NLoS} are the shadow fading random variable for LoS link and NLoS link, respectively. The probability of the LoS and NLoS link can be expressed as

$$P_i^{\mathrm{LoS}} = \frac{1}{1 + a\exp(-b(\theta_i - a))}, \quad (2)$$

$$\text{and} \quad P_i^{\mathrm{NLoS}} = 1 - P_i^{\mathrm{LoS}}, \quad (3)$$

where a and b denote the environment constants, $\theta_i = \sin^{-1}(\frac{h_i}{d_i(x,y)})$ is the elevation angle. The traditional path loss model is considered for the TBS-user link [11]. The path loss between TBS j and user at location (x, y) is

$$PL_j^{TBS}(x, y) = K_o d_j^n, \quad (4)$$

where d_j is the distance between TBS and user, and n is the path loss exponent. We denote the path loss between TBS or UAV k and user at location (x, y) as $PL_k(x, y) = \begin{cases} PL_k^{uav}(x, y), & \text{if } k \in \mathcal{K}_u, \\ PL_k^{TBS}(x, y), & \text{if } k \in \mathcal{K}_b. \end{cases}$ We denote the bandwidth as $B_k = \begin{cases} B_u^k, & \text{if } k \in \mathcal{K}_u, \\ B_b^k, & \text{if } k \in \mathcal{K}_b, \end{cases}$ where B_u^k and B_b^k denotes bandwidth of UAV and TBS, respectively. Given the path loss from the TBS or UAV, we can express the rate of a user at (x, y) uploading tasks to a UAV or TBS k as

$$R_k = \frac{B_k}{N_k}\log_2\left(1 + \frac{P_{user}}{PL_k(x, y)\nu^2}\right), \quad (5)$$

where ν^2 is the power of AWGN, p_{user} is the transmit power of the user.

We denote the location of user as $\boldsymbol{e} = (x, y, 0)$, the location of TBS or UAV as $\boldsymbol{s_k}, k \in \mathcal{K}$. The delay of communication is calculated as

$$T_c(\boldsymbol{e}, \boldsymbol{s_k}) = \frac{M}{R_k}. \quad (6)$$

B. MEC Computing Model

Users upload tasks to the associated TBS or UAV for processing, while the computing capability of MEC servers are limited. Each MEC server is assumed to process one task at a time [12], while other tasks waiting in a task queue. The users waiting delay needs to be considered.

We assume that each user generate tasks at rate α following a Poisson process. Thus, all the users' tasks are uploaded, queued and finally processed at the TBSs or UAVs, which are also following a Poisson process of rate λ_k, equal to αN_k. Following first-come-fist-served discipline, tasks are processed by MEC server at a deterministic rate μ_k, equal to $\frac{C_k}{L}$, where L is the computational load for each task [8]. The MEC computation capacity in sub-region D_k is expressed as $C_k = \begin{cases} C_u^k, & \text{if } k \in \mathcal{K}_u, \\ C_b^k, & \text{if } k \in \mathcal{K}_b, \end{cases}$ where C_u^k and C_b^k denote computation capacity of UAV MEC server and TBS MEC server, respectively.

Thus, an M/D/1 queuing model is developed for task queuing at each MEC server. For any $\rho_k = \frac{\lambda_k}{\mu_k} < 1$, the steady-state distribution exists, where the average number of tasks in the queue is given by

$$l_k^q = \frac{1}{2}\left(\frac{\rho_k^2}{1 - \rho_k}\right). \quad (7)$$

The average time for a task generated by user at (x, y), involving the queuing time and processing time, is given by

$$T_p(\boldsymbol{e}, \boldsymbol{s_k}) = \frac{l_k^q}{\lambda_k} + \frac{1}{\mu_k} = \frac{2LC_k - \alpha L^2 N_k}{2C_k^2 - 2\alpha L N_k C_k}. \quad (8)$$

C. Problem Formulation

The task delay in MEC networks mainly consists of three parts, i.e., transmission delay (including task uploading and results downloading), computation delay and waiting delay at MEC. As the computing results are relatively small compared with the task upload, we ignore the task results downloading delay from the TBS or UAV [13]. Thus, we mainly consider delay of task uploading and delay at the MEC server.

