IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 30 January 2024, accepted 1 March 2024, date of publication 7 March 2024, date of current version 15 March 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3374384

RESEARCH ARTICLE

Reinforcement Learning Placement Algorithm for Optimization of UAV Network in Wireless Communication

SAHAR BAGHDADY[1], SEYED MASOUD MIRREZAEI[1], AND RASHID MIRZAVAND[2], (Senior Member, IEEE)

[1]Electrical Engineering Department, Shahrood University of Technology, Shahrood 36199-95161, Iran
[2]Electrical and Computer Engineering Department, University of Alberta, Edmonton, AB T6G 1H9, Canada

Corresponding author: Seyed Masoud Mirrezaei (sm.mirrezaei@shahroodut.ac.ir)

ABSTRACT Unmanned aerial vehicles (UAVs) or drones have attracted much attention in wireless communication networks because of their agility, unique flexibility, low cost of implementation, and the high strength of the line-of-sight (LoS) channel. They are widely used in different scenarios. In many environments with complex geographical conditions or in situations where areas are affected by natural disasters, UAVs can be used as base stations (BSs) for downlink ground users. The article proposes a communication system using multiple UAV-mounted BSs to improve coverage rate and minimize the number of required UAVs. The problem is formulated as a mixed-integer programming problem with constraints on the quality of service (QoS) and serviceability of each UAV. A three-step method is developed to solve the problem, which includes deriving the maximum service radius of UAVs using the Karush-Kuhn-Tucker (KKT) method, minimizing the number of required UAVs using reinforcement learning (RL) algorithm, and designing the three-dimensional (3D) position and frequency band of each UAV to increase signal power and reduce interference. The simulation results show that the RLP algorithm outperforms other algorithms in terms of coverage rate, user clustering, increased signal, reduced interference, and processing time required to find the optimal solution.

INDEX TERMS Wireless communication, unmanned aerial vehicles (UAVs), base stations (BSs), three-dimensional (3D) deployment, user clustering.

I. INTRODUCTION

In recent years, the world has seen unprecedented demands for high-quality wireless services that have put incredible strain on existing cellular communication networks. UAVs offer significant advantages, such as wide operational coverage, flexible and highly controllable mobility, and comparatively low implementation costs. UAVs have been widely used in different aspects, especially for natural disaster management, monitoring public areas, military areas, and rural and remote areas that need communication. UAVs have been the focus of numerous research studies due to their remarkable mobility and capability to establish LoS connections on wireless communication networks [1], [2], [3], [4], [5], [6], [7], [8], [9], [10]. The purpose of determining the minimum number of UAVs in communication between users in UAV-based networks and improving the user's coverage is to create a secure and reliable network infrastructure for crisis management that maximizes energy efficiency, reduces interference power, and optimizes the network's lifetime.

This paper proposes a new algorithm called Reinforcement Learning Placement (RLP) that uses the RL algorithm to cluster users. Then, each cluster or UAV is assigned a frequency band using the frequency band allocation algorithm. In the next step, by optimizing the 3D deployment of drones, there will be a decrease in the number of drones used in the network. The RL algorithm is a type of machine learning that deals with how intelligent agents should perform actions in an environment to maximize the concept of cumulative reward [9].

The associate editor coordinating the review of this manuscript and approving it for publication was Oussama Habachi.

A. RELATED WORKS

The artificial bee colony-based (ABC) algorithm proposed in [10] is a meta-heuristic optimization algorithm inspired by the foraging behavior of honeybees. The algorithm aims to optimize the deployment of drones as base stations (BSs) in a wireless network to improve user coverage rates and overall network performance. One of the key features of this algorithm is its ability to simultaneously optimize multiple parameters, including the number of required drones, received power, interference power, UAVs' 3D locations, and frequency band allocation. This multi-objective optimization approach ensures that the deployment of drones is both efficient and effective in meeting the needs of users. Another important feature of the algorithm is its consideration of the serviceability of each drone. The algorithm takes into account the maximum service radius and optimal altitude of each drone to ensure that users are assigned to the appropriate drone for efficient service. Additionally, the signal-to-interference-plus-noise ratio (SINR) threshold is considered to ensure that users receive a high-quality signal without interference from other drones or external sources. In [11], a novel 3D deployment method for UAVs is introduced, utilizing the ant colony optimization (ACO) meta-heuristic algorithm to optimize the required number of UAVs for user service, thereby enhancing network performance and coverage. The use of UAVs as base stations is a notable aspect of this approach. The article focuses on obtaining the 3D positions of the drones, but it does not consider the serviceability of each UAV. This can lead to some UAVs being overloaded while others are underutilized, which can degrade the quality of service for users. Additionally, the SINR threshold is not considered in this approach. In [12], a meta-heuristic particle swarm optimization (PSO) algorithm is proposed to determine the 3D locations of UAVs acting as BSs. The article focuses on obtaining the 3D positions of the drones, without considering the serviceability of each UAV and the SINR threshold. One of the advantages of this approach is its ability to find the optimal 3D location of UAVs as base stations while minimizing the number of uncovered users. This improves the network performance and coverage while reducing the cost and complexity of deploying and managing the network.

In [13], the deployment of dynamic movement for UAVs in multi-UAV networks is explored. The paper proposes a genetic algorithm to obtain user cell division and a learning-based algorithm for the 3D positioning of drones. Additionally, a learning-based algorithm is suggested for acquiring 3D dynamic movement for UAVs. The article covers the 3D positioning of drones and considers the SINR threshold, which can help to ensure that the quality of service for users is not degraded due to interference. However, the serviceability of each UAV is not considered in this approach. This can lead to some UAVs being overloaded while others are underutilized, which can affect the overall network performance. In [14], the efficient placement of drones as wireless BS for optimal user coverage is examined. The theory of circular packing is proposed to determine the 3D locations of drones to maximize coverage. However, the article does not consider the serviceability of each UAV and the SINR threshold. In [15], the edge prior placement (EPP) algorithm was examined for the 3D deployment of drones acting as BSs. The algorithm aims to cover all ground users while minimizing the number of required drones in the communication network. The article takes into account the 3D positions of the drones and considers the serviceability of each drone. However, the SINR threshold is not considered in this particular investigation. Without considering the SINR threshold, there is a risk of interference between drones or from external sources, which can degrade the quality of the signal received by users. Overall, the EPP algorithm offers a cost-effective solution for the 3D deployment of drones as BSs to cover all ground users. However, it may not be suitable for applications that require high-quality signal transmission and reception due to its lack of consideration for the SINR threshold.

In [16], a spiral mobile base station deployment algorithm is presented to minimize the number of mobile base stations required to serve all ground users by optimizing drone deployment. The article focuses on obtaining the 2D position of drones and does not consider the serviceability of each UAV and the SINR threshold. In [17], the elephant herd optimization algorithm [18] is employed for the fixed deployment of drones to minimize the number of drones needed to serve all users in the network. The article focuses on obtaining the 2D position of drones and does not consider the serviceability of each UAV and the SINR threshold. In [19], a 3D deployment algorithm for multiple mobile base stations mounted on UAVs, based on quality of service (QoS) for ground user coverage, has been proposed. The algorithm takes into account the QoS requirements of each user. The article aims to optimize the altitude and coverage radius of drones. It considers the 3D position of the drones and the serviceability of each drone but does not address the SINR threshold.

In [20], an improved particle swarm optimization algorithm was proposed to optimize the positions of drones in cellular networks. The goal of this optimization was to minimize the number of drones required to serve all users effectively. Additionally, the article suggests using a clustering method based on the K-means algorithm to determine the locations of the drones. This approach helps in generating the 3D positions of the drones. However, the serviceability of each UAV and the SINR threshold are not considered in this approach. These factors are important in ensuring that the network performance meets the desired QoS requirements of the users. In [21], the focus is on the 3D deployment of base station-based drones in a cellular network and the management of backhaul to minimize the number of drones required to cover ground users. The approach used in the article involves stochastic optimization techniques. In this article, the 3D position of the drones has been obtained. However, the serviceability of each UAV and the SINR

threshold are not considered in this approach. In [22], the focus is on investigating two drone deployment algorithms based on BSs to minimize the number of drones required to serve users and enhance the coverage for ground users in a cellular network. The determination of the 3D positions of the drones is addressed in this study. However, it is noted that the serviceability of each UAV is not considered. Therefore, while the positions of the drones are obtained, the article does not take into account the capability of each drone to effectively serve the users on the ground. Moreover, the SINR threshold, which is an important metric in evaluating the quality of wireless communication, is not considered in this article.

In [23], the focus is on investigating a novel meta-heuristic method for the 3D dynamic deployment of multiple base station-based UAVs in a cellular network. The article proposes the use of the PSO algorithm and the electromagnetism-like algorithm (EML) to address this deployment issue. The determination of the 3D positions of the drones is considered in this study. However, it is important to note that the serviceability of each UAV is not taken into account. Additionally, the SINR threshold, which is a crucial factor in evaluating the quality of wireless communication, is not considered in this article. In [24], the optimization of UAV deployment in wireless networks is investigated, considering base stations and user power allocation, with the assistance of non-orthogonal multiple access (NOMA). The article proposes the utilization of the K-means algorithm for grouping ground users and the balanced gray wolf optimization (B-GWO) algorithm to solve the non-confrontational optimization problem. The optimization problem is formulated as a non-confrontational problem, where the objective function is to maximize the sum rate of all users while ensuring that each user's QoS requirements are met. The determination of the 3D positions of the drones is addressed in this study. However, it is important to note that the serviceability of each UAV is not taken into account. Furthermore, the SINR threshold, which is a crucial parameter in evaluating the quality of wireless communication, is not considered in this article. In [25], the Q-learning algorithm is used to optimize the placement of the drones in 3D space, taking into account the location of the users and the available resources. In this paper, SDQ-H is an approach that uses distributed learning to help drones serve users. The proposed approach focuses on maximizing the number of users served by each drone while ensuring that the drones do not interfere with each other. The determination of the 3D positions of the drones is addressed in this study. However, it is important to note that the serviceability of each UAV is not taken into account. Furthermore, the SINR threshold, which is a crucial parameter in evaluating the quality of wireless communication, is not considered in this article.

In [26], 3D position optimization has been investigated for both uplink and downlink transmissions, which could provide a more comprehensive solution for the UAV-assisted communication networks. However, the proposed approach in the article considers several factors such as the number of users served by each drone, the distance between drones and users, and the interference between drones. However, the lack of consideration for the serviceability of each UAV and the SINR threshold could limit the applicability of the approach in practical scenarios. The serviceability of each UAV is an important factor in ensuring that the network can operate reliably, and the SINR threshold is critical for maintaining a certain level of quality of service for the users. In [27], a multi-agent deep reinforcement learning (DRL) algorithm is used to optimize task offloading and resource allocation in multi-UAV-enabled IoT edge networks. By optimizing the task offloading and resource allocation process, the study aims to improve the overall system performance, which in turn will lead to better QoS for users. This is achieved by reducing task completion time and improving network throughput, which are two key factors that contribute to the overall user experience. The determination of the 2D positions of the drones is addressed in this study. However, it is important to note that the serviceability of each UAV is not taken into account. Furthermore, the SINR threshold, which is a crucial parameter in evaluating the quality of wireless communication, is not considered in this article. In [28], a deep reinforcement learning (DRL) algorithm is used to optimize the placement and mobility of UAVs for content caching in mobile edge networks. The goal is to maximize the QoS of users by reducing content delivery latency and improving the network throughput. The DRL-based approach involves training a UAV agent to learn how to make decisions about content caching and placement based on the current state of the network. The determination of the 3D positions of the drones is addressed in this study. However, it is important to note that the serviceability of each UAV is not taken into account. Furthermore, the SINR threshold, which is a crucial parameter in evaluating the quality of wireless communication, is not considered in this article.

B. MOTIVATION AND CONTRIBUTIONS

In this paper, a multi-UAV system has been investigated to serve and cover fixed-ground users. Additionally, UAVs have been used as BSs. A new algorithm is proposed to optimize the placement of multiple UAVs as BSs to provide coverage to ground users. Our goal is to minimize the number of drones required by clustering users appropriately, improving the coverage rate by 3D optimization of the drones and adjusting the appropriate altitude of drones, and frequency band allocation to avoid inter-cluster interference. By clustering users and optimizing the placement of drones, the algorithm aims to achieve maximum coverage with the minimum number of drones required.

In this article, compared to related articles, a smaller number of drones have been used to serve ground users, and the coverage rate of drones has increased, so that a coverage

rate of 100% has been achieved. The received power has increased and the interference power has decreased. Also, the execution time of the program to solve the problem is less than other algorithms.

This article's main contributions are as follows:

1) The proposed solution involves formulating a 3D deployment problem that takes into account both the service ability of UAVs and the QoS requirements of each user. This problem is shown to be a mixed-integer programming problem with NP-hard complexity, which means that finding an optimal solution is computationally difficult. Therefore, the proposed solution offers a sub-optimal approach that involves three steps: determining the maximum service radius, clustering users based on their QoS requirements, and deploying UAVs in a 3D space to maximize the number of users served while minimizing interference and collisions.
2) In the first step of the proposed approach, the proposed solution uses the Karush-Kuhn-Tucker (KKT) conditions to derive the optimal altitude and corresponding maximum service radius for each UAV. The KKT conditions are a set of necessary conditions for optimization problems with constraints, and they allow us to find the optimal solution by considering both the objective function and the constraints. In this case, the objective is to maximize the number of users served while ensuring that the QoS requirements are met, and the constraint is the altitude limit of the UAVs. By using the KKT conditions, the optimal altitude and service radius can be determined for each UAV, which maximizes the coverage area and ensures fairness in the number of users served by each UAV.
3) In the second step of the proposed approach, the goal is to minimize the number of user clusters while taking into account the maximum service radius and the serviceability of each UAV. To achieve this goal, a user clustering algorithm based on the RLP algorithm is proposed. The RLP algorithm is used to optimize the placement of UAVs to maximize the number of users served while minimizing the number of drones required. The user clustering algorithm then uses this optimized placement to group users into clusters that can be served by a single UAV. The goal is to minimize the number of clusters while ensuring that each cluster can be served by a single UAV within its service radius and serviceability constraints.
4) In the third step of the proposed approach, the appropriate altitude is determined by considering various factors such as the coverage area, interference, and energy consumption. The altitude of each UAV is optimized to maximize coverage while minimizing interference and energy consumption. Once the optimal 3D location and altitude of each UAV are determined, the frequency bands are assigned to each cluster. The goal is to ensure that no two identical frequency bands are adjacent to each other, as this can cause interference and degrade the QoS of users.
5) In this step, the efficiency of the RLP algorithm in the network is described and compared with other methods. According to the simulation results, the RLP algorithm has performed better than other algorithms in minimizing the number of drones used in the communication system, increasing coverage rate, increasing received signal power, reducing interference signal power, and then having the lowest system processing time.

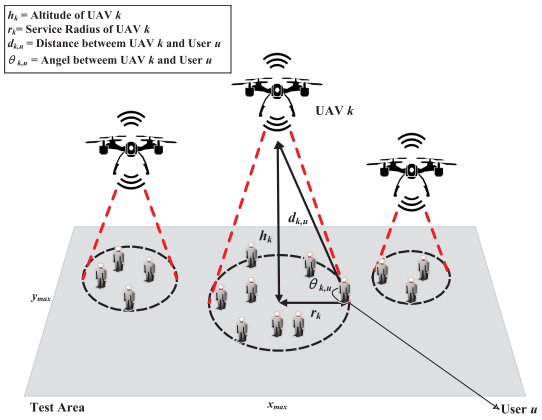


FIGURE 1. System model.

C. ORGANIZATION

The system model for drone deployment and problem formulation is introduced in Section II. The solution to the problem is presented in section III. Additionally, the RLP algorithm is proposed for user clustering and minimizing the number of drones. The simulation results and comparison of the proposed method with other algorithms are presented in section IV. Finally, the conclusion is presented in the V section.

II. SYSTEM MODEL AND PROBLEM FORMULATION

Figure 1 shows a downlink wireless communication system with several drones that are used as BSs, and transmit data and information to ground users. In our communication system, network backbone connections between the drones and access links are used between drones and ground users [29], [30]. Users in the system are randomly distributed in a 2D area $\mathcal{A} = \{(x_\mathcal{A}, y_\mathcal{A}) | x_{min} \leq x_\mathcal{A} \leq x_{max}, y_{min} \leq y_\mathcal{A} \leq y_{max}\}$.

A set of users is represented by the set $\mathcal{U} = \{1, 2, \ldots, U\}$ in the system. Each user $u \in \mathcal{U}$ has a fixed position in the 2D environment, and the fixed location of every user is denoted by $p_u = [x_u, y_u]^T \in \mathbb{R}^{2\times 1}$. UAVs must be deployed in the 3D region such that $\mathcal{P}_k = \{(x_k, y_k, h_k) | x_{min} \leq x_k \leq x_{max}, y_{min} \leq y_k \leq y_{max}, h_{min} \leq h_k \leq h_{max}\}$, and $x_{min} < x_{max}, y_{min} < y_{max}, h_{min} < h_{max}$. The set of UAVs is represented by $\mathcal{K} = \{1, 2, \ldots, K\}$. The 3D location of every UAV is represented $k \in \mathcal{K}$ by $p_k = [x_k, y_k, h_k]^T \in \mathbb{R}^{3\times 1}$.

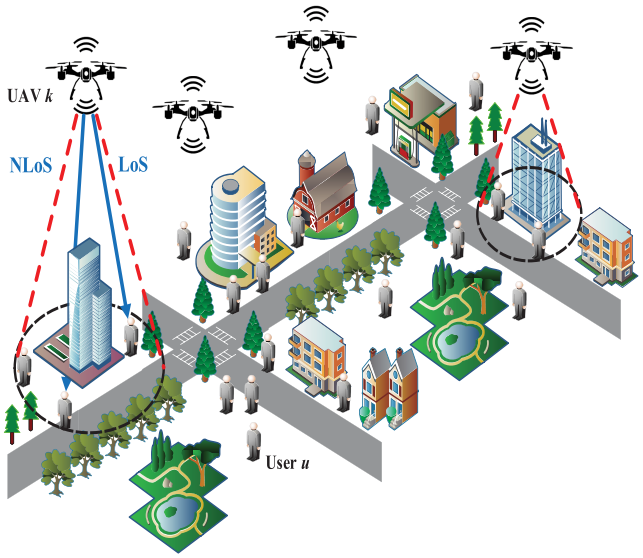


FIGURE 2. Channel model.

The set of frequency bands in the network is represented by the set $\mathcal{F} = \{1, 2, \ldots, F\}$, and it is assumed that the same frequency bands are not assigned to two adjacent clusters so that there is less interference in the system. Each UAV uses only one frequency in the communication system to serve the fixed users. The frequency band of the drone k is presented by f_k, and $\mathcal{K}_f$ represents a set of drones that work with the f band. $\mathcal{S}_{max}$ shows the serviceability of each UAV. Therefore, $\mathcal{S}_{max}$ is the maximum number of users that the drones can support in the system.

A. CHANNEL MODEL

In wireless communication, the environment is not always LoS, and because of obstacles such as buildings and trees, the communication between drones and users may be blocked, and the existence of these obstacles creates a non-line-of-sight (NLoS) environment. A complex urban environment is considered, so the channel model includes a mixed environment of LoS and NLoS environments. Figure 2 shows the communication of UAVs with ground users.

Therefore, between the drone k and the user u, the channel scale factor is defined according to [31] by the equation (1):

$$\xi_{k,u}(d_{k,u}) = \begin{cases} \xi_0 d_{k,u}^{-\nu}, & \text{LoS environment,} \\ \tau \xi_0 d_{k,u}^{-\nu}, & \text{NLoS environment,} \end{cases} \tag{1}$$

In equation (1), for the LoS environment, ξ_0 represents the path loss in d_0, where $d_0 = 0$ m is the reference distance, τ indicates the attenuation loss of NLoS environments, and ν represents modeling the parameter that relates to path loss, and $d_{k,u}$ represents the distance between the drone k and the user u, $d_{k,u}$ is defined as an equation (2):

$$d_{k,u} = \sqrt{{h_k}^2 + {s_{k,u}}^2} = \frac{s_{k,u}}{\cos\theta_{k,u}}, \tag{2}$$

In equation (2), h_k represents the altitude of UAVs, $s_{k,u}$ represents the predicted 2D distance between UAV k and user u, which is defined as $s_{k,u} = \sqrt{(x_u - x_k)^2 + (y_u - y_k)^2}$, and $\theta_{k,u}$ represents the altitude angle between the UAV k and the user u.

The probability of the LoS link between drone k and the user u according to [31] is defined as an equation (3):

$$Pr_{k,u}^{LoS}(\theta_{k,u}) = \frac{1}{1 + \alpha \exp(-\beta(\theta_{k,u} - \alpha))}, \tag{3}$$

In equation (3), α and β are represent the environment modeling parameters.

The probability of the NLoS link between the drone k and the user u is defined as an equation (4):

$$Pr_{k,u}^{NLoS}(\theta_{k,u}) = 1 - Pr_{k,u}^{LoS}(\theta_{k,u}), \tag{4}$$

The average probability of path loss is defined as an equation (5) by [32]:

$$\Gamma = Pr_{k,u}^{LoS} \times PL_{k,u}^{LoS} + Pr_{k,u}^{NLoS} \times PL_{k,u}^{NLoS}, \tag{5}$$

In equation (5), $PL_{k,u}^{LoS}$ and $PL_{k,u}^{NLoS}$ indicate path loss of the LoS link and the NLoS link, respectively, are obtained by equations (6) and (7) [32]:

$$PL_{k,u}^{LoS} = 20\log(\frac{4\pi f_c d_{k,u}}{c}) + \zeta_{LoS}, \tag{6}$$

$$PL_{k,u}^{NLoS} = 20\log(\frac{4\pi f_c d_{k,u}}{c}) + \zeta_{NLoS}, \tag{7}$$

In equations (6) and (7), f_c is the carrier frequency.

By combining environmental constant parameters and equations (3), (4), (6), and (7) in (5), equation (8) is obtained [32]:

$$PL_{max} = \frac{W}{1 + \alpha \exp(-\beta[\arctan(\frac{h_k}{r_k}) - \alpha])} + 10\log({h_k}^2 + {r_k}^2) + Z, \tag{8}$$

In the above equation, h_k and r_k are the altitudes and the coverage radius or the service radius of the drone respectively, and also $W = \zeta_{LoS} - \zeta_{NLoS}$ and $Z = 20\log f_c + 20\log(\frac{4\pi}{c}) + \zeta_{NLoS}$.

According to equation (8), there is a relationship between the radius of drone coverage and altitude, and it is defined as equation (9) [32]:

$$\frac{\partial r_k}{\partial h_k} = 0, \tag{9}$$

According to the above equation, if the coverage radius of a drone is available, the optimal height of the 3D deployment of drones can be obtained.

The channel gain between the drone k and the user u is defined as equation (10) [31]:

$$\begin{aligned} \bar{g}_{k,u}(d_{k,u}, \theta_{k,u}) &\triangleq Pr_{k,u}^{LoS}(\theta_{k,u})\xi_0 {d_{k,u}}^{-\nu} \\ &\quad + Pr_{k,u}^{NLoS}(\theta_{k,u})\tau\xi_0 {d_{k,u}}^{-\nu} \\ &= \hat{P}_{k,u}^{LoS}(\theta_{k,u})\xi_0 {d_{k,u}}^{-\nu}, \end{aligned} \tag{10}$$

In equation (10), $\hat{P}_{k,u}^{LoS}(\theta_{k,u}) = Pr_{k,u}^{LoS}(\theta_{k,u}) + Pr_{k,u}^{NLoS}(\theta_{k,u})\tau$ represents a regular LoS link [29], which includes both environments LoS and NLoS.

For every user u, SINR represents the measurement of communication quality and is defined by equation (11) [10]:

$$N_{k,u} = \frac{P_{k,u}}{I_u + \sigma^2}, \tag{11}$$

In equation (11), $P_{k,u}$ represents the received signal power by user u from drone k, I_u represents the power of interference in user u, and σ^2 represents the Gaussian white noise power. $P_{k,u}$, and I_u can be defined by equations (12) and (13), respectively [10]:

$$P_{k,u} = \bar{g}_{k,u} \times P_T, \tag{12}$$

$$I_u = \sum_{k' \in \mathcal{K}_b \setminus k} P_{k',u} \tag{13}$$

In equations (12) and (13), P_T represents the signal power sent to every user u and f represents the band used by drone k. The power of the transmitted signal is assumed to be the same for every user in the communication system. Therefore, the maximum transmitted power for every drone is at the upper limit with $\mathcal{S}_{max}P_T$. When $N_{k,u} \geq N_0$, user u has successfully communicated with UAV k, and N_0 represents the SINR threshold.

The index function $\eta_{k,u}$ to show the relationship between UAVs and users is defined as equation (14) [10]:

$$\eta_{k,u} = \begin{cases} 1, & \text{user } u \text{ is covered by drone } k, \\ 0, & \text{otherwise}, \end{cases} \tag{14}$$

In equation (14), when $\eta_{k,u} = 1$, this means that the user u is supported by drone k, and otherwise, if $\eta_{k,u} = 0$, it means that user u is not served by drone k. Then, the coverage rate $\mathcal{C}$ is obtained by equation (15), which represents the percentage of users successfully supported by UAVs [10]:

$$\mathcal{C} = \frac{\sum_{k \in \mathcal{K}} \sum_{u \in \mathcal{U}} \eta_{k,u}}{U}, \tag{15}$$

B. PROBLEM FORMULATION

This article aims to utilize a minimum number of drones to support a maximum number of users. To improve user coverage, more drones are needed to serve users. As a result, to establish a balance between the number of drones and the drone coverage rate, the objective function should be formulated as a weighted sum between the number of drones and the drone coverage rate. The location of drones, communication between the drones and all ground users, as well as the allocation of frequency bands, is optimized as an equation (16) [10]:

$$\min_{\{p_k\},\{\eta_{k,u}\},\{\mathcal{K}_b\}} \rho_1 |\mathcal{M}| + \frac{\rho_2}{\mathcal{C}} \tag{16}$$

$$s.t.\ \ p_k \in \mathcal{P}_k, \quad \forall k \in \mathcal{K}, \tag{16a}$$

$$\eta_{k,u} \in \{0,1\}, \quad \forall u \in \mathcal{U},\ \forall k \in \mathcal{K}, \tag{16b}$$

$$\sum_{k \in \mathcal{K}} \eta_{k,u} \leq 1, \quad \forall u \in \mathcal{U}, \tag{16c}$$

$$\sum_{u \in \mathcal{U}} \eta_{k,u} \leq \mathcal{S}_{max}, \quad \forall k \in \mathcal{K}, \tag{16d}$$

$$\bigcup_{b \in \mathcal{B}} \mathcal{K}_b = \mathcal{K}, \tag{16e}$$

$$\mathcal{K}_f \bigcap \mathcal{K}_q = \emptyset, \quad \forall f, q \in \mathcal{F},\ b \neq q, \tag{16f}$$

$$N_{k,u} \geq N_0 \eta_{k,u}, \quad \forall u \in \mathcal{U},\ \forall k \in \mathcal{K}, \tag{16g}$$

In equation (16), ρ_1, and ρ_2 represent weight coefficients and $\rho_1 \ll \rho_2$ shows the priority of coverage rate over reducing the number of UAVs. Equation (16a) limits the operating area of the 3D position of UAVs. Equation (16b) shows the range of the $\eta_{k,u}$. Equation (16c) shows that every user is supported with a maximum of one drone. Equation (16d) shows that the number of users covered by every drone should not exceed $\mathcal{S}_{max}$. Equations (16e) and (16f) show that every drone can use a frequency band to support users. Equation (16g) shows the SINR requirement, that is, the QoS for communication between the drone k and the user u.

III. PROBLEM SOLUTION

A non-optimal solution is used to solve the problem (16). It is a mixed integer programming problem with NP-hard complexity [33], proposed by [10]. In the communication system, because the coverage rate is prioritized over the number of drones, drones serve all ground users. Therefore, fewer drones are used to provide services. Additionally, the altitude of drones is optimized to obtain the largest coverage radius or service radius and reduce the maximum interference power, and also the number of users who are covered by drones must be close to $\mathcal{S}_{max}$ [10]. So, the RLP algorithm is proposed for user clustering to minimize the number of clusters or drones in the communication system. Then, the 2D locations of the drones are optimized to have the maximum received power. To establish the SINR of every user, a frequency band is assigned to every drone.

A. MAXIMUM COVERAGE RADIUS

The maximum coverage radius or R_{max} is the biggest 2D distance between a user u and a drone k [10]. According to equation (10), the maximum coverage radius or service radius is achieved by optimizing the flight altitude of UAVs. In this subsection, a solution to the maximum coverage radius of the drones is presented [10]. P_0 is defined as the minimum required power, and $\bar{g}_0 = \frac{P_0}{P_T}$ is defined as the minimum gain of the channel.

Equation (17) is defined to maximize the service radius of the UAVs [10]:

$$\max_{r_k, h_k} r_k \tag{17}$$

$$s.t.\ \ \frac{1 + \tau\alpha e^{-\beta(\arctan\frac{h_k}{r_k} - \alpha)}}{1 + \alpha e^{-\beta(\arctan\frac{h_k}{r_k} - \alpha)}} \xi_0 ({r_k}^2 + {h_k}^2)^{-\frac{v}{2}} \geq \bar{g}_0, \tag{17a}$$

$$h_{min} \leq h_k \leq h_{max}, \tag{17b}$$

In equation (17), h_k shows the altitude of the drone and r_k represents the service radius of the drone. Problem (17) is non-convex. It can be solved using the KKT condition [10]. The parameter $\theta_{k,u}$ indicates the altitude angle of a UAV relative to the ground user. Using the trigonometric equation $h_k = r_k \tan\theta_{k,u}$, equation (17) is defined as a solvable equation (18) [10]:

$$\min_{r_k,h_k} -r_k \tag{18}$$

$$s.t.\ \ \bar{g}_0 - \frac{1+\tau\alpha e^{-\beta(\theta_{k,u}-\alpha)}}{1+\alpha e^{-\beta(\theta_{k,u}-\alpha)}}\xi_0\left(\frac{r_k}{\cos\theta_{k,u}}\right)^{-\nu} \le 0, \tag{18a}$$

$$r_k \tan\theta_{k,u} - h_{max} \le 0, \tag{18b}$$

$$h_{min} - r_k \tan\theta_{k,u} \le 0, \tag{18c}$$

The Lagrange function is defined as equation (19) [10]:

$$\begin{aligned} &L(r_k,\theta_{k,u},\gamma_1,\gamma_2,\gamma_3) \\ &= -r_k + \gamma_1[\bar{g}_0 - \frac{1+\tau\alpha e^{-\beta(\theta_{k,u}-a)}}{1+\alpha e^{-\beta(\theta-\alpha)}}\xi_0\left(\frac{r}{\cos\theta_{k,u}}\right)^{-\nu}] \\ &\quad + \gamma_2(r_k\tan\theta - h_{max}) + \gamma_3(h_{min} - r_k\tan\theta_{k,u}), \end{aligned} \tag{19}$$

In equation (19), coefficients γ_1, γ_2, and γ_3 represent Lagrange coefficients.

To simplify the equations, intermediate variables are used, which are defined by equations (20), (21), (22), and (23), respectively [10]:

$$\bar{g}^* = \frac{1+\tau\alpha e^{-\beta(\theta^*_{k,u}-\alpha)}}{1+\alpha e^{-\beta(\theta^*_{k,u}-\alpha)}}\xi_0\left(\frac{R_{max}}{\cos\theta^*_{k,u}}\right)^{-\nu}, \tag{20}$$

$$\bar{g}'_1 = \frac{\xi_0\left(\frac{R_{max}}{\cos\theta^*_{k,u}}\right)^{-\nu}}{[1+\alpha e^{\Theta}]^2}, \tag{21}$$

$$\bar{g}'_2 = \frac{180}{\pi}(1-\tau)\alpha\beta e^{\Theta} - \nu\tan\theta^*_{k,u}(1+\tau\alpha e^{\Theta})(1+\alpha e^{\Theta}), \tag{22}$$

$$\Theta = [-\beta(\frac{180}{\pi}\theta^*_{k,u} - \alpha)]. \tag{23}$$

In equation (20), $\bar{g}^*$ indicates the optimal channel gain, which is obtained by replacing R_{max} and $\theta^*_{k,u}$ in the equation (10).

Therefore, the parameters must satisfy all the conditions of equations (24a) to (24i) according to the conditions of KKT [10]:

$$\bar{g}_0 - \bar{g}^* \le 0, \tag{24a}$$

$$R_{max}\tan\theta^*_{k,u} - h_{max} \le 0, \tag{24b}$$

$$h_{min} - R_{max}\tan\theta^*_{k,u} \le 0, \tag{24c}$$

$$\gamma_1^* \ge 0, \gamma_2^* \ge 0, \gamma_3^* \ge 0, \tag{24d}$$

$$\gamma_1^*(\bar{g}_0 - \bar{g}^*) = 0, \tag{24e}$$

$$\gamma_2^*(R_{max}\tan\theta^*_{k,u} - h_{max}) = 0, \tag{24f}$$

$$\gamma_3^*(h_{min} - R_{max}\tan\theta^*_{k,u}) = 0, \tag{24g}$$

$$\frac{\partial L}{\partial\theta_{k,u}} = -\gamma_1^*\bar{g}'_1\bar{g}'_2 + \frac{\gamma_2^* R_{max}}{\cos^2\theta^*_{k,u}} - \frac{\gamma_3^* R_{max}}{\cos^2\theta^*_{k,u}} = 0, \tag{24h}$$

$$\frac{\partial L}{\partial r_k} = -1 + \frac{\gamma_1^*\nu}{\cos\theta^*_{k,u}}\frac{1+\tau\alpha e^{\Theta}}{1+\alpha e^{\Theta}}\left(\frac{R_{max}}{\cos\theta^*_{k,u}}\right)^{-\nu-1} + \gamma_2^*\tan\theta^*_{k,u} - \gamma_3^*\tan\theta^*_{k,u} = 0, \tag{24i}$$

The optimal values of R_{max} and $\theta^*_{k,u}$ parameters are obtained according to three theorems 1, 2, and 3 [10].

Theorem 1: If $\gamma_2^* = \gamma_3^* = 0$, the value of parameter R_{max} is obtained from the equation (25):

$$R_{max} = \left(\frac{\bar{g}_0}{\hat{P}^{LoS}_{k,u}(\theta^*_{k,u})\xi_0}\right)^{-\frac{1}{\nu}}\cos\theta^*_{k,u}. \tag{25}$$

Proof: If $\gamma_2^* = \gamma_3^* = 0$, then $h_{min} < R_{max}\tan\theta^*_{k,u} < h_{max}$ is obtained according to the equations (24f) and (24g). Therefore, equations (24i) and (24h) are defined as equations (26) and (27):

$$\frac{\partial L}{\partial r_k} = -1 + \frac{\gamma_1^*\nu}{\cos\theta^*_{k,u}}\frac{1+\tau\alpha e^{\Theta}}{1+\alpha e^{\Theta}}\left(\frac{R_{max}}{\cos\theta^*_{k,u}}\right)^{-\nu-1} = 0, \tag{26}$$

$$\frac{\partial L}{\partial\theta_{k,u}} = -\gamma_1^*\bar{g}'_1\bar{g}'_2 = 0. \tag{27}$$

According to the equation (26), the value of parameter γ_1^* is given by the equation (28):

$$\gamma_1^* = \frac{1}{\frac{\nu}{\cos\theta^*_{k,u}}\frac{1+\tau\alpha e^{\Theta}}{1+\alpha e^{\Theta}}\left(\frac{R_{max}}{\cos\theta^*_{k,u}}\right)^{-\nu-1}}, \tag{28}$$

According to the above equation, the value of parameter γ_1^* is positive. Therefore, according to equation (24e), equation (29) must be satisfied:

$$\bar{g}_0 - \bar{g}^* = 0. \tag{29}$$

Since $\gamma_1^* > 0$ and $\bar{g}'_1 > 0$, equation (27) is simplified as $\bar{g}'_2 = 0$, where the parameter $\theta^*_{k,u}$ is our only unknown variable, and by putting $\theta^*_{k,u}$ in equation (29), its value is obtained, and also, the value of the parameter R_{max} is obtained by equation (25).

Theorem 2: If $\gamma_2^* > 0$, $\gamma_3^* = 0$, then the value of the R_{max} the parameter is obtained by the equation (30):

$$R_{max} = \frac{\cos\theta^*_{k,u}}{\log_{\nu+1}\frac{(1-\gamma_2^*\tan\theta^*_{k,u})\cos\theta^*_{k,u}(1+\alpha e^{\Theta})}{\gamma_1^*\nu(1+\tau\alpha e^{\Theta})}}. \tag{30}$$

Proof: If $\gamma_2^* > 0$, $\gamma_3^* = 0$, then according to equations (24f) and (24g), $R_{max}\tan\theta^*_{k,u} = h_{max}$ is obtained. Therefore, equations (24i) and (24i) are defined as equations (31) and (32):

$$\frac{\partial L}{\partial r_k} = -1 + \frac{\gamma_1^*\nu}{\cos\theta^*_{k,u}}\frac{1+\tau\alpha e^{\Theta}}{1+\alpha e^{\Theta}}\left(\frac{R_{max}}{\cos\theta^*_{k,u}}\right)^{-\nu-1} + \gamma_2^*\tan\theta^*_{k,u} = 0, \tag{31}$$

$$\frac{\partial L}{\partial\theta_{k,u}} = -\gamma_1^*\bar{g}'_1\bar{g}'_2 + \frac{\gamma_2^* R_{max}}{\cos^2\theta^*_{k,u}} = 0. \tag{32}$$

According to the equation (32), the value of parameter γ_1^* is obtained by the equation (33):

$$\gamma_1^* = \frac{\lambda_2^* R_{max}}{\cos^2\theta^*_{k,u}\bar{g}'_1\bar{g}'_2}, \tag{33}$$

According to the equation (24d), the value of the parameter γ_1^* is positive. Therefore, equation (29) must also be satisfied.

By simultaneously solving equations (29), (31), (32) and since it is from the $R_{max}\tan\theta_{k,u}^* = h_{max}$ theorem 2, the value of parameters $\theta_{k,u}^*$, γ_1^*, and γ_2^* has been obtained. Also, the value of parameter R_{max} is obtained by the equation (30).

Theorem 3: If $\gamma_2^* = 0$, $\gamma_3^* > 0$, then the value of the R_{max} parameter is obtained by the equation (34):

$$R_{max} = \frac{\cos\theta_{k,u}^*}{\log_{\nu+1}\frac{(1+\gamma_3^*\tan\theta_{k,u}^*)\cos\theta_{k,u}^*(1+\alpha e^{\Theta})}{\gamma_1^*\nu(1+\tau\alpha e^{\Theta})}}, \tag{34}$$

Proof: If $\gamma_2^* = 0$, $\gamma_3^* > 0$, then, according to the equations (24f) and (24g), $R_{max}\tan\theta_{k,u}^* = h_{min}$ is obtained. Therefore, the equations (24i) and (24h) are defined as equations (35) and (36):

$$\frac{\partial L}{\partial r_k} = -1 + \frac{\gamma_1^*\nu}{\cos\theta_{k,u}^*}\frac{1+\tau\alpha e^{\Theta}}{1+\alpha e^{\Theta}}\left(\frac{R_{max}}{\cos\theta_{k,u}^*}\right)^{(-\nu-1)} - \gamma_3^*\tan\theta_{k,u}^* = 0, \tag{35}$$

$$\frac{\partial L}{\partial\theta_{k,u}} = -\gamma_1^*\bar{g}_1'\bar{g}_2' - \frac{\gamma_3^* R_{max}}{\cos^2\theta_{k,u}^*} = 0. \tag{36}$$

According to the equation (35), the value of parameter γ_1^* is obtained by the equation (37):

$$\gamma_1^* = \frac{\gamma_3^* R_{max}}{\cos^2\theta_{k,u}^*\bar{g}_1'\bar{g}_2'}, \tag{37}$$

According to the equation (24d), the value of the parameter γ_1^* is positive. Therefore, the equation of (29) must also be satisfied.

By simultaneously solving equations (29), (35), (36) and since it is from the $R_{max}\tan\theta_{k,u}^* = h_{min}$ theorem 3, the value of parameters $\theta_{k,u}^*$, γ_1^*, and γ_3^* has been obtained. Also, the value of parameter R_{max} is obtained by the equation (34).

The optimal height of the drone is defined as $h_k^* = R_{max}\tan\theta_{k,u}^*$, and since $s_{k,u} = R_{max}$, it must always satisfy $\bar{g}_{k,u} = \bar{g}_0$.

$$\eta_{k,u}s_{k,u} \le R_{max}, \tag{38}$$

The above equation expresses the distance between the UAV k and the user u served by that UAV should not exceed the value of the R_{max} parameter.

Therefore, equation (16) is rewritten as equation (39):

$$\min_{\{p_k\},\{\eta_{k,u}\}} \quad |\mathcal{K}| \tag{39}$$

$$s.t. \quad p_k \in \mathcal{P}, \quad \forall k \in \mathcal{K}\mathcal{M}, \tag{39a}$$

$$\eta_{k,u} \in \{0,1\}, \quad \forall u \in \mathcal{U}, \forall k \in \mathcal{K}, \tag{39b}$$

$$\sum_{k\in\mathcal{K}} \eta_{k,u} = 1, \quad \forall u \in \mathcal{U}, \tag{39c}$$

$$\sum_{u\in\mathcal{U}} \eta_{k,u} \le \mathcal{S}_{max}, \quad \forall k \in \mathcal{K}, \tag{39d}$$

$$\eta_{k,u}s_{k,u} \le R_{max}, \quad \forall u \in \mathcal{U}, \forall k \in \mathcal{K}. \tag{39e}$$

Solving the equation (39) is a challenging task. Therefore, to solve the equation (39), the RLP algorithm is presented.

The RLP algorithm is composed of two parts. In the first part, users are clustered, so that every cluster is assigned a UAV. In the second part, according to the clustering of users, the 3D deployment of UAVs is optimized in such a way that the power received by the users from the UAVs is increased and the interference power of the system is reduced. Additionally, frequency bands are assigned to every UAV or every cluster in such a way as to avoid inter-cluster interference.

B. USER CLUSTERING

The RLP algorithm is proposed to solve the problem of user clustering, which is a combination of the EPP algorithm [15] and the RL algorithm. According to the serviceability of every drone, drone k serves users of group u. The total number of users that every drone or every cluster can cover is equal to $\mathcal{S}_{max}$. The largest area covered by every drone consists of a circular area. The center of the circle is the location of the drone and the radius of the circle is the coverage radius of the drone. The goal is to use the smallest number of drones to cover the maximum number of users. So every drone should have the largest coverage radius, and the number of users served by every drone should be close to $\mathcal{S}_{max}$.

The issue of clustering users can be expressed as follows:

$$\min_{\{\eta_{k,u}\}} \quad |\mathcal{K}| \tag{40}$$

$$s.t. \quad \eta_{k,u} \in \{0,1\}, \quad \forall u \in \mathcal{U}, \forall k \in \mathcal{K}, \tag{40a}$$

$$\sum_{k\in\mathcal{K}} \eta_{k,u} = 1, \quad \forall u \in \mathcal{U}, \tag{40b}$$

$$\sum_{u\in\mathcal{U}} \eta_{k,u} \le \mathcal{S}_{max}, \quad \forall k \in \mathcal{K}, \tag{40c}$$

$$\eta_{k,u}r_{k,u} \le R_{max}, \quad \forall u \in \mathcal{U}, \forall k \in \mathcal{K}. \tag{40d}$$

According to Algorithm 1, users cluster in different groups, and one drone is assigned to each cluster. As a result, all users are grouped in the communication system.

The RLP algorithm is described as follows:

1) Data Collection: This step contains information such as 2D location and users' labels, the characteristics of UAVs such as the location, maximum radius of coverage, number of users, and maximum number of learning repetitions.
2) Searching Boundaries: In this step, it obtains the search boundary and then selects the best point in that boundary using the learning method. To determine the search domain, the first task is to find the user-uncovered user center $\mathcal{U}_{un}$. According to the following formula, the center of all uncovered users can be found:

$$C_0 = \frac{1}{|\mathcal{U}_{un}|}\sum u_i, u_i \in \mathcal{U}_{un}, \tag{41}$$

C_0 is obtained and the users in the boundary area are found and their numbers are stored in L_{max}. If the value of L_{max} is smaller than the maximum learning iteration, it sets L_{max} to the maximum iteration in order not to

Algorithm 1 RLP Algorithm for User Clustering

Input: $\mathcal{U}$ user set, p_u user's location, $\mathcal{S}_{max}$, R_{max}, L_{max}
Output: Number of UAVs $\mathcal{K}, \mathcal{U}^k$

Find the center of unallocated users by solving (41)
Find users in boundaries area
Save number of the users in the boundaries region in L_{max}
if L_{max} is lower than the maximum iteration **then**
  Set the maximum iteration number to L_{max}
end if
for each user u in the boundaries area, **do**
  Calculate the distance from the center of unallocated users by solving (42)
end for
Find the furthest user based on the distance by solving (43)
Save maximum distance in Dist
Define search ring based-no users in distance area [0.95 × Dist, Dist]
Set the k number to 0
while adding one number to k **do**
  for each iteration, **do**
    Find initial coverage at maximum distance and Save it in $Base_{cov}$
    if it=1, **then**
      Set the initial probability of users in the search ring (44)
    else
      if coverage is greater than $Base_{cov}$, **then**
        Update and increase the probability of users in the search ring
      end if
      if coverage is lower than $Base_{cov}$, **then**
        Update and decrease the probability of users in the search ring
      end if
    end if
  end for
  Find the user with the highest probability in the search ring
  Set coverage number to $\mathcal{S}_{max}$
  for each user allocation try, to **do**
    Find $\mathcal{S}_{max}$ users near the user with the highest probability
    Solve optimization problem by CVX toolbox to cover selected users
    if the minimum radius is greater than R_{max}, **then**
      Reduce $\mathcal{S}_{max}$ by one number
    else
      Remove selected users from unallocated users
      Break
    end if
  end for
  if there are no unallocated users, **then**
    Break
  end if
end while
return $\mathcal{K}, \mathcal{U}^k$

perform this search too much. Then, for each user in the boundary area, it obtains the distance of each user from the center according to the following formula:

$$\mathcal{D}_0 = \sqrt{\Delta x^2 + \Delta y^2}, \tag{42}$$

After obtaining the distance, it calculates the farthest point according to the formula below and stores the greatest distance in (Dist):

$$u_0 = \max(\mathcal{D}_{10}, \mathcal{D}_{20}, \ldots), \tag{43}$$

One circle is formed with the maximum distance, and another circle is formed with 95% of the distance. Therefore, the search circle is obtained according to the maximum and minimum distance.

3) Learning process: After finding the search circle, the number of drones k is set to 0 at this stage. Adds a number to k and assigns the first UAV to cover a cluster of users. In each iteration, it finds the initial coverage and stores it in $Base_{cov}$. Now, if it is the initial iteration, it sets the probability of all users in the search circle to be selected equal to the initial value, because it is the first iteration and has no opinion about which point is better. The following formula can be used to calculate the probability of the initial iteration:

$$prob = \frac{1}{|\mathcal{U}_{cov}|}, \tag{44}$$

$\mathcal{U}_{cov}$ are users covered.
Now, if it is not in the first iteration and it is selected, it checks the coverage to see if the coverage has improved or worsened. So, it checks that the coverage it received is greater than $Base_{cov}$. It updates the probability of users inside the search loop and finds the best user for each iteration. Finds the user with the highest probability in the search circle. The largest number of users that any drone can cover is $\mathcal{S}_{max}$. Therefore, it finds $\mathcal{S}_{max} - 1$ users closest to the user with the highest probability and clusters the users. In the following, it optimizes the position of the UAVs using the CVX toolbox [34]. Minimizes the coverage radius of UAVs. If the minimum coverage radius is greater than R_{max}, it uses $\mathcal{S}_{max} - 1$, otherwise, it removes the selected users from the set of uncovered users. These steps are carried out until all users are covered by UAVs.

4) Output Data: The output of the algorithm will be the number of drones, the radius of service or the coverage radius of every drone, how many users every UAV has supported, and also which UAV served which user.

Figure 3 shows how users are clustered in the RLP algorithm in the UAV network. First, it finds the center of the region it wants to cluster and then starts clustering from the boundary region.

The complexity of an RLP algorithm for user clustering in a UAV network can vary depending on the specific approach used and the size of the problem being solved. In the case

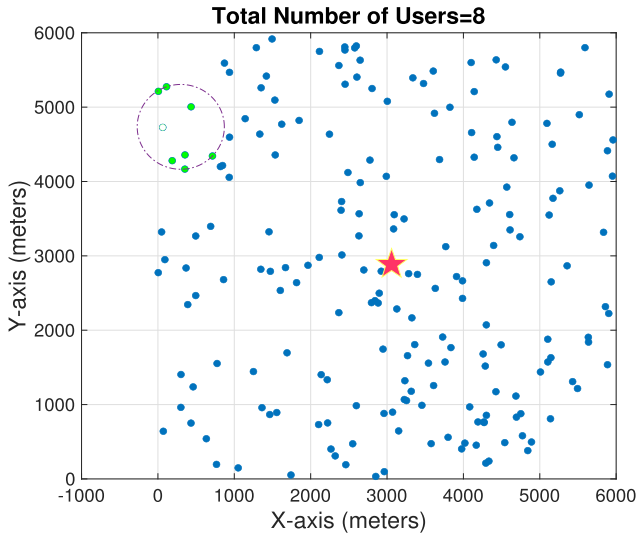


FIGURE 3. User clustering method with RLP algorithm.

of user clustering in a UAV network, the complexity of the RLP algorithm will depend on several factors such as the number of users, the number of UAVs, the size of the search space, and the complexity of the reward function. The computational complexity of the proposed algorithm for user clustering is $\mathcal{O}(L_{max}\,\mathcal{U}\,|\mathcal{K}|)$, where L_{max} is the number of training episodes, $\mathcal{U}$ is the number of users, and $|\mathcal{K}|$ is the number of UAVs.

C. 3D DEPