

Research Article

Performance Optimization in UAV-Assisted Wireless Powered mmWave Networks for Emergency Communications

Jinsong Gui , Nansen Jin , and Xiaoheng Deng

School of Computer Science and Engineering, Central South University, Changsha 410083, China

Correspondence should be addressed to Jinsong Gui; jsgui06@163.com

Received 4 April 2021; Revised 21 May 2021; Accepted 25 May 2021; Published 8 June 2021

Academic Editor: Laurie Cuthbert

Copyright © 2021 Jinsong Gui et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, we explore how a rotary-wing unmanned aerial vehicle (UAV) acts as an aerial millimeter wave (mmWave) base station to provide recharging service and radio access service in a postdisaster area with unknown user distribution. The addressed optimization problem is to find out the optimal path starting and ending at the same recharging point to cover a wider area under limited battery capacity, and it can be transformed to an extended multiarmed bandit (MAB) problem. We propose the two improved path planning algorithms to solve this optimization problem, which can improve the ability to explore the unknown user distribution. Simulation results show that, in terms of the total number of served user equipment (UE), the number of visited grids, the amount of data, the average throughput, and the battery capacity utilization level, one of our algorithms is superior to its corresponding comparison algorithm, while our other algorithm is superior to its corresponding comparison algorithm in terms of the number of visited grids.

1. Introduction

With the development of computers and robotics, various intelligent unmanned vehicles have been developed and applied to various fields. In the ocean, autonomous underwater vehicles (AUVs) have been used as both undersea mobile base stations (BSs) to assist the construction of ocean sensor networks and underwater mobile collectors to facilitate the information collection [1, 2]. In the land, unmanned aerial vehicles (UAVs) play an important role in the similar applications. In particular, UAVs have significant advantages for the rapid construction of wireless emergency communication networks in postdisaster areas.

The disasters caused by earthquake, flood, typhoon, tsunami, and tornado are often unpredictable, and their sudden occurrences will damage the terrestrial radio infrastructures and prevent the victims from communicating with each other or the outside world. Since the 72 hours after the disaster is the most critical, there is a need to deploy wireless networks for emergency communications, which can provide assistance to the rescue workers or teams in the affected areas [3].

Under these emergency communications circumstances, UAVs will be well suited as the platform equipped with small BSs to provide broadband wireless services for postdisaster areas [4, 5], where any rotary-wing UAV will hover over the target area and serve the terrestrial users within a certain horizontal distance. In order to serve more terrestrial users, one effective approach is to increase the number of UAVs for collaborative coverage, which is not always available due to increased networking costs. Another alternative is to fly a UAV along a planned path, covering the target area and serving more terrestrial users.

According to the investigation of the existing works, the main challenge faced by the UAV for the path planning in the emergency communications is that the UAV does not know the distribution of terrestrial users within the affected areas. Therefore, a UAV cannot plan the optimal path in advance to maximize the service efficiency. In addition, due to the limited capacity of UAV batteries [6], it is crucial to find the energy-efficient flight trajectory in UAV-assisted emergency communications [7].

In response to the above challenges, the authors in [8] developed an online learning framework to solve a UAV's

path planning problem for postdisaster emergency communications, in which the UAV serves as a flying BS to serve terrestrial users along its planned path. In [8], the objective of optimization task is to maximize the sum data rate under the capacity constraints of the UAV battery, and it is modeled as an extended multiarmed bandit (MAB) problem [9], where the two path planning schemes were proposed to gradually learn an energy-efficient path while serving terrestrial users within the UAV coverage.

All the above schemes work in the traditional sub-6 GHz frequency bands, which is difficult to meet the demand of high-throughput communication due to the limited spectrum resources [10, 11]. Since the communication duration may be very short when a terrestrial user is in a UAV's coverage each time, there is the increasing demand for ultrahigh-speed data transmission in UAV-assisted emergency communications, which cannot be satisfied by the traditional sub-6 GHz frequency bands because of the limited spectrum resources. The spectrum resources in millimeter wave (mmWave) bands are very rich, which can support ultrahigh-speed data transmission [12]. Therefore, UAV-assisted mmWave communication is becoming one of the most prominent technologies in emergency communication systems.

However, mmWave communications usually rely on beamforming technology to overcome its inherent weakness, such as severe path loss. The beamforming training mechanism is usually used to determine the beam orientation and beam width at both the transmission end and receiving end [13]. Because of the unknown distribution of terrestrial users, it is difficult to determine the number of response slots required for the beamforming training process. If the number of response slots is selected at random, it is very likely that the number of response slots will not match the user distribution density, resulting in either waste of response slots or high probability of conflicts when terrestrial users compete for response slots, which is a new challenge that needs to be addressed in UAV-assisted mmWave networks.

In addition, a user equipment (UE) carried by a terrestrial user trapped in a postdisaster area may not be recharged in a timely manner, so it should save as much battery energy as possible. Especially for an energy-constrained sensing device [14] in a postdisaster area, its residual energy level is the most critical influencing factor for ensuring communication with the outside world. With the popularization of terminals equipped with wireless recharging function, a UE can not only share battery energy each other but also conveniently harvest radiofrequency (RF) energy radiated by BSs [15]. Therefore, the UAV hovering above the target area can also act as a recharger to recharge UEs in its coverage area, and in turn, the recharged UEs can use the harvested energy to send data to the BS in the UAV.

For a UAV's each flight, it is very desirable to both find the most effective path to serve as many UEs as possible and maximize each UE's uplink throughput in UAV-assisted wireless powered mmWave networks for emergency communications. In fact, it is difficult to achieve both goals simultaneously when the distribution of terrestrial users is unknown. Therefore, in this paper, we focus on how to

improve the entire system uplink performance while satisfying the UAV's battery energy constraint and each UE's throughput constraint. We address these challenges in this paper, and our main contributions are as follows:

- (1) We propose the effective performance optimization scheme in UAV-assisted wireless powered mmWave networks for emergency communications, which can improve the entire system uplink performance while satisfying the UAV's battery energy constraint and each UE's throughput constraint even if the distribution of terrestrial users is unknown
- (2) Compared with the scheme in [8], each UE makes full use of the energy it harvests to send uplink data in our scheme and thus extends its battery life due to saving its own battery energy
- (3) In our scheme, the number of response slots can be adjusted according to the currently observed performance results, and thus, it can match the user distribution density as the number of rounds of online learning increases

The rest of this paper is organized as follows. The related works are reviewed in Section 2. The system model and the problem statements are detailed in Section 3. The algorithms for solving the problem are described in Section 4. The performance evaluation results are presented in Section 5, while the conclusions are summarized in Section 6.

2. Related Work

In recent years, the application of UAVs has become more and more extensive, and the researching works in terms of UAV-assisted networks have become a hot spot. In [7], the authors gave an overview of UAV-assisted wireless communications and introduced the basic framework and channel characteristics related to UAV-assisted wireless communications by narrating three cases: UAVs for improving coverage quality, UAVs for information dissemination, and UAVs acting as BSs or relays. Based on the known distribution of terrestrial users, through the joint design of sensing nodes' scheduling scheme, power allocation strategy, and UAV flight trajectory, the scheme in [16] is aimed at minimizing the total power consumption of a UAV, while the work in [17] attempted to minimize the amount of UAVs acting as flying BSs under the condition of satisfying the coverage of ground terminals.

Some works assumed that terrestrial users follow some kind of probability distribution and explored the performance of UAV-assisted networks. The authors in [18] explored how to maximize the coverage probability and system throughput in a single-UAV-assisted downlink network underlaid with device-to-device (D2D) communications, in which the distribution of D2D users follows a homogeneous Poisson point process (PPP). The work of [19] is similar to that in [18], but it considered a multi-UAV-assisted downlink network, in which the UAVs are employed to forward data from satellites to the terrestrial vehicles in a postdisaster

area. The authors in [20] designed a prototype of UAV-assisted emergency Wi-Fi-based network, which is aimed at accelerating search and rescue operation and doing on-site surveillance over a postdisaster area. By using a stochastic geometry framework, they derived analytical approximate expressions for the outage probability. In a postdisaster area, it is not always feasible to know the distribution of terrestrial users. As mentioned above, the authors in [8] developed an online learning framework to improve performance of UAV-assisted wireless communications without the knowledge of the distribution of ground terminals.

Some works focused on connectivity of perceptual networks in a postdisaster area. The authors in [21] explored the deployment of UAV-mounted BSs to assist rescue, where machine-to-machine (M2M) communication is used to establish connectivity and send rescue messages among the human portable/wearable machine-type devices (MTDs) in either relay mode via a UAV or direct mode. For emergency communications based on heterogeneous Internet of Things, the authors in [22] explored how to accommodate the link between the UAV and the terrestrial cellular device and the D2D link outside the UAV's line of sight area in the same frequency band. There are also some open research issues and challenges in UAV-assisted emergency networks, such as intelligent distributed optimization for the trajectory and scheduling for UAVs, and interference management for UAV-assisted emergency networks [23].

All the above works do not involve the application of mmWave communication technology. However, UAV-enabled cellular architecture over mmWave frequency band has been recognized as one of the best solution for on-demand high data rate service. Therefore, the many existing works focused on UAV-assisted mmWave communications [24–38].

The authors of [24] designed a hierarchical beamforming codebook structure with different beam widths to speed up beamforming training and tracking, which is aimed at enhancing the capacity of mmWave UAV-assisted cellular networks. The authors of [25] proposed a training-based beam tracking scheme to keep the connection between UAV and terrestrial BSs. The authors of [26] focused on BS-to-UAV backhaul communications and proposed a beam tracking method, which reduces training overhead by adopting wider beam width at the cost of lower beamforming gain. The authors of [27] proposed a connectivity constraint-based path planning and beam tracking method, by which the UAV can start from a random location and reach its destination within a BS coverage by learning a trajectory while keeping better connectivity.

The works in [28–30] mainly explored the theoretical frameworks to derive the coverage probability and average throughput expression. The authors of [28] assumed that the UAVs acted as the parent nodes and were a three-dimensional (3D) spatial deployed at the same height, while the UEs acted as the daughter nodes and were distributed in a Poisson cluster process (PCP) mode. In [29], the UAVs and BSs were assumed to be distributed in a PPP mode while the UEs were distributed in a PCP mode. In [30], although the UEs are modeled by using a PCP mode, downlink simul-

taneous wireless information and power transfer (SWIPT) scenario and uplink information transmission scenario are jointly considered.

The authors of [31] focused on network coverage and system energy efficiency optimization problem for UAV-assisted mmWave cellular networks, but they did not discuss spatially moveable UAVs acting as flying BSs or dynamic UAV placement. The authors of [32] analyzed the secrecy rate performance of UAV-assisted mmWave networks by utilizing Matern hardcore point process, which can also ensure the safety distance between UAVs. The authors of [33] designed a spectrum management architecture for mmWave swarm networks with UAVs acting as flying BSs, which considered the issues of interference, energy consumption, and UAV mobility. The authors of [34] derived the multicell probability and the volume spectral efficiency for UAV-assisted mmWave cellular networks by using stochastic geometry.