Consider $c(\boldsymbol{e}, \boldsymbol{s_k})$ as the delay of task generated by user at $\boldsymbol{e} = (x, y)$ served by TBS or UAV $\boldsymbol{s_k}$, which is expressed as

$$c(\boldsymbol{e}, \boldsymbol{s_k}) = T_c(\boldsymbol{e}, \boldsymbol{s_k}) + T_p(\boldsymbol{e}, \boldsymbol{s_k}). \quad (9)$$

In the UAV-aided MEC networks, we try to minimize the task delay by optimizing user association and UAV deployment. The average task delay in the whole region depends on the region partition $\mathbf{D} = \{D_k, k \in \mathcal{K}\}$ and the UAV location $\mathbf{L} = \{l_k, k \in \mathcal{K}_u\}$, which is formulated as

$$T(\mathbf{D}, \mathbf{L}) = \frac{\sum_{k=1}^{K} \iint_{D_k} c(\boldsymbol{e}, \boldsymbol{s_k}) \cdot f(x, y) dxdy}{N}. \tag{10}$$

The optimization problem can be formulated as

$$\begin{aligned} \mathbf{P1}: \min_{\mathbf{D},\mathbf{L}} \quad & T(\mathbf{D}, \mathbf{L}), && \text{(11a)} \\ \text{s.t.} \quad & \mathcal{D}_m \cap \mathcal{D}_n = \varnothing \quad \forall m, n \in \mathcal{K}, && \text{(11b)} \\ & \bigcup_{k \in \mathcal{K}} \mathcal{D}_k = \mathcal{D}, && \text{(11c)} \\ & x_u^{\min} \le x_k \le x_u^{\max}, \forall k, && \text{(11d)} \\ & y_u^{\min} \le y_k \le y_u^{\max}, \forall k, && \text{(11e)} \\ & h_u^{\min} \le h_k \le h_u^{\max}, \forall k, && \text{(11f)} \end{aligned}$$

where (11b), (11c) constrain that sub-regions do not overlap and cover the entire area, and (11d), (11e), (11f) limit the deployment location of UAVs.

III. Problem Analysis and Algorithm Design

Problem **P1** is an NP-hard problem, with the coupled region partition **D** and UAV location **L**. To solve this problem, we partition problem **P1** into two subproblems, namely user association optimization and UAV deployment optimization. The procedure is as follows.

A. Optimization of User Association

The region partitions $D_k, \forall k \in \mathcal{K}$ are continuous and coupled in the sub-problem **P2**. To solve this sub-problem, we first prove the optimal solution exists via introducing optimal transport theory [14], [15], and then characterize the optimal solution. At last, we propose a low-complexity iterative algorithm to approach the optimal region partition. The user association sub-problem can be rewritten as

$$\begin{aligned} \mathbf{P2}: \min_{\mathbf{D}} \quad & T(\mathbf{D}, \widehat{\mathbf{L}}), && \text{(12a)} \\ \text{s.t.} \quad & \text{(11b)}, \text{(11c)} && \text{(12b)} \end{aligned}$$

As users follow a continuous distribution $f(x, y)$, and TBSs as well as UAVs can be regarded as discrete points, sub-problem **P2**, with fixed UAV location **L**, can be seen as a semi-discrete optimal transport problem. Thus, sub-problem **P2** is equivalent to matching users to the TBSs and UAVs with the minimum average task delay cost.

Theorem 1: Problem **P2** has an optimal solution.

Proof: Let $d_k = \int_{D_k} f(x, y)\ \mathrm{d}x\mathrm{d}y$, for $\forall k \in \mathcal{K}$, $E = \{\mathbf{d} = (d_1, d_2, \ldots, d_K) \in \mathbb{R}^K; d_k \ge 0, \sum_{k=1}^{K} d_k = 1\}$, and $F(\boldsymbol{e}, \boldsymbol{s_k}) = \frac{M}{\log_2(1+\frac{P_{user}}{PL_k(x,y)\ \sigma^2})}$. For any given $\boldsymbol{s}_k$, delay $c(\boldsymbol{e}, \boldsymbol{s}_k)$ is continuous. We have $\liminf_{\boldsymbol{e} \to \boldsymbol{e_0}} c(\boldsymbol{e}, \boldsymbol{s_k}) \ge c(\boldsymbol{e_0}, \boldsymbol{s_k})$, so the delay cost function $c(\boldsymbol{e}, \boldsymbol{s_k})$ is lower semi-continuous. Then, lemma 1 is used from optimal transport theory [16]:

Lemma 1: Consider continuous probability measure f and discrete probability measure ζ in $\mathbf{\Omega}$. Let $L : \mathbf{\Omega} \to \mathbf{\Omega}$ be a transport map from f to ζ and $C(x, T(x)) : \mathbf{\Omega} \times \mathbf{\Omega} \to [0, \infty)$ be the cost function of L. Then, for any semi-continuous cost function, the optimal transport map from f to ζ exists, which minimizes the total transport cost $\int_{\mathbf{\Omega}} C(x, T(x)) f(x) \mathrm{d}x$.

According to lemma 1, the sub-problem **P2** has an optimal solution over E. ■

Theorem 2: To achieve minimum delay in the UAV-aided MEC networks, the optimal region partition is given by

$$\begin{aligned} D_k^* = \{(x, y) :\ & \frac{N_k}{B_k} F(\boldsymbol{e}, \boldsymbol{s_k}) + \frac{2LC_k - \alpha L^2 N_k}{2C_k^2 - 2\alpha L N_k C_k} \\ & \le \frac{N_n}{B_n} F(\boldsymbol{e}, \boldsymbol{s_n}) + \frac{2LC_n - \alpha L^2 N_n}{2C_n^2 - 2\alpha L N_n C_n}, \forall n \ne k \in \mathcal{K}\}, \end{aligned} \tag{13}$$

Proof: According to Theorem 1, optimal region partitions $D_k^*, k \in \mathcal{K}$ exists, which are the solutions to problem (12). Now, we consider another region partition scheme $\widetilde{D}_k, k \in \mathcal{K}$ as an example. Taking a coordinate $z_0 = (x_0, y_0) \in D_m$ and a circle area B_τ with the center z_0 and radius $\tau > 0$. Then, the sub-region $\widetilde{D}_k, k \in \mathcal{K}$ is generated from the optimal partition as

$$\begin{cases} \widetilde{D}_m = D_m \backslash B_\tau(v_0), \\ \widetilde{D}_n = D_n \cup B_\tau(v_0), \\ \widetilde{D}_k = D_k, \quad k \ne m, n. \end{cases} \tag{14}$$

Denote $d_\tau = \iint_{B_\tau} f(x, y)\ dxdy$ and $\widetilde{d}_k = \iint_{\widetilde{\mathcal{D}}_k} f(x, y)\ dxdy$. As the region partition $D_k^*, k \in \mathcal{K}$ is optimal, thus, a better solution cannot be achieved by any variation of the optimal partitions $\widetilde{D}_k, k \in \mathcal{K}$. We have

$$\begin{aligned} & \sum_{k=1}^{K} \int_{D_k} c(\boldsymbol{e}, \boldsymbol{s}_k) \cdot f(x, y) dxdy \\ & \quad \le \sum_{k=1}^{K} \int_{\widetilde{D}_k} c(\boldsymbol{e}, \boldsymbol{s_k}) \cdot f(x, y) dxdy. \end{aligned} \tag{15}$$

Now, we subtract the common items on both sides of the equation, yielding

$$\begin{aligned} & \int_{D_m} c(\boldsymbol{e}, \boldsymbol{s_m}) \cdot f(x, y) dxdy \\ & \quad + \int_{D_n} c(\boldsymbol{e}, \boldsymbol{s_n}) \cdot f(x, y) dxdy \\ & \quad \le \int_{D_m \backslash B_\tau(v_0)} c(\boldsymbol{e}, \boldsymbol{s_m}) \cdot f(x, y) dxdy \\ & \quad + \int_{D_n \cup B_\tau(v_0)} c(\boldsymbol{e}, \boldsymbol{s_n}) \cdot f(x, y) dxdy. \end{aligned} \tag{16}$$

$$\begin{aligned} & \int_{B_\tau(v_0)} c(\boldsymbol{e}, \boldsymbol{s_m}) \cdot f(x, y) dxdy \\ & \quad \le \int_{B_\tau(v_0)} c(\boldsymbol{e}, \boldsymbol{s_n}) \cdot f(x, y) dxdy. \end{aligned} \tag{17}$$

Divide both sides of the inequality by d_τ and take the limit when $\tau \to 0$, we have

$$c(\boldsymbol{v_0}, \boldsymbol{s_m}) \le c(\boldsymbol{v_0}, \boldsymbol{s_n}). \tag{18}$$

Equation (18) shows how we assign a user at (x_0, y_0) to a BS. Consequently, a tractable expression of the optimal region partition is given as Theorem 2. ■

Algorithm 1 Iterative Algorithm for User Association

Input: Number of users N and distribution function $f(x,y)$, the UAV location set $\mathbf{L}$.
Output: The optimal region partition $\mathcal{D}_i^t, \forall i \in \mathcal{K}$.

Let $t = 1$, initialize user association and let $\phi_i^t(x,y) = 0, \forall i \in \mathcal{K}$.
while t $\leq$ Z **do**
  $\gamma \leftarrow 1 - 1/t$
  Compute $\phi_i^{t+1}(x,y)$
  $$= \begin{cases} 1-\gamma\big(1-\phi_i^t(x,y)\big), & \text{if } (x,y) \in \mathcal{D}_i^t. \\ \gamma\phi_i^t(x,y), & \text{otherwise.} \end{cases}$$
  Compute $d_i = \int_{\mathcal{D}} \phi_i^{(t+1)}(x,y)\, f(x,y)\, \mathrm{d}x\mathrm{d}y, \forall i \in \mathcal{K}$
  $t \leftarrow t+1$
  Update user association using (13).
end while
$\mathcal{D}_i^* \leftarrow \mathcal{D}_i^t, \forall i \in \mathcal{K}$.

B. Optimization of UAV Deployment

Sub-problem of UAV deployment optimization is given as **P3**. As $T(\widehat{\mathbf{D}},\mathbf{L})$ is non-convex, we adopt the classical PSO algorithm [17], a widely used non-gradient algorithm, to search the optimal UAV location.

$$\mathbf{P3}: \min_{\mathbf{L}} \quad T(\widehat{\mathbf{D}},\mathbf{L}), \tag{19a}$$
$$\text{s.t.} \quad (11d), (11e), (11f) \tag{19b}$$

The potential UAV locations are represented by particles $X = [l_k, k \in \mathcal{K}_u]$. Each particle adjusts its position X_i and velocity v_i iteratively, guided by its own personal best solution $pbest_i$ and the global best solution *gbest* found by the swarm, which is expressed as

$$X_i^{t+1} = X_i^t + v_i^{t+1}, \tag{20}$$
$$v_i^{t+1} = w * v_i^t + c_p * r_p * (pbest_i^t - X_i^t) + c_g * r_g * (gbest^t - X_i^t), \tag{21}$$

w, c_p, c_g are acceleration coefficients, and r_p, r_g are random numbers between 0 and 1. In the t-th iteration, the fitness function, represented by task delay, updates $pbest_i^t, gbest$ as

$$pbest_i^t = \min_t \Big\{ T\Big(\hat{D}, X_i^t\Big)\Big\}, gbest^t = \min_i \big\{pbest_i^t\big\}. \tag{22}$$

The joint user association and UAV deployment optimization is shown in Algorithm 2.

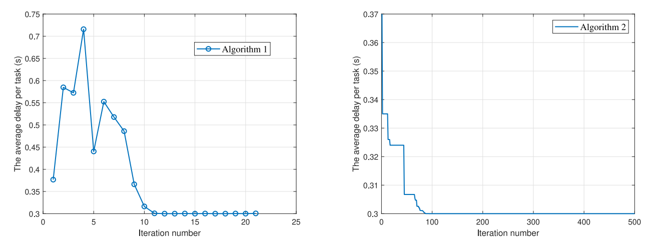
IV. Numerical Results

In our simulations, we assume that two hotspot areas appear at the edge of the cellular network in a region of size 1 km $\times$ 1km, where 3 UAVs are sent to assist the TBS. In this case, we adopt bi-modal truncated Gaussian distribution [14] to simulate user distribution with mean values $\mu_{x_1} = \mu_{y_1} = 330$, $\mu_{x_2} = \mu_{y_2} = 660$ and variance values $\sigma^2 = \sigma_{x_1}^2 = \sigma_{y_1}^2 = \sigma_{x_2}^2 = \sigma_{y_2}^2 = 20000$. The carrier frequency $f_c = 2$ GHz, $c = 3 \times 10^8$ m/s, bandwidth and transmit power of TBS and UAV is 20 MHz, 20 W and 5 MHz, 5 W, respectively. Also, $N = 100$, $M = 1$ Mb, $L = 60$ GFLOP, $C_b = 8000$ GFLOP and $C_u = 5000$ GFLOP [8]. The transmit power of each user

Algorithm 2 Joint User Association and Deployment Optimization for Delay-Minimized UAV-Aided MEC Networks

Input: number of particles S, number of users N and distribution function $f(x, y)$.
Output: UAV location $L^* = Gbest$, the optimal region partition $\mathcal{D}^*$ with L^*.

Randomize the initial position and velocity of particle swarm
while t $\leq$ Z **do**
  for each particle i **do**
    Optimize user association $\mathcal{D}_i^*$ with fixed UAV location X_i^t by **Algorithm** 1.
    Calculate fitness $T\Big(\hat{D}, X_i^t\Big)$.
    Update $pbest_i^t, gbest_i, v_i^t, X_i^t$ by (20), (21), (22).
  end for
  $t \leftarrow t+1$
end while



(a) Algorithm 1 convergence speed (b) Algorithm 2 convergence speed

Fig. 2. Convergence speed of Algorithm 1 and 2.

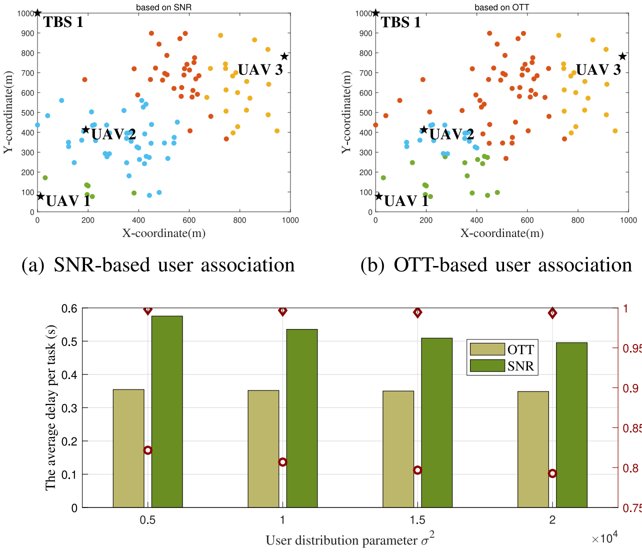
is 2 W, and the noise power spectral density is -170 dBm/Hz. Besides, we consider a dense urban environment with $n = 3$, $\mu_{\text{LoS}} = 3$ dB, $\mu_{\text{NLoS}} = 23$ dB, $a = 8.96$, and $b = 0.04$.

Fig. 2 shows the convergence performance of Algorithm 1 and Algorithm 2. It can be observed that the Algorithm 1 converges within 12 iterations and the Algorithm 2 converges within 100 iterations. The computational complexities for each iteration in Algorithm 1 and Algorithm 2 are given by $\mathcal{O}(KN)$ and $\mathcal{O}(SK_uKN)$, where S represents the number of particles within the swarm.

Then, the proposed Algorithm 2 is compared with the benchmarks as follows. SNR is short for signal-to-noise ratio.

- PSO + SNR: UAVs are deployed by PSO algorithm and users access the BS with the largest SNR.
- Uniform + OTT: UAVs are deployed uniformly and users access the BS by OTT algorithm.
- Uniform + SNR: UAVs are deployed uniformly and users access the BS with the largest SNR.

Fig. 3 shows the superiority of the proposed OTT-based association (Algorithm 1) with the given UAV locations. In the traditional SNR based association, the focus is largely placed on the power strength of the received signal from various BSs. However, this approach overlooks the considerable bandwidth and computing resource variance between the UAVs and the TBS. As shown in Fig. 3(a), this results in an imbalance, with only 30% of users are served by TBS. On the contrary, the proposed OTT-based scheme offers more balanced association with nearly half of the users served by TBS as shown in



(a) SNR-based user association

(b) OTT-based user association

(c) Jain's fairness index for users' task delay

Fig. 3. Comparison of different user association.

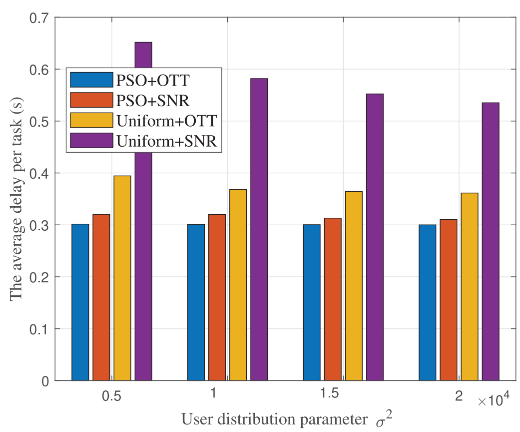


Fig. 4. Average task delay versus user distribution parameter.

Fig. 3(b), which aligns with the available communication and computing resources of TBS and UAVs. In Fig. 3(c), Jain's index is utilized to evaluate the algorithm fairness [14]. The Jain's fairness index of the proposed OTT algorithm is higher than that of the SNR scheme, which shows that the resources of UAVs and TBS are fairly shared by users, and the improvement in the average task delay doesn't come at the expense of some users' services.

Fig. 4 shows the superiority of the Algorithm 2, which jointly optimizes user association and UAV deployment. At $\sigma^2 = 10000$, compared to the uniform+SNR scheme, optimizing user association with the OTT based algorithm can reduce delay by 36.75%, and the proposed Algorithm 2 can reduce delay by 48.44%. The results indicate that the system performance can be improved by individually using OTT based user association algorithm, with any given UAV deployment location. For further enhancement of system performance, it's crucial to jointly optimize both the association of users and the deployment of UAVs.

V. Conclusion

In this letter, we investigated the task delay minimization problem in UAV-aided MEC networks, where TBSs and UAVs have different capabilities and the MEC queuing delay is considered. In particular, we propose an algorithm to jointly optimize user association and UAV deployment. In doing so, PSO algorithm is adopted to find the deployment of multiple UAVs. The fitness of PSO is determined by formulating the optimal association scheme and subsequently computing the corresponding average delay. The existence and characteristics of the optimal user association are obtained by using optimal transport theory, and an iteration algorithm is developed to approach the optimal user association. Numerical results show that the proposed algorithm can reduce the average task delay by up to 48.44%.

References

[1] T. Xu, T. Zhou, J. Tian, J. Sang, and H. Hu, "Intelligent spectrum sensing: When reinforcement learning meets automatic repeat sensing in 5G communications," *IEEE Wireless Commun.*, vol. 27, no. 1, pp. 46–53, Feb. 2020.

[2] C. Wang et al., "Covert communication assisted by UAV-IRS," *IEEE Trans. Commun.*, vol. 71, no. 1, pp. 357–369, Jan. 2023.

[3] Q. Hu, Y. Cai, G. Yu, Z. Qin, M. Zhao, and G. Y. Li, "Joint offloading and trajectory design for UAV-enabled mobile edge computing systems," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 1879–1892, Apr. 2019.

[4] Z. Yu, Y. Gong, S. Gong, and Y. Guo, "Joint task offloading and resource allocation in UAV-enabled mobile edge computing," *IEEE Internet Things J.*, vol. 7, no. 4, pp. 3147–3159, Apr. 2020.

[5] H. Peng and X. Shen, "Multi-agent reinforcement learning based resource management in MEC- and UAV-assisted vehicular networks," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 1, pp. 131–141, Jan. 2021.

[6] Q. Zhang, J. Chen, L. Ji, Z. Feng, Z. Han, and Z. Chen, "Response delay optimization in mobile edge computing enabled UAV swarm," *IEEE Trans. Veh. Technol.*, vol. 69, no. 3, pp. 3280–3295, Mar. 2020.

[7] C. Zhu, G. Zhang, and K. Yang, "Fairness-aware task loss rate minimization for multi-UAV enabled mobile edge computing," *IEEE Wireless Commun. Lett*, vol. 12, no. 1, pp. 94–98, Jan. 2023.

[8] A. Traspadini, M. Giordani, G. Giambene, and M. Zorzi, "Real-time HAP-assisted vehicular edge computing for rural areas," *IEEE Wireless Commun. Lett*, vol. 12, no. 4, pp. 674–678, Apr. 2023.

[9] X. Liu et al., "Placement and power allocation for NOMA-UAV networks," *IEEE Wireless Commun. Lett.*, vol. 8, no. 3, pp. 965–968, Jun. 2019.

[10] 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, Dec. 2014.

[11] H. Min, T. Zhou, T. Xu, Y. Huang, and H. Hu, "Stratified federated learning-based Beamforming design for heterogeneous IRSs-aided networks," *IEEE Wireless Commun. Lett.*, vol. 12, no. 1, pp. 1–5, Jan. 2023.

[12] P. Dai, K. Hu, X. Wu, H. Xing, F. Teng, and Z. Yu, "A probabilistic approach for cooperative computation offloading in MEC-assisted vehicular networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 2, pp. 899–911, Feb. 2022.

[13] C. Sun, W. Ni, and X. Wang, "Joint computation offloading and trajectory planning for UAV-assisted edge computing," *IEEE Trans. Wireless Commun.*, vol. 20, no. 8, pp. 5343–5358, Aug. 2021.

[14] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Wireless communication using unmanned aerial vehicles (UAVs): Optimal transport theory for hover time optimization," *IEEE Trans. Wireless Commun.*, vol. 16, no. 12, pp. 8052–8066, Dec. 2017.

[15] C. Villani, *Topics in Optimal Transportation*. Boston, MA, USA: AMS, 2003.

[16] G. Crippa, C. Jimenez, and A. Pratelli, "Optimum and equilibrium in a transport problem with queue penalization effect," *Adv. Calculus Variat.*, vol. 2, no. 3, pp. 207–246, 2009.

[17] D. Bratton and J. Kennedy, "Defining a standard for particle swarm optimization," in *Proc. IEEE Swarm Intell. Symp.*, Apr. 2007, pp. 120–127.