The authors of [35] derived analytical channel models to evaluate the benefit of mmWave links associated with UAVs, which were further assessed by using outage probability. The authors of [36] investigated the rate performance, fairness, and their tradeoff in UAV swarm connected to mmWave cellular networks, where all UAVs are located in a 3D area and distributed according to a homogeneous PPP. The authors of [37] studied beamforming training and tracking for a UAV-assisted mmWave system. They designed a beam training codebook according to the angular domain converted from the known location distribution of users and presented the two beam tracking methods in terms of both random and inertial user mobility to predict the beam direction.

The authors of [34] proposed a 3D location distribution model to characterize the positions of the UAVs by using stochastic geometry and considered multicell coverage probability and volume spectral efficiency as metrics to study the performance of mmWave networks with UAVs acting as aerial BSs. The beam formed by the mmWave antenna array has strong directivity and needs multiple beam scans to cover the whole region, so the authors of [38] proposed a sectoring approach to ensure coverage of the whole region. Since the side lobe gain of the antenna array is considered, it leads to substantial interference in other sectors. They set a threshold on power spillage from adjacent sectors to limit interference in concurrent transmission strategy.

The above solutions for UAV-assisted mmWave cellular networks need to know the probability distribution of terrestrial users or UAVs, which does not always match the reality in the affected area. As far as we know, the work in [8] is most closely related to ours. As mentioned above, the work in [8] cannot be transplanted to a UAV-assisted mmWave emergency communication network, which encourages us to study a new solution for it.

3. System Model and Problem Statement

3.1. System model. As shown in Figure 1, we consider a single rotary-wing UAV acting as a flying BS to serve a postdisaster area without available ground infrastructure. The UAV takes off in the area where it can be easily recharged, and it will

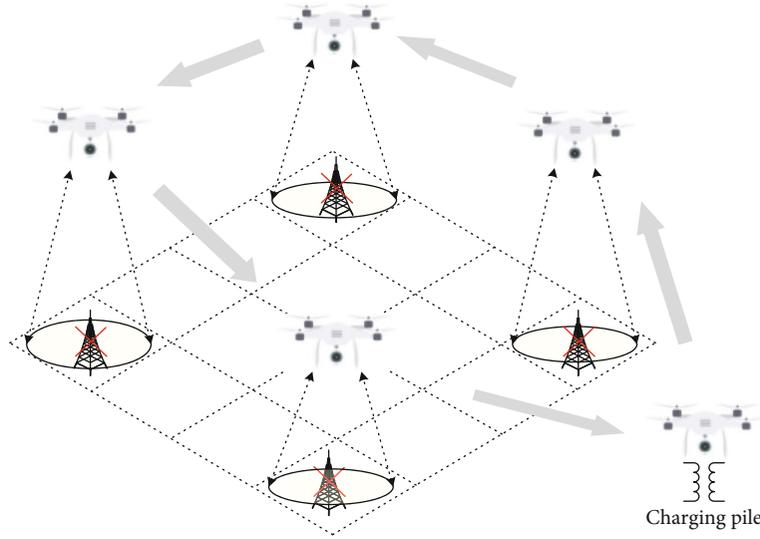


FIGURE 1: UAV-assisted emergency communication, where the hovering UAV is used as an aerial BS covering a certain area.

come back to the recharging point to recharge before the battery runs out. We assume that the UAV flies to a position and then hovers above the position to offer service to ground UEs within horizontal distance d . Therefore, the UAV should alternatively fly and hover to cover the disaster-affected area. We use v_f to denote the flying speed of the UAV and adopt t_h to denote its hovering interval. Also, we use e_{\max} to denote the maximum battery capacity of the UAV, where the UAV's engine powers for hovering and flying are denoted as p_h and p_f respectively.

We divide the considered area into $N \times N$ equal size grids. Also, it is assumed that the UAV can serve all the ground UEs in the grid while hovering over the center of the grid. We use $\mathbb{N} = \{1, 2, \dots, N^2\}$ to denote the set of all the grids. Also, we adopt $l_K = \{g_0, g_1, \dots, g_K\}$ to represent a UAV path, where each element g_k ($g_k \in \mathbb{N}$ and $1 \leq k \leq K$) denotes a grid in this area and g_0 is the recharging point. The path starts at g_0 , serves K grids, and returns to g_0 to recharge. The set of all the possible paths that start and end at the recharging point is denoted as \mathcal{L} . The ground UEs are located randomly in the considered area, and these UE users need wireless access service with the same probability. We use D_{g_k} to denote the data volume of the ground UEs within grid g_k .

The distribution of UEs in each grid is unknown to the UAV. However, when the UAV selects to hover above the center of a grid, it can get the distribution of UEs in this grid through the mmWave beam training process. We assume that the UAV starts a mmWave beam training process at the transmission power p_c . During the mmWave beam training process, the UEs that do not have data delivery requirement can turn on the energy harvesting mode, while those that have data delivery requirement can respond to the training initiation package by feeding back the response packages with location information. After obtaining the location information of all the UEs of the grid, the UAV can adjust its transmission beam width to cover exactly all the UEs of the

grid and then start the downlink recharging process at power p_c . We assume that these UEs only use the harvested energy to send uplink data. The optimization task can be given by

$$\begin{cases} \max_{l_K \in \mathcal{L}} \sum_{k=1}^K D_{g_k}, \\ \text{s.t.} & Kp_h t_h + \sum_{k=1}^K \left(\eta_{g_k} p_c t_h + p_f \frac{d_{g_{k-1}g_k}}{v_f} \right) \leq e_{\max}, \end{cases} \quad (1)$$

where $d_{g_{k-1}g_k}$ represents the horizontal distance between the centers of grid g_{k-1} and grid g_k , and η_{g_k} is the ratio of charging time in the hovering interval t_h in grid g_k .

3.2. Time-Slotted Structure. In general, the hovering interval t_h consists of the beam training interval t_b , the recharging interval t_c , and the transmission interval t_d . As shown in Figure 2, the beamforming training interval t_b consists of one starting slot for training process, multiple response slots, and one confirming slot for training results, where it satisfies the following relation:

$$(I+2)\tau_b = t_b, \quad (2)$$

where τ_b is the size of a slot, which is a predetermined value. The number of response time slots depends on the number of UEs participating in the beamforming training. If the number of UEs is large, the response slots need to be set more accordingly.

The recharging interval t_c and the transmission interval t_d are closely related and bound by a communication cycle. That is, $t_c + t_d = \tau$. In each communication cycle τ , the UAV can recharge UEs during the recharging interval t_c , while it can receive data from UEs during the transmission interval t_d .

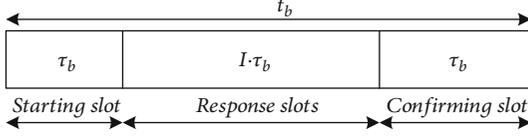


FIGURE 2: The structure of time slots for beamforming training interval t_b .

To allow multiple ground UEs to communicate with the UAV during each communication cycle τ , we need to divide τ into a downward recharging slot (i.e., $\epsilon_0\tau$) and multiple upward transmission slots (i.e., $\epsilon_j\tau$, $1 \leq j \leq J$) based on time division multiple access mode, which is shown in Figure 3. The downward recharging slot is allocated to the UAV to transfer radio energy to its serving UEs, while the upward transmission slots are allocated to UEs for data transmission, during which each UE transmits its own data to the UAV by using the assigned upward transmission slot.

Since $\epsilon_0, \epsilon_1, \dots, \epsilon_j$ are the time ratios in communication cycle τ allocated to the UAV and its serving UEs (e.g., UE₁, UE₂, ..., UE_J) for energy transfer downward and data transfer upward, respectively, which meets the following relation:

$$\sum_{j=0}^J \epsilon_j \leq 1. \quad (3)$$

Therefore, η_{g_k} ($1 \leq k \leq K$) can be estimated by the following formula:

$$\eta_{g_k} = \frac{t_h - I\tau_b - \sum_{j=1}^J \epsilon_j\tau}{t_h}. \quad (4)$$

3.3. Beam Alignment and Signal Propagation Model. To know the distribution of UEs in a grid and align the mmWave beam between each UE and the UAV, the UAV needs to start a beamforming training process between the UAV and the UEs. Firstly, the UAV broadcasts a directional beacon frame in its mmWave beam facing the ground, where all the UEs remain in a quasinidirectional listening mode. Then, each UE that wants to communicate with the UAV chooses a response slot that is similar to the association beamforming training of the UE side (i.e., it transmits the directional beacon frames in all its beams). After receiving the directional beacon frames from UEs, the UAV can obtain the beamforming training information and the location information of every UE. Finally, the UAV broadcasts a confirmation frame to announce the conflict-free UEs. By setting a reasonable number of training slots, we expect to avoid the collision of UEs' beacon frames while keeping training time as short as possible. If there is the collision of UEs' beacon frames, the number of training slots will be increased based on a certain strategy when the beamforming training is needed again. Otherwise, if the number of responding UEs is significantly less than the total number of response slots, it will be

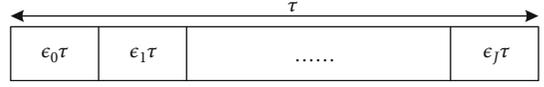


FIGURE 3: The structure of time slots for communication cycle τ .

reduced. We employ the following mmWave signal propagation model:

$$p_{s,u}^r = p_s^t G_{s,u}^t G_{s,u}^r G_{s,u}^c, \quad (5)$$

where p_s^t is the transmission power at the mmWave transmission beam of the UAV s , which is set to p_c ; $p_{s,u}^r$ is the received power at the mmWave link $s \rightarrow u$ from the UAV s to the UE u ; $G_{s,u}^t$ is the directional transmission gain at the mmWave transmission beam of the UAV s , while $G_{s,u}^r$ is the directional reception gain at $s \rightarrow u$; $G_{s,u}^c$ is the channel gain at $s \rightarrow u$. When the mmWave beam between the UAV s and the UE u is aligned, $G_{s,u}^r$ is obtained by the following formula [39]:

$$G_{s,u}^r = \frac{2\pi - (2\pi - \varphi_{s,u}^r)\epsilon}{\varphi_{s,u}^r}, \quad (6)$$

where $\varphi_{s,u}^r$ is the reception beam width of the UE u ; ϵ is the gain of the side lobe and $0 < \epsilon \ll 1$, while $\varphi_{s,u}^r$ is the main lobe in radian. $G_{s,u}^t$ is obtained by the following formula [39]:

$$G_{s,u}^t = \frac{2\pi - (2\pi - \varphi_{s,u}^t)\epsilon}{\varphi_{s,u}^t}, \quad (7)$$

where $\varphi_{s,u}^t$ is the transmission beam width of the UAV s , and it is the main lobe in radian. The channel gain $G_{s,u}^c$ is obtained by the following formula [40]:

$$G_{s,u}^c = |\chi_{s,u}^c \delta(\tau - \tau_{s,u})|^2, \quad (8)$$

where $\delta(\cdot)$ is the Dirac delta function; $\tau_{s,u}$ and $\chi_{s,u}^c$ are the propagation delay and the amplitude of $s \rightarrow u$, respectively. $\tau_{s,u}$ is obtained by the following formula:

$$\tau_{s,u} = \frac{d_{s,u}}{c}, \quad (9)$$

where $d_{s,u}$ is the distance of $s \rightarrow u$, and c is the speed of light. When $s \rightarrow u$ is a line-of-sight (LOS) link, the amplitude is obtained by the following formula [40]:

$$\chi_{s,u}^{c, \text{los}} = \frac{\lambda}{4\pi d_{s,u}}, \quad (10)$$

where λ is the wavelength and $\lambda = c/f_c$, f_c is the carrier frequency. if $s \rightarrow u$ is a non-line-of-sight (NLOS) link, it will be eliminated during the beam training process, so we do not discuss its estimation method. $d_{s,u}$ is estimated by the

following formula:

$$d_{s,u} = \sqrt{\begin{matrix} (x_s - x_u)^2 \\ + (y_s - y_u)^2, \\ + (z_s - z_u)^2 \end{matrix}}, \quad (11)$$

where (x_s, y_s, z_s) is the three-dimensional coordinate of the UAV s , while (x_u, y_u, z_u) is that of the UE u .

3.4. Throughput Maximization for UEs in Single Grid. After obtaining the number (e.g., J) of conflict-free UEs participating in the beamforming training process, the UAV (e.g., s) wirelessly powers this set of UEs at the power p_c , where the amount of energy harvested by each UE (e.g., u) in the downward recharging slot ϵ_0 is obtained by the following formula:

$$e_{s,u} = \eta p_c G_{s,u}^t G_{s,u}^r G_{s,u}^c \epsilon_0 \tau, \quad (12)$$

where $e_{s,u}$ is the amount of energy harvested by the UE u from the UAV s , while p_c denotes the recharging power of the UAV s . After the UE u supplement its energy during the downward phase, in the subsequent upward phase, it transfers its own data to the UAV s in its allocated time slot ϵ_u ($1 \leq u \leq J$). Because of the recharging efficiency problem, not all of the energy harvested by each UE can be adopted for its data transmission in the upward. We use κ_u to denote the power conversion efficiency factor for the UE u , $0 < \kappa_u \leq 1$. Thus, the average transmission power according to the harvested energy for the UE u is given by

$$p_{s,u}^t = \frac{\kappa_u e_{s,u}}{\epsilon_u \tau}. \quad (13)$$

To simplify the formulaic expression of problem modeling, we assume $\kappa_u = 1, \forall u$. Although changing κ_u can affect the range of values, it has no effect on the overall trend of the simulation results described in the following text. Based on (12)~(13) and the Shannon theorem, the achievable upward throughput from the UE u to the UAV s in bits/second (bps) can be given by

$$r_{s,u}(\epsilon) = \epsilon_u H \log_2 \left(1 + \vartheta_{s,u} \frac{\epsilon_0}{\epsilon_u} \right), \quad (14)$$

where $\epsilon = [\epsilon_0, \epsilon_1, \dots, \epsilon_u, \dots, \epsilon_J]$, H is the mmWave bandwidth, and $\vartheta_{s,u}$ is expressed as follows:

$$\vartheta_{s,u} = \frac{\eta p_c G_{s,u}^t G_{s,u}^r G_{s,u}^c G_{u,s}^t G_{u,s}^r G_{u,s}^c}{\sigma_s^2}, \quad (15)$$

where σ_s^2 is the ambient noise at the UAV s , and $G_{u,s}^t$, $G_{u,s}^r$, and $G_{u,s}^c$ are the directional transmission gain, the directional reception gain, and the channel gain at the mmWave link $u \rightarrow s$ from the UE u to the UAV s , respectively. From (14), it is seen that $r_{s,u}(\epsilon)$ decreases with ϵ_0 for a given ϵ_u . Moreover, it is observed that $r_{s,u}(\epsilon)$ increases with ϵ_u for a

given ϵ_0 . However, due to the total time constraint in (2), ϵ_0 and ϵ_u cannot be increased simultaneously. Therefore, there exists an optimal time allocation (i.e., ϵ_0^*) for maximizing the throughput. When ϵ_0 is less than ϵ_0^* , the throughput gets larger with ϵ_0 ; otherwise, it gets smaller with ϵ_0 .

For J UEs, if we do not limit their lowest throughput, the ‘‘doubly near-far’’ phenomenon explained in [41] will occur while optimizing their total throughput, which is generated by the downward and upward distance-dependent signal propagation attenuation. Furthermore, the result generated from the common-throughput maximization scheme in [41] deviates far from the optimal value. According to (14), the total throughput of all the J UEs is given by

$$R_{g_k}(\epsilon) = \sum_{u=1}^J r_{s,u}(\epsilon), \quad (16)$$

where $R_{g_k}(\epsilon)$ is the total throughput of all the UEs of the grid g_k served by the UAV s and it is also a function of the time allocation vector ϵ . Thus, the total throughput maximum problem based on time slot threshold constraint is modeled as follows:

$$\begin{cases} \max_{\epsilon} R_{g_k}(\epsilon), \\ \text{s.t.} & \sum_{u=0}^J \epsilon_u \leq 1, \\ \epsilon_u \geq \epsilon_{th}, \forall u \in [1, J], \end{cases} \quad (17)$$

Lemma 1. $r_{s,u}(\epsilon)$ is a concave function of ϵ for any given $u \in \{1, 2, \dots, J\}$.

Proof. Please refer to the proof of Lemma 3.1 in [41]. \square

Based on Lemma 1, $R_{g_k}(\epsilon)$ is also a concave function of ϵ because it is the sum of $r_{s,u}(\epsilon)$. Thus, the optimization problem (17) is convex, which can be solved by using convex optimization methods. Lemma 3.2 in [41] can be used to solve the optimization problem (17), which is repeated as follows. If A is more than 0, there is a unique $z^* > 1$ that is the solution of $f(z) = A$, in which $f(z) \triangleq z \cdot \ln z - z + 1$ and $z \geq 0$. Thus, $f(z)$ is convex function with respect to $z > 0$, in which the minimum value is get at $z = 1$ with $f(1) = 0$. If A is more than 0 and not more than 1, there are the two different solutions for $f(z) = A$. If A is more than 1, there is only one solution for $f(z) = A$, which is more than 1, i.e., $z^* > 1$.

Proposition 2. The suboptimal time allocation solution for the optimization problem (17), which is denoted by $\epsilon^* = [\epsilon_0^*, \epsilon_1^*, \dots, \epsilon_u^*, \dots, \epsilon_J^*]$, is given by

$$\epsilon_u^* = \begin{cases} \frac{z^* - 1}{A + z^* - 1}, & u = 0, \\ \frac{\vartheta_{i,j}}{A + z^* - 1}, & 1 \leq u \leq J, \end{cases} \quad (18)$$

where $A_s \triangleq \sum_{u=1}^J \vartheta_{s,u} > 0$ and $z_s^* > 1$ is the corresponding solution of $f(z_s) = A_s$.

Proof. The Lagrangian of the optimization problem (17) is given by

$$\mathcal{L}_s(\epsilon, \varsigma, \zeta) = R_{g_s}(\epsilon) - \varsigma \left(\sum_{u=0}^J \epsilon_u - 1 \right) - \sum_{u=1}^J \zeta_u (\epsilon_{th} - \epsilon_u), \quad (19)$$

where $\varsigma \geq 0$ and $\zeta_u > 0$ ($u \in [1, J]$) are the Lagrange multipliers associated with the constraints in the optimization problem (17). Thus, the dual function of the optimization problem (17) is given by

$$\Theta(\varsigma, \zeta) = \min_{\epsilon \in \Omega} \mathcal{L}_s(\epsilon, \varsigma, \zeta) \quad (20)$$

where Ω is the feasible set of ς and ζ specified by the constraints in the optimization problem (17). \square

It is seen from the optimization problem (17) that there is a $\epsilon \in \Omega$ with $\epsilon_u > \epsilon_{th}$ ($u \in [0, J]$), meeting $\sum_{u=0}^J \epsilon_u < 1$. Therefore, the optimization problem (17) has a strong duality due to the Slater's condition [42] and thus is a convex optimization problem, where the Karush-Kuhn-Tucker (KKT) conditions are both necessary and sufficient for the global optimality of the optimization problem (17), which are given by

$$\sum_{u=0}^J \epsilon_u^* \leq 1, \quad (21)$$

$$\varsigma^* \left(\sum_{u=0}^J \epsilon_u^* - 1 \right) = 0, \quad (22)$$

$$\epsilon_u^* \geq \epsilon_{th}, \quad u \in [0, J], \quad (23)$$

$$\zeta_u^* (\epsilon_{th} - \epsilon_u^*) = 0, \quad u \in [0, J], \quad (24)$$

$$\frac{\partial R_{g_s}(\epsilon^*)}{\partial \epsilon_u} - \varsigma^* + \zeta_u^* = 0, \quad u \in [0, J], \quad (25)$$

where ϵ_u^* , ς^* , and ζ_u^* represent the optimal and dual solutions of the optimization problem (17), respectively. It is easily verified that $\sum_{u=0}^J \epsilon_u^* = 1$ must hold for the optimization problem (17) and thus from (22) without loss of generality, we assume $\varsigma^* > 0$. Also, it is easily verified that $\epsilon_u^* > \epsilon_{th}$ must hold for the optimization problem (17) and thus from (24) without loss of generality, we set $\zeta_u^* = 0, u \in [0, J]$. From (25), it follows that

$$\sum_{u=1}^J \frac{\vartheta_{s,u}}{1 + \vartheta_{s,u}(\epsilon_0^*/\epsilon_u^*)} = \frac{(\varsigma^* - \zeta_0^*) \ln 2}{H}, \quad (26)$$

$$t \left(\vartheta_{s,u} \frac{\epsilon_0^*}{\epsilon_u^*} \right) = \frac{(\varsigma^* - \zeta_u^*) \ln 2}{H}, \quad 1 \leq u \leq J, \quad (27)$$

where $t(x)$ is defined as follows:

$$t(x) \triangleq \ln(1+x) - \frac{x}{1+x}, \quad x \geq 0. \quad (28)$$

Given $1 \leq u, u' \leq J$, from (26), we have

$$t \left(\vartheta_{s,u} \frac{\epsilon_0^*}{\epsilon_u^*} \right) + \zeta_u^* \ln 2 = t \left(\vartheta_{s,u'} \frac{\epsilon_0^*}{\epsilon_{u'}^*} \right) + \zeta_{u'}^* \ln 2, \quad u \neq u'. \quad (29)$$

It is shown that $t(x)$ is a monotonically increasing function with respect to $x \geq 0$ because $dt(x)/dx = x(1+x)^{-2} \geq 0$ for $x \geq 0$. Thus, the equality in (29) holds if and only if $\vartheta_{s,u}(\epsilon_0^*/\epsilon_u^*) = \vartheta_{s,u'}(\epsilon_0^*/\epsilon_{u'}^*), 1 \leq u, u' \leq J$, i.e.,

$$\frac{\vartheta_{s,1}}{\epsilon_1^*} = \frac{\vartheta_{s,2}}{\epsilon_2^*} = \dots = \frac{\vartheta_{s,J}}{\epsilon_J^*} = C. \quad (30)$$

It is worth noting that $1 - \epsilon_0^* = \sum_{u'=1}^J \epsilon_{u'}^*$ and $\epsilon_{u'}^* = (\vartheta_{s,u'}/\vartheta_{s,u}) \epsilon_u^*$ from (21) and (30), respectively. Thus, ϵ_u^* is given by

$$\epsilon_u^* = \frac{1 - \epsilon_0^*}{\sum_{u'=1}^J \vartheta_{s,u'}} \vartheta_{s,u} = \frac{1 - \epsilon_0^*}{A} \vartheta_{s,u}, \quad (31)$$

where $A = \sum_{u'=1}^J \vartheta_{s,u'}$. Furthermore, it follows from (27)~(31) that

$$\ln(1 + C\epsilon_0^*) - \frac{C\epsilon_0^*}{1 + C\epsilon_0^*} = \frac{A}{1 + C\epsilon_0^*}, \quad 1 \leq u \leq J, \quad (32)$$

where C is defined in (30). Because $C = A/(1 - \epsilon_0^*)$ from (31), we can rewrite (32) as

$$z \ln z - z - A + 1 = 0, \quad (33)$$

where $z = 1 + (A\epsilon_0^*/(1 - \epsilon_0^*))1 - \epsilon_0^*$. It is seen that z is more than 1 if A is more than 0 and $0 < \epsilon_0^* < 1$. On the basis of Lemma 3.2 in [41], there is a unique $z^* > 1$ that is the solution of (33). Thus, the optimal time allocation to the downward radio energy transfer is estimated as follows:

$$\epsilon_0^* = \frac{z^* - 1}{A + z^* - 1}. \quad (34)$$

Furthermore, on the basis of (31) and (34), the optimal time allocation to the upward radio information transmission is estimated as follows:

$$\epsilon_u^* = \frac{\vartheta_{s,u}}{A + z^* - 1}, \quad 1 \leq u \leq J. \quad (35)$$

Therefore, it proves Proposition 2.

In order to make the suboptimal time allocation solution (i.e., the formula (18)) for the optimal problem (17) meet the minimum throughput requirement of the UE with the

longest communication distance, the following relationships must be satisfied:

$$\frac{H}{L} \geq \rho_{\text{th}}, \quad (36)$$

where the meanings of H and L are shown in Figure 4, and ρ_{th} may be selected from 0.7 to 0.9 based on the actual experience. Therefore, the following relationships must be satisfied accordingly:

$$\phi \leq \cot^{-1} \rho_{\text{th}}. \quad (37)$$

From Figure 4, it can be known that the ground-oriented mmWave beam width of the UAV should not exceed $2 \cot^{-1} \rho_{\text{th}}$. After the beam training interval t_b , the UAV gets the coordinates of all the conflict-free UEs. So, it can also estimate the distance from itself to each UE according to formula (11). If the longest distance is less than L , it can further narrow the beam width facing the ground, which makes the emitted energy be more concentrated and also further inhibits the adverse effects of the ‘‘doubly near-far’’ phenomenon.

3.5. UAV Flight Path Planning for Throughput Optimization. For optimization task (1), the throughput optimization problem for the UEs in a single grid (e.g., the grid g_k) has been solved in Section 3.4, but it does not involve how to choose an optimal flight path for the UAV to maximize the throughput it receives from the UEs of all the grids in this path. In this subsection, we will address it.

The UAV is unaware of the distribution of UEs in a grid and their communication requirements until a beamforming training process is performed on a selected grid. Due to the limited battery capacity of the UAV, only a subset of all the grids can be selected to form its flight path during a flight. The selection problem of an optimized subset can be transformed into a MAB problem, where there exists a bandit with multiple arms denoted by $\mathbb{N} = \{1, 2, \dots, N^2\}$ and each arm offers a reward (e.g., R_{g_k} , $g_k \in \mathbb{N}$). In this paper, the UAV is the gambler and each grid is its arm.

Based on the MAB theory, in order to maximize the total reward of all the trials, the gambler needs to explore different arms to find the most value one and also exploit the arm with most value as many times as possible so as to maximize the total reward. It is desirable to find a balance between exploration and exploitation.

Unlike the real gambler who can select each arm arbitrarily, the UAV need to fly to different grids to serve UEs. So, it needs to consider not only the battery energy consumption during a flight but also the number of potential UEs in the target grid. The former depends on the horizontal distance between its current grid and the target grid, while the latter requires a reasonable forecast of the number of UEs in the target grid. Since the classic algorithms for the MAB problem do not consider the number of potential UEs in the target grid, we extend them to solve our formulated UAV path planning task. For the distance-aware upper con-

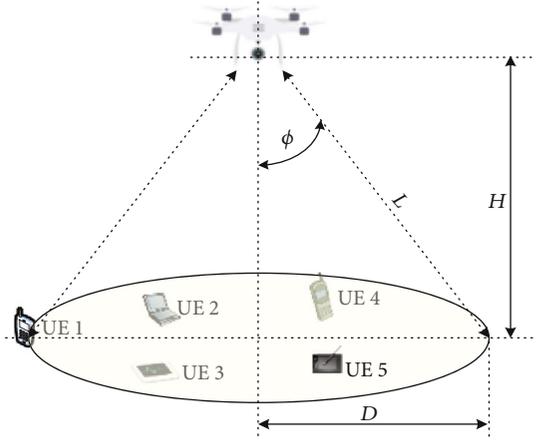


FIGURE 4: An example of the ground-oriented mmWave beam width of the UAV and the coverage area.

fidence bound (D-UCB) algorithm in [8], the UAV first serves each grid once, and then, it selects a grid from $\mathbb{N} = \{1, 2, \dots, N^2\}$ according to the following formula:

$$n_{D-UCB}^* = \operatorname{argmax}_{n \in \mathbb{N}} \left\{ \begin{array}{l} \bar{r}_{n,t} + \alpha \sqrt{\frac{2 \ln t}{T_{n,t}}} - \\ \beta d_{n_{t-1},n} - \gamma \frac{d_{n,t}}{B_r} \end{array} \right\}, \quad (38)$$

where t represents the index of serving rounds and the UAV serves a grid in each round, $\bar{r}_{n,t}$ is the average reward of grid n , $T_{n,t}$ is the times that the UAV visited grid n , $d_{n_{t-1},n}$ represents the horizontal distance between the next grid and the current one, $d_{n,t}$ represents the horizontal distance between the next grid and the recharging point, B_r represents the remaining battery capacity of the UAV, and α , β , and γ are the relevant empirical parameters.

In (38), if $\bar{r}_{n,t}$ is large, the UAV tends to exploit this high-profit grid to get the possible maximal throughput; if $T_{n,t}$ is large, the confidence interval $\sqrt{2 \ln t / T_{n,t}}$ decreases, and thus, the UAV tends to explore other less selected grids. In addition, $\beta d_{n_{t-1},n}$ represents the flight cost and it prompts the UAV to select the nearby grids to save the flight energy consumption, while $\gamma(d_{n,t}/B_r)$ represents the remaining battery level of the UAV and it prompts the UAV to select the grids near the recharging point to return to the recharging point with less energy consumption.

Obviously, formula (38) does not take into account the communication requirement of the UEs who have conflicts during the current round of beamforming training process. Ignoring such potential UEs will result in unfair distribution of access opportunities to the UAV. In addition, in order to avoid the response frame conflict of beamforming training process, the number of response slots will be increased, which in turn may make the number of response slots greatly exceed the actual number of responding UEs due to the dynamic distribution of UEs. For any grid which occurs in

```

Run at the UAV s
Input:  $e_{\max}, t_h, p_h, p_c, p_f, v_f$ 
Output:  $\sum_{k=1}^K D_{g_k}$ 
1:  $t = 0; e_r = e_{\max}; D_{\text{sum}} = 0; g_k = g_0$ 
2: For each grid  $g_{k+1} \in \mathbb{N}$  except for the current grid  $g_k$  do
3:   Initialize  $I_{g_{k+1}}$  to an empirical value
4:    $T_{g_{k+1}} = 0; M_{g_{k+1}} = 0; J_{g_{k+1}} = 0$ 
5:   Invoke Algorithm2 to get  $D_{g_{k+1}}$  and update  $M_{g_{k+1}}, J_{g_{k+1}}, \eta_{g_{k+1}}$ 
6:    $e_r = e_r - p_h t_h - \eta_{g_{k+1}} p_c t_h - (p_f/v_f) d_{g_k, g_{k+1}}$ 
7:    $t = t + 1; T_{g_{k+1}} = T_{g_{k+1}} + 1; D_{\text{sum}} = D_{\text{sum}} + D_{g_{k+1}}; g_k = g_{k+1}$ 
8: End for
9:  $\mathbb{N}_r = \emptyset$ 
10: For each grid  $g_{k+1} \in \mathbb{N}$  except for the current grid  $g_k$  do
11:    $e = p_h t_h + \eta_{g_{k+1}} p_c t_h + (p_f/v_f) (d_{g_k, g_{k+1}} + d_{g_0, g_{k+1}})$ 
12:   If  $e_r > e$  then add  $g_{k+1}$  to  $\mathbb{N}_r$  End if
13: End for
14: If  $\mathbb{N}_r \neq \emptyset$  then
15:   Select a grid  $g_{k+1}$  from  $\mathbb{N}_r$  according to formula (39)
16:   If  $M_{g_{k+1}} > 0$  then
17:      $I_{g_{k+1}} = I_{g_{k+1}} + M_{g_{k+1}}$ 
18:   Else
19:      $I_{g_{k+1}} = I_{g_{k+1}} - \lfloor (I_{g_{k+1}} - J_{g_{k+1}}) / 2 \rfloor$ 
20:   End if
21:   Invoke Algorithm2 to get  $D_{g_{k+1}}$  and update  $M_{g_{k+1}}, J_{g_{k+1}}, \eta_{g_{k+1}}$ 
22:    $e_r = e_r - p_h t_h - \eta_{g_{k+1}} p_c t_h - (p_f/v_f) d_{g_k, g_{k+1}}$ 
23:    $t = t + 1; T_{g_{k+1}} = T_{g_{k+1}} + 1; D_{\text{sum}} = D_{\text{sum}} + D_{g_{k+1}}; g_k = g_{k+1}$ 
24:   Go to 9
25: End if
26: Return  $D_{\text{sum}}$ 

```

ALGORITHM 1: FPP-IDA.

this case, if the number of response slots is reduced, the corresponding communication cycle τ will be increased. That is, this grid has the potential for greater throughput. Therefore, we propose the improved grid selection model based on distance-aware as follows:

$$g_{\text{ID-UCB}}^* = \operatorname{argmax}_{g_k \in \mathbb{N}} \left\{ \begin{array}{l} \widehat{R}_{g_{k+1}}^{\min \{t_0, t\}} + \alpha \sqrt{\frac{2 \ln t}{T_{g_{k+1}}}} - \\ \beta d_{g_k, g_{k+1}} - \gamma \frac{d_{g_0, g_{k+1}}}{e_r} + \\ \chi_D M_{g_{k+1}} + \psi_D \frac{I_{g_{k+1}} - J_{g_{k+1}}}{I_{g_{k+1}}} \end{array} \right\}, \quad (39)$$

where $\widehat{R}_{g_{k+1}}^{\min \{t_0, t\}}$ is the average throughput of grid g_{k+1} during the $\min \{t_0, t\}$ latest rounds; t_0 is fixed to a positive integer; $T_{g_{k+1}}$ is the cumulative times that the UAV visited grid g_k ; $d_{g_k, g_{k+1}}$ represents the horizontal distance between the next grid g_{k+1} and the current grid g_k ; $d_{g_0, g_{k+1}}$ represents the horizontal distance between the next grid g_{k+1} and the recharging point g_0 ; e_r represents the remaining battery

energy of the UAV; χ_D and ψ_D are the relevant empirical parameters; $M_{g_{k+1}}$ is the number of response slots with conflicting signal in the next grid g_{k+1} , which records the information about the last conflict that occurred in this grid; $J_{g_{k+1}}$ is the latest number of conflict-free UEs in the next grid g_{k+1} ; and $I_{g_{k+1}}$ is the latest number of total response slots in the next grid g_{k+1} . $\widehat{R}_{g_{k+1}}^{\min \{t_0, t\}}$ is given by

$$\widehat{R}_{g_{k+1}}^{\min \{t_0, t\}} = \begin{cases} \frac{\sum_{i=t-t_0+1}^t R_{g_{k+1}}(i)}{t_0}, & \text{if } t > t_0, \\ \frac{\sum_{i=1}^t R_{g_{k+1}}(i)}{t}, & \text{if } t \leq t_0, \end{cases} \quad (40)$$

where $R_{g_{k+1}}(i)$ is the throughput of grid g_{k+1} during the i th round.

Based on formulas (39) and (40), we design the flight path planning algorithm based on improved distance-aware (FPP-IDA) for a fully recharged UAV to find the most effective path to serve as many UEs as possible. The FPP-IDA algorithm requires that a UAV must first visit each grid once in order to know the rewards from all the grids. After that, it can select out the next grid with the greatest potential reward

```

Run at the UAV  $s$ 
Input:  $g_k, e_r, \tau_b, t_h, p_h, p_c, p_f, v_f, I_{g_k}$ 
Output:  $R_{g_k}, M_{g_k}, J_{g_k}, \eta_{g_k}$ 
1: Initialize  $L_r$  to 0
2: Broadcasts a directional beacon frame in the mmWave beam facing the ground during the starting slot
3: Waiting for response frames during the response slots
4: Broadcasts a confirmation frame to announce the conflict-free UEs during the confirming slot
5: Observe the number of conflict-free UEs
6: Update  $J_{g_k}$  with the number of the current conflict-free UEs
7: Observe the number of response slots with conflicting signal
8: Update  $M_{g_k}$  with the number of conflicting response slots
9:  $\tau = t_h - (I_{g_k} + 2)\tau_b$ 
10:  $\eta_{g_k} = t_h - I_{g_k}\tau_b - \sum_{j=1}^{J_{g_k}} e_j\tau/t_h$ 
11: For each UE  $u \in J_{g_k}$  do
12:   Calculate  $d_{s,u}$  according to formula (11)
13:   If  $d_{s,u} > L_r$  then  $L_r = d_{s,u}$  End If
14: End for
15: If  $L_r < L$  then
16:   Update the UAV's beam width according to  $\phi_r = \cot^{-1}(H/L_r)$ 
17: End If
18: Get  $e^* = [e_0^*, e_1^*, \dots, e_{J_{g_k}}^*]$  by invoking Algorithms 1–2 in [43]
19: Calculate  $R_{g_k}(e^*)$  according to formula (16)
20:  $D_{g_k} = (1 - e_0^*)\tau R_{g_k}$ 
21: Return  $D_{g_k}, M_{g_k}, J_{g_k}, \eta_{g_k}$ 

```

ALGORITHM 2: Beam alignment and first-charging-then-receiving.

according to the historical average reward of each grid, the number of the served times of each grid, and the other relevant parameters. When the FPP-IDA algorithm judges that the remaining energy after the UAV's visiting the next grid is not enough to make the UAV fly back to the recharging point, it instructs that the UAV stops visiting the next grid and flies back to the recharging point. The further details for the FPP-IDA algorithm are described in Section 4.

For the ∂ -exploration algorithm in [8], the UAV selects a grid with the probability ∂ or selects the grid that generate maximum throughput with the probability $1-\partial$. After checking the remaining battery capacity at each round t , with probability ∂ , the UAV selects a grid from $\mathbb{N} = \{1, 2, \dots, N^2\}$ according to a softmax function which converts the average throughput of the grid n into the probability.

$$P(n) = \frac{e^{\bar{r}_{n,t}/Q}}{\sum_{n'=1}^{N^2} e^{\bar{r}_{n',t}/Q}}, \quad (41)$$

where Q is more than 0 and it is the temperature parameter of the softmax function. If this variable Q goes to infinity (i.e., high temperature), all the grids will be selected by nearly the same probability. Otherwise, if this variable Q goes to 0 (i.e., low temperature), the probability of choosing the grid with the highest average throughput tends to 1. Also, with probability $1-\partial$, the UAV selects a grid from $\mathbb{N} = \{1, 2, \dots,$

$N^2\}$ according to the following formula:

$$n_{\partial-e}^* = \operatorname{argmax}_{n \in \mathbb{N}} \left\{ \bar{r}_{n,t} - \theta d_{n_{t-1},n} \right\} \quad (42)$$

Since the ∂ -exploration algorithm tends to choose the nearby grids according to formula (42), we propose the improved grid selection model based on ∂ -exploration as follows:

$$g_{1-\partial-e}^* = \operatorname{argmax}_{g_k \in \mathbb{N}} \left\{ \widehat{R}_{g_{k+1}}^{\min\{t_0,t\}} - \theta d_{g_k,g_{k+1}} + \chi_{\partial} M_{g_{k+1}} + \psi_{\partial} \frac{I_{g_{k+1}} - J_{g_{k+1}}}{I_{g_{k+1}}} \right\}, \quad (43)$$

where χ_{∂} and ψ_{∂} are the relevant empirical parameters. In addition, we also rewrite formula (41) as follows:

$$P(g_k) = \frac{e^{R_{g_k}(t)/Q}}{\sum_{g'_k \in \mathbb{N}} e^{R_{g'_k}(t)/Q}}. \quad (44)$$

Based on formulas (40), (43), and (44), we design the flight path planning algorithm based on improved ∂ -exploration (FPP-I ∂ E) for a fully recharged UAV to find the most effective path to serve as many UEs as possible. Unlike the FPP-IDA algorithm, the UAV running the FPP-I ∂ E algorithm does not serve each grid in advance once since it is based on probability ∂ to balance exploration and exploitation, where the larger the ∂ is, the more inclined it is to

```

Run at the UAV s
Input:  $e_{\max}, t_h, p_h, p_c, p_f, v_f, \partial, \varrho$ 
Output:  $\sum_{k=1}^K D_{g_k}$ 
1:  $t = 0; e_r = e_{\max}; D_{\text{sum}} = 0; g_k = g_0$ 
2: For each grid  $g_{k+1} \in \mathbb{N}$  except for the current grid  $g_k$  do
3:   Initialize  $I_{g_{k+1}}$  and  $\eta_{g_{k+1}}$  to an empirical value
4:    $R_{g_{k+1}} = 0; T_{g_{k+1}} = 0; M_{g_{k+1}} = 0; J_{g_{k+1}} = 0;$ 
5: End for
6:  $\mathbb{N}_r = \emptyset$ 
7: For each grid  $g_{k+1} \in \mathbb{N}$  except for the current grid  $g_k$  do
8:    $e = p_h t_h + \eta_{g_{k+1}} p_c t_h + (p_f/v_f)(d_{g_k, g_{k+1}} + d_{g_0, g_{k+1}})$ 
9:   If  $e_r > e$  then add  $g_{k+1}$  to  $\mathbb{N}_r$  End If
10: End if
11: If  $\mathbb{N}_r \neq \emptyset$  then
12:   Generate a random number  $\varepsilon$  from 0 to 1
13:   If  $\varepsilon > \partial$  then
14:     Select a grid  $g_{k+1}$  from  $\mathbb{N}_r$  according to formula (43)
15:     If  $M_{g_{k+1}} > 0$  then
16:        $I_{g_{k+1}} = I_{g_{k+1}} + M_{g_{k+1}}$ 
17:     Else
18:        $I_{g_{k+1}} = I_{g_{k+1}} - \lfloor I_{g_{k+1}} - J_{g_{k+1}}/2 \rfloor$ 
19:     End if
20:     Invoke Algorithm2 to get  $D_{g_{k+1}}$  and update  $M_{g_{k+1}}, J_{g_{k+1}}, \eta_{g_{k+1}}$ 
21:      $e_r = e_r - p_h t_h - \eta_{g_{k+1}} p_c t_h - (p_f/v_f) d_{g_k, g_{k+1}}$ 
22:      $t = t + 1; T_{g_{k+1}} = T_{g_{k+1}} + 1; D_{\text{sum}} = D_{\text{sum}} + D_{g_{k+1}}$ 
23:      $g_k = g_{k+1}$ 
24:     Go to 6
25:   Else
26:     Select a grid  $g_{k+1}$  from  $\mathbb{N}_r$  according to formula (44)
27:     Go to 15
28:   End if
29: End if
30: Return  $D_{\text{sum}}$ 

```

ALGORITHM 3: FPP-I ∂ E.

exploitation. The further details for the FPP-I ∂ E algorithm are described in Section 4.

4. Problem Solving Algorithm Description

The pseudocode for FPP-IDA is listed in Algorithm 1. In Algorithm 1, after initializing some variables (see line 1), each grid must be visited once (see lines 2~8), where parameter variables associated with each grid are initialized (see lines 3~4), and then, Algorithm 2 is invoked to get the throughput for each grid and update the corresponding parameters (see line 5). After the execution of Algorithm 2, the remaining energy of the UAV is updated (see line 6), the number of times that each grid is visited is recorded, and the amount of data received by the UAV is accumulated (see line 7).

After each grid is serviced once, it is necessary to determine the next most appropriate grid to service, where the grids that meet the conditions are firstly identified (see lines 9~13), and then the most suitable one is selected (see lines 14~25). In the identification phase, if the amount of energy

that the UAV takes to serve a grid and then return to a recharging point is less than its current remaining energy, this grid can be served by the UAV and thus added to the candidate set \mathbb{N}_r (see lines 11~12). In the selection and service phase, the algorithm can adjust the serving parameters of the currently selected grid based on its historical data (see lines 16~20), where the number of response slots should be increased accordingly if this grid has any conflicting historical record (see lines 16~17); otherwise, the number of response slots should be reduced appropriately if there are too many idle response slots (see lines 18~20).

Lines 2~8 of Algorithm 2 mainly complete the beam training task and observe and record the number of conflicting response slots and the information of nonconflicting UEs. According to the known position information of nonconflicting UEs, Algorithm 2 adjusts the beam width of the UAV (see lines 11~17), and invokes Algorithms 1~2 in [43] to get $e^* = [e_0^*, e_1^*, \dots, e_{J_{g_k}}^*]$ (see line 18). From this, the data collected by the UAV from the grid can be calculated (see lines 19~20).

TABLE 1: Simulation parameters.

Symbol	Description	Value
α	Empirical parameter related to visited grids	1
β	Empirical parameter related to distance between grids (FPP-IDA)	2
γ	Empirical parameter related to distance between current grid and charging point	1
χ_D	Empirical parameter related to number of slots in conflict (FPP-IDA)	2
ψ_D	Empirical parameter related to number of response slots in idle state (FPP-IDA)	1
θ	Empirical parameter related to distance between grids (FPP-I ∂ E)	0.5
χ_{∂}	Empirical parameter related to number of slots in conflict (FPP-I ∂ E)	2
ψ_{∂}	Empirical parameter related to number of response slots in idle state (FPP-I ∂ E)	1
∂	Exploration probability	0.2
τ_b	Basic slot size	0.1 s
d	The UAV's covering radius	50 m
H	The UAV's altitude above ground	100 m
ϕ	Half value of beam width	27°
v_f	Flying speed	20 km/h
P_f	Flying engine power	2 w
P_h	Hovering engine power	4 w
P_c	Recharging power	1 w
t_h	Hovering interval	120 s
Q	Temperature parameter	1
$N \times N$	The number of grids	9×9
t_0	The positive integer	2

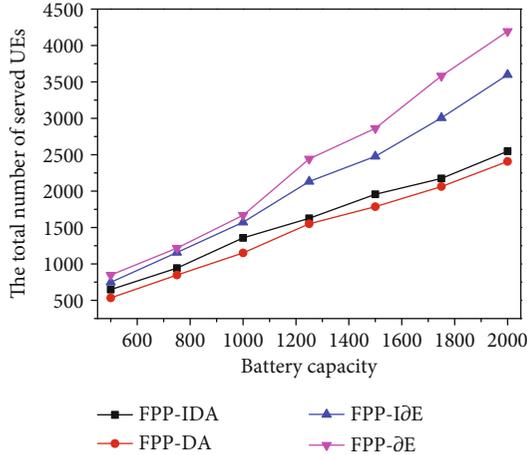


FIGURE 5: The total number of served UEs as a function of battery capacity.

Based on Algorithm 1 and Algorithm 2, we know that the FPP-IDA algorithm needs K rounds of calculation for a UAV flight path $l_K = \{g_0, g_1, \dots, g_K\}$, where $N^2 - 1$ grids need to be calculated and compared to select the next service grid in each round, and thus, the computational complexity of the FPP-IDA algorithm is $O(K \times N^2)$. In order to get the optimal solution, we must enumerate each possible solution in the

solution space to determine which is optimal, and thus, the computational complexity of obtaining the optimal solution is $O(N^{2K})$. Clearly, the computational complexity of the FPP-IDA algorithm is significantly less than that of the optimal solution mathematically.

The pseudo-code for FPP-I ∂ E is listed in Algorithm 3, which is basically similar to Algorithm 1, except that it does not require that each grid must be served once at the beginning of algorithm execution process and its grid selection strategy (see lines 13~14 and lines 25~26) is different from that of Algorithm 1.

5. Performance Evaluation

5.1. Experimental Parameter Settings. We study the performance of our two algorithms (i.e., FPP-IDA and FPP-I ∂ E) through simulation in terms of the total number of served UEs, the number of visited grids, the amount of data, the average throughput, and the battery capacity utilization level. The battery capacity utilization level is defined as the ratio of the amount of energy used for wireless recharging and hovering to the initial amount of energy of the UAV, which is used to measure the percentage of energy to serve UEs. When it gets bigger, it means that the UAV is more about serving UEs instead of flight.

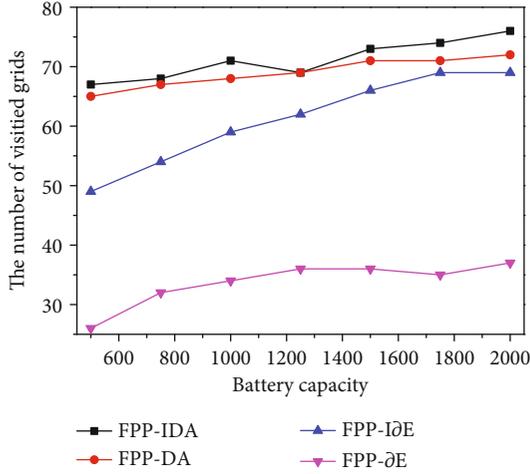


FIGURE 6: The total number of visited grids as a function of battery capacity.

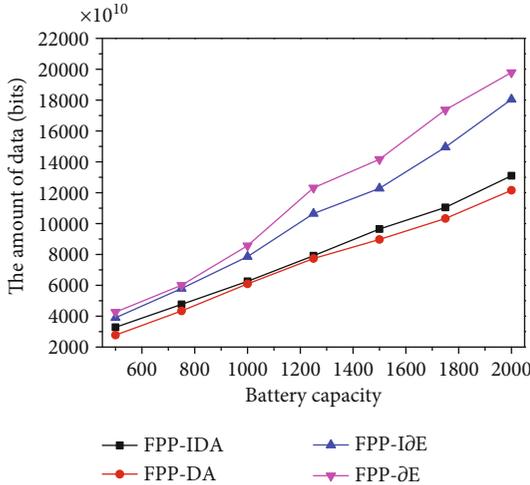


FIGURE 7: The amount of data as a function of battery capacity.

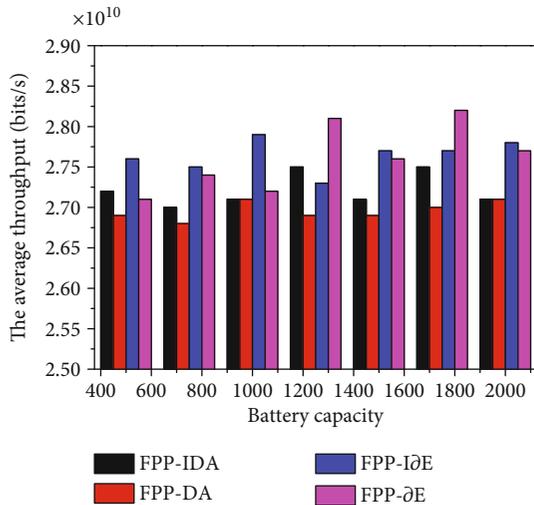


FIGURE 8: The average throughput as a function of battery capacity.

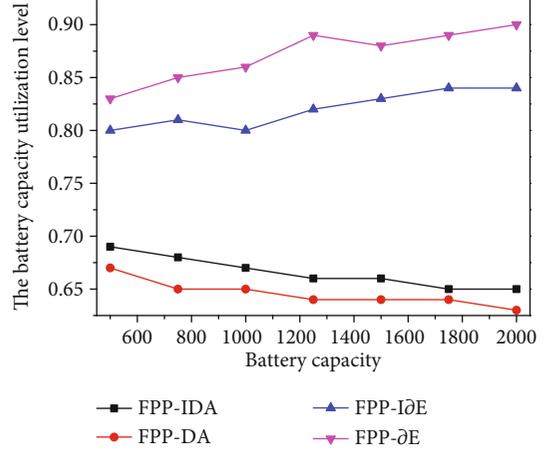


FIGURE 9: The battery capacity utilization level as a function of battery capacity.

To compare the strengths and weaknesses of our two grid selection models, we consider the flight path planning algorithm based on distance-aware (FPP-DA) and the flight path planning algorithm based on ∂ -exploration (FPP- ∂ E) as the comparison algorithms. FPP-DA selects the next grid by formula (38) instead of formula (39), and the others are the same as those of FPP-IDA. Similarly, FPP- ∂ E selects the next grid by formulas (41) and (42) instead of formulas (43) and (44), and the others are the same as those of FPP-I Δ E.

The UEs are randomly distributed and the traffic load in any grid follows binomial distribution $B(u_k, P)$, where u_k is the number of UEs in grid g_k , P is the possibility that the UE needs radio access service, and the possibility that each UE needs radio access service is fixed to 0.2. In addition, each UE will transfer the given amount of data when it wants to get access service, where the length of the sent data is determined randomly. If any UE gets an access service in a round but it does not complete all the data transfers in this round, it will compete for access service in each subsequent encountered beamforming training opportunity.

The key parameters of the proposed algorithms and the comparison algorithms are selected via a series of numerical simulations, and these parameters and the other main parameters are listed in Table 1.

5.2. Experimental Results and Analysis. We first consider the five simulations, which are shown in Figures 5–9. Here, the number of UEs is fixed to 1000, and the distance between the recharging point and the center of the covered area is fixed to 900 m.

We simulate the performance trend of the four algorithms as the initial amount of energy (i.e., the battery capacity) varies. From Figures 5–7, we can see that all the four algorithms can serve more UEs, visit more grids, and get more amount of data as the battery capacity increases. This is because the longer paths can be planned when the UAV has more energy.

Figure 8 shows that there is not a significant relationship interaction between the throughput and the battery capacity, which is consistent with the intuition. Still, there is a small

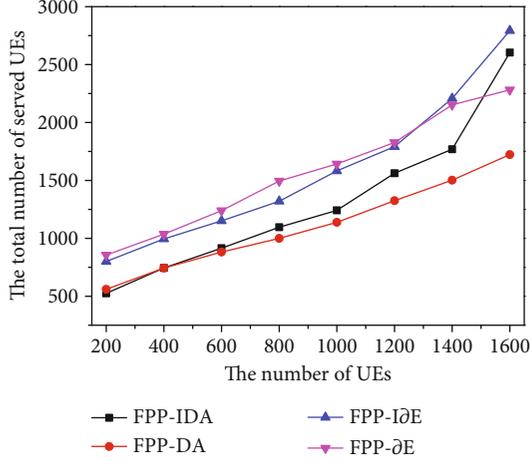


FIGURE 10: The total number of served UEs as a function of UEs.

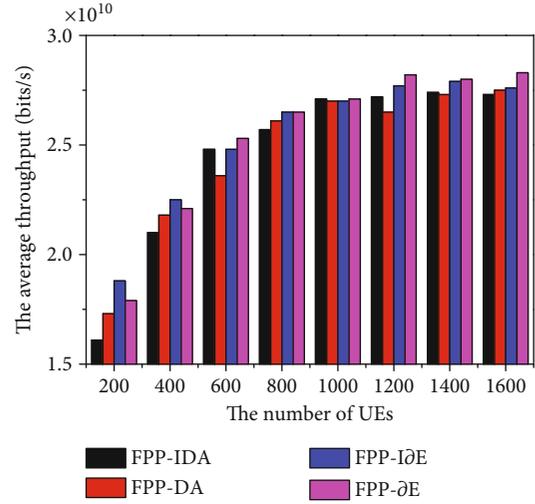


FIGURE 13: The average throughput as a function of UEs.

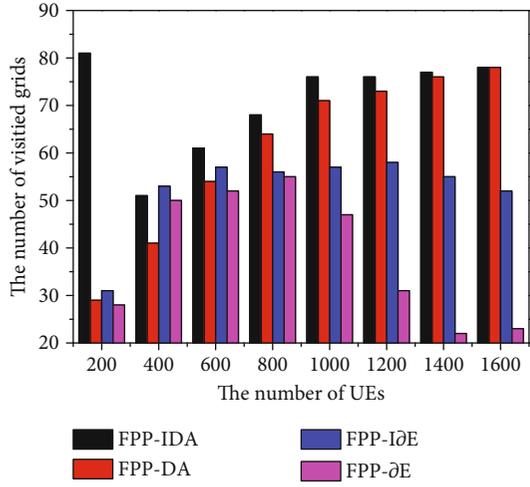


FIGURE 11: The total number of visited grids as a function of UEs.

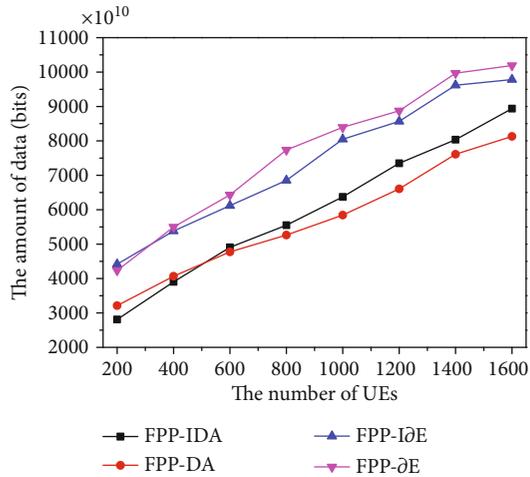


FIGURE 12: The amount of data as a function of UEs.

variation trend. When the battery capacity gets large from a small value, the UAV has more opportunities to find the grids with higher throughput. However, when it continues to get large, this advantage is no longer obvious or even vanishes since it may not find the more grids with higher throughput.

From Figure 9, the battery capacity utilization level in FPP-IDA and FPP-DA decreases with the battery capacity, while that of FPP-I∂E and FPP-∂E increases with the battery capacity. In FPP-IDA and FPP-DA, once a grid is visited, the probability of its being visited again is reduced. Therefore, the UAV needs to fly more different grids to serve UEs, and thus, it increases its flying distance and consumes more flight energy. In FPP-I∂E and FPP-∂E, once the UAV has spotted a few nearby high-throughput grids, it tends to repeat the aircraft over the neighboring area since the exploration probability ∂ is very small, which is helpful to reduce the flying distance and flight energy.

Figures 5–9 show that FPP-IDA outperforms FPP-DA on the five performance metrics. This is mainly because we have taken into account the avoidance of conflicts in the beamforming training phase and the dynamic adjustment of response slots. From Figures 5–9, we also see that these considered factors have little effect on FPP-I∂E. As mentioned above, FPP-I∂E tends to make the UAV repeat the aircraft over the neighboring area once the UAV has spotted a few nearby high-throughput grids. So the effect of these considered factors is also compromised when the UAV always serve the UEs in those familiar grids. When compared with FPP-∂E, FPP-I∂E allows the UAV to explore more unknown grids and thus extend the service range, though it has worse performance in terms of the total number of served UEs, the amount of data, and the battery capacity utilization level.

In addition, when compared with FPP-I∂E and FPP-∂E, FPP-IDA and FPP-DA have the wider service range and thus have the better fairness. This is mainly because in FPP-IDA and FPP-DA, the increase in the number of times of being served will reduce the probability of being served again, which in turn limits the performance improvement in terms

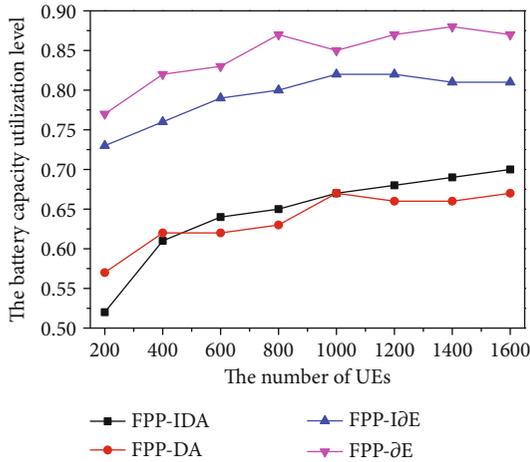


FIGURE 14: The battery capacity utilization level as a function of UEs.

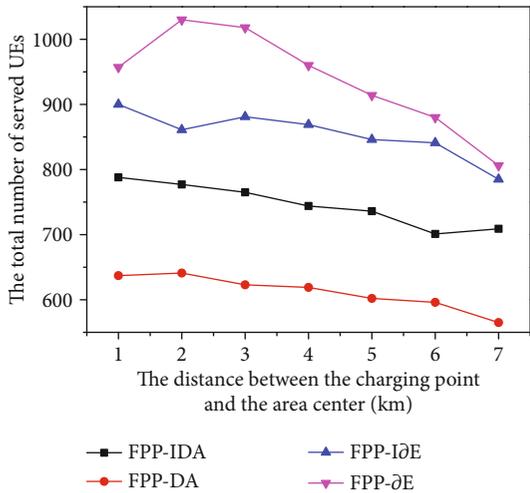


FIGURE 15: The total number of served UEs as a function of distance.

of the total number of served UEs, the amount of data, and the battery capacity utilization level due to more flight energy consumption. In Figures 10–14, the battery capacity is fixed as 1000, while the distance between the recharging point and the center of the covered area is fixed as 900 m. We simulate the performance trend of the four algorithms as the number of UEs varies.

From Figures 10–14, we can see that, as the number of UEs increases, four of all the five performance metrics of the four algorithms show an improved trend. This is due to the fact that the four algorithms can select out better planned paths in a high-density user environment, which is helpful to improve the total number of served UEs, the amount of data, the average throughput, and the battery capacity utilization level. However, in terms of the number of visited grids, the change of the number of UEs has a greater impact on the instability of FPP-∂E, especially when the number of UEs is larger, the UAV is more likely to be confined to a smaller range of flight. The main reason is that the exploration prob-

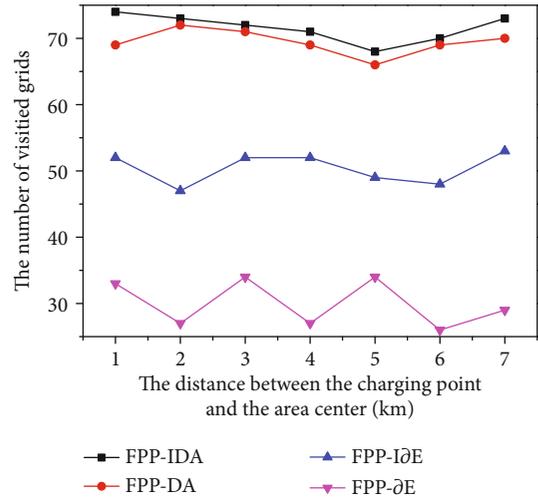


FIGURE 16: The total number of visited grids as a function of distance.

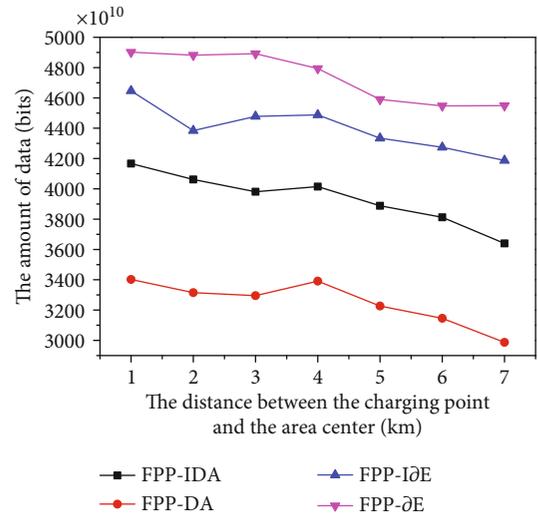


FIGURE 17: The amount of data as a function of distance.

ability ∂ is too small in FPP-∂E. For the same exploration probability, FPP-I∂E overcomes this weakness of FPP-∂E since it increases the visiting probability to the grids with conflicting records and thus extends the service range.

Figures 10–14 show that FPP-IDA is on the whole better than FPP-DA while FPP-I∂E is slightly worse than FPP-∂E in terms of the total number of served UEs, the amount of data, and the battery capacity utilization level. However, FPP-I∂E is superior to FPP-∂E in terms of the total number of visited grids. In addition, because of the larger service range, the total number of served UEs in FPP-I∂E exceeds that in FPP-∂E when the number of UEs distributed on the ground is large. The main reason of the above phenomena is the same as that explained in Figures 5–9.

In Figure 15–19, the number of UEs is fixed as 1000, while the battery capacity is fixed as 100. As shown in Figures 15, 17, and 19, the performances of all the four algorithms in terms of the total number of served UEs, the

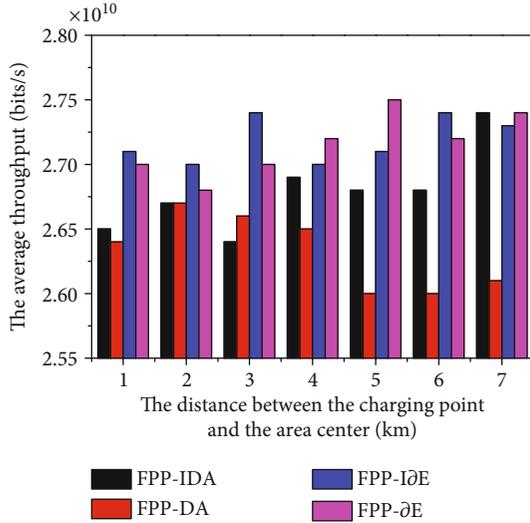


FIGURE 18: The average throughput as a function of distance.

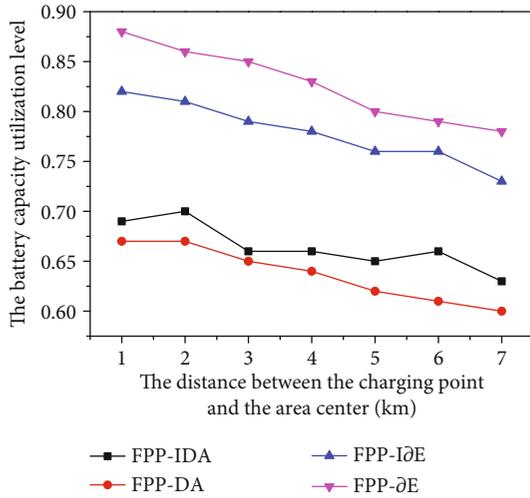


FIGURE 19: The battery capacity utilization level as a function of distance.

amount of data, and the battery capacity utilization level decrease as the distance between the recharging point and the center of the covered area increases. The main reason is that the longer distance between the recharging point and the center of the covered area requires more energy to be reserved to return to the recharging point, so there is less energy to serve the UEs when the battery capacity is fixed.

Figure 16 shows that there is no significant correlation between the number of visited grids and the distance between the recharging point and the center of the covered area, while Figure 18 shows that there is no significant correlation between the average throughput and the distance between the recharging point and the center of the covered area. In fact, the former mainly depends on the strategy function for selecting grids, while the latter mainly relies on the selected grids.

6. Conclusion

In this paper, we have addressed the performance optimization problem in UAV-assisted wireless powered mmWave networks for emergency communications. The optimization task of concern is transformed into an extended MAB problem, for which we have proposed FPP-IDA and FPP-IDE to solve efficiently, respectively. Our algorithms improve the ability to explore the unknown user distribution. Simulation results show that FPP-IDA outperforms FPP-DA in terms of all the five metrics while FPP-IDE outperforms FPP-DE in terms of the number of visited grids.

Data Availability

The simulation data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (No. 61873352).

References

- [1] Z. Fang, J. Wang, C. Jiang, Q. Zhang, and Y. Ren, "AoI inspired collaborative information collection for AUV assisted internet of underwater things," *IEEE Internet of Things Journal*, 2021.
- [2] C. Wang, J. Lu, X. Ding, C. Jiang, J. Yang, and J. Shen, "Design, modeling, control, and experiments for a fish-robot-based IoT platform to enable smart ocean," *IEEE Internet of Things Journal*, vol. 8, no. 11, pp. 9317–9329, 2021.
- [3] M. Erdelj and E. Natalizio, "UAV-assisted disaster management: applications and open issues," in *2016 international conference on computing, networking and communications (ICNC)*, pp. 1–5, Kauai, HI, USA, 2016.
- [4] J. Wang, C. Jiang, Z. Han, Y. Ren, R. G. Maunder, and L. Hanzo, "Taking drones to the next level: cooperative distributed unmanned aerial-vehicular networks for small and mini drones," *IEEE vehicular technology magazine*, vol. 12, no. 3, pp. 73–82, 2017.
- [5] A. Merwaday, A. Tuncer, A. Kumbhar, and I. Guvenc, "Improved throughput coverage in natural disasters: unmanned aerial base stations for public-safety communications," *IEEE vehicular technology magazine*, vol. 11, no. 4, pp. 53–60, 2016.
- [6] X. H. Deng, Y. J. Liu, C. X. Zhu, and H. G. Zhang, "Air-ground surveillance sensor network based on edge computing for target tracking," *Computer Communications*, vol. 166, pp. 254–261, 2021.
- [7] Y. Zeng, R. Zhang, and T. J. Lim, "Wireless communications with unmanned aerial vehicles: opportunities and challenges," *IEEE Communications Magazine*, vol. 54, no. 5, pp. 36–42, 2016.
- [8] Y. Lin, T. Y. Wang, and S. W. Wang, "UAV-assisted emergency communications: an extended multi-armed bandit

- perspective,” *IEEE Communications Letters*, vol. 23, no. 5, pp. 938–941, 2019.
- [9] J. Vermorel and M. Mohri, “Multi-armed bandit algorithms and empirical evaluation,” in *European conference on machine learning*, pp. 437–448, Berlin, Heidelberg, 2005.
 - [10] X. H. Deng, Z. H. Sun, and J. Luo, “User-centric computation offloading for edge computing,” *IEEE Internet of Things Journal*, 2021.
 - [11] L. L. Wang, J. S. Gui, X. H. Deng, F. Zeng, and Z. F. Kuang, “Routing algorithm based on vehicle position analysis for internet of vehicles,” *IEEE Internet of Things Journal*, vol. 7, no. 12, pp. 11701–11712, 2020.
 - [12] X. W. Dai and J. S. Gui, “Joint access and backhaul resource allocation for D2D-assisted dense mmWave cellular networks,” *Computer Networks*, vol. 183, p. 107602, 2020.
 - [13] P. Zhou, X. M. Fang, X. B. Wang, Y. Long, R. He, and X. Han, “Deep learning-based beam management and interference coordination in dense mmWave networks,” *IEEE Transactions on Vehicular Technology*, vol. 68, no. 1, pp. 592–603, 2019.
 - [14] C. Huang, R. Wang, M. Xie, and S. Zhang, “A parallel joint optimized relay selection protocol for wake-up radio enabled WSNs,” *Physical Communication*, vol. 47, p. 101320, 2021.
 - [15] J. S. Gui, L. H. Hui, N. N. Xiong, and J. Wu, “Improving spectrum efficiency of cell-edge devices by incentive architecture applications with dynamic charging,” *IEEE Transactions on Industrial Informatics*, vol. 17, no. 2, pp. 795–808, 2021.
 - [16] M. Hua, Y. Wang, Z. Zhang, C. Li, Y. Huang, and L. Yang, “Power-efficient communication in UAV-aided wireless sensor networks,” *IEEE Communications Letters*, vol. 22, no. 6, pp. 1264–1267, 2018.
 - [17] J. Lyu, Y. Zeng, R. Zhang, and T. J. Lim, “Placement optimization of UAV-mounted mobile base stations,” *IEEE Communications Letters*, vol. 21, no. 3, pp. 604–607, 2017.
 - [18] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, “Unmanned aerial vehicle with underlaid device-to-device communications: performance and tradeoffs,” *IEEE Transactions on Wireless Communications*, vol. 15, no. 6, pp. 3949–3963, 2016.
 - [19] S. W. Zhang and J. J. Liu, “Analysis and optimization of multiple unmanned aerial vehicle-assisted communications in post-disaster areas,” *IEEE Transactions on Vehicular Technology*, vol. 67, no. 12, pp. 12049–12060, 2018.
 - [20] K. G. Panda, S. Das, D. Sen, and W. Arif, “Design and deployment of UAV-aided post-disaster emergency network,” *IEEE Access*, vol. 7, pp. 102985–102999, 2019.
 - [21] X. Liu and N. Ansari, “Resource allocation in UAV-assisted M2M communications for disaster rescue,” *IEEE Wireless Communications Letters*, vol. 8, no. 2, pp. 580–583, 2019.
 - [22] M. Liu, J. Yang, and G. Gui, “DSF-NOMA: UAV-assisted emergency communication technology in a heterogeneous internet of things,” *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 5508–5519, 2019.
 - [23] N. Zhao, W. D. Lu, M. Sheng et al., “UAV-assisted emergency networks in disasters,” *IEEE Wireless Communications*, vol. 26, no. 1, pp. 45–51, 2019.
 - [24] Z. Xiao, P. Xia, and X. G. Xia, “Enabling UAV cellular with millimeter wave communication: potentials and approaches,” *IEEE Communications Magazine*, vol. 54, no. 5, pp. 66–73, 2016.
 - [25] L. Yang and W. Zhang, “Beam tracking and optimization for UAV communications,” *IEEE Transactions on Wireless Communications*, vol. 18, no. 11, pp. 5367–5379, 2019.
 - [26] L. Yang and W. Zhang, “Hierarchical codebook and beam alignment for UAV communications,” in *2018 IEEE Globecom Workshops (GC Wkshps)*, pp. 1–6, Abu Dhabi, United Arab Emirates, 2018.
 - [27] P. Susarla, Y. S. Deng, G. Destino et al., “Learning-based trajectory optimization for 5G mmWave uplink UAVs,” in *2020 IEEE International Conference on Communications Workshops (ICC Workshops)*, pp. 1–7, Dublin, Ireland, 2020.
 - [28] W. Yi, Y. Liu, M. ElKashlan, and A. Nallanathan, “Modeling and coverage analysis of downlink UAV networks with mmWave communications,” in *2019 IEEE International Conference on Communications Workshops (ICC Workshops)*, pp. 1–6, Shanghai, China, 2019.
 - [29] W. Yi, Y. Liu, E. Bodanese, A. Nallanathan, and G. K. Karagiannidis, “A unified spatial framework for UAV-aided mmWave networks,” *IEEE Transactions on Communications*, vol. 67, no. 12, pp. 8801–8817, 2019.
 - [30] X. Wang and M. C. Gursoy, “Coverage analysis for energy-harvesting UAV-assisted mmWave cellular networks,” *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 12, pp. 2832–2850, 2019.
 - [31] J. Chakareski, S. Naqvi, N. Mastrorarde, J. Xu, F. Afghah, and A. Razi, “An energy efficient framework for UAV-assisted millimeter wave 5G heterogeneous cellular networks,” *IEEE Transactions on Green Communications and Networking*, vol. 3, no. 1, pp. 37–44, 2019.
 - [32] Y. Zhu, G. Zheng, and M. Fitch, “Secrecy rate analysis of UAV-enabled mmWave networks using Matern hardcore point processes,” *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 7, pp. 1397–1409, 2018.
 - [33] Z. Feng, L. Ji, Q. Zhang, and W. Li, “Spectrum management for mmWave enabled UAV swarm networks: challenges and opportunities,” *IEEE Communications Magazine*, vol. 57, no. 1, pp. 146–153, 2019.
 - [34] C. H. Liu, K. H. Ho, and J. Y. Wu, “mmWave UAV networks with multi cell association: performance limit and optimization,” *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 12, pp. 2814–2831, 2019.
 - [35] M. T. Dabiri, H. Safi, S. Parsaeeafard, and W. Saad, “Analytical channel models for millimeter wave UAV networks under hovering fluctuations,” *IEEE Transactions on Wireless Communications*, vol. 19, no. 4, pp. 2868–2883, 2020.
 - [36] B. Yang, T. Taleb, Y. L. Shen, X. H. Jiang, and W. D. Yang, “Performance, fairness and tradeoff in UAV swarm underlaid mmWave cellular networks with directional antennas,” *IEEE Transactions on Wireless Communications*, vol. 20, no. 4, pp. 2383–2397, 2021.
 - [37] W. Z. Zhang, W. Zhang, and J. Wu, “UAV beam alignment for highly mobile millimeter wave communications,” *IEEE Transactions on Vehicular Technology*, vol. 69, no. 8, pp. 8577–8585, 2020.
 - [38] S. Kumar, S. Suman, and S. De, “Dynamic resource allocation in UAV-enabled mmWave communication networks,” *IEEE Internet of Things Journal*, 2020.
 - [39] Q. Xue, X. Fang, and C. X. Wang, “Beamspace SU-MIMO for future millimeter wave wireless communications,” *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 7, pp. 1564–1575, 2017.
 - [40] P. Liu, J. Blumenstein, N. S. Perović, M. Di Renzo, and A. Springer, “Performance of generalized spatial modulation MIMO over measured 60 GHz indoor channels,” *IEEE Transactions on Communications*, vol. 66, no. 1, pp. 133–148, 2018.

- [41] H. Ju and R. Zhang, "Throughput maximization in wireless powered communication networks," *IEEE Transactions on Wireless Communications*, vol. 13, no. 1, pp. 418–428, 2014.
- [42] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge Univ. Press, Cambridge, U.K., 2004.
- [43] J. S. Gui, L. H. Hui, and X. R. Zhou, "Improving lifetime of cell-edge smart sensing devices by incentive architecture based on dynamic charging," *IEEE Access*, vol. 7, no. 1, pp. 72703–72715, 2019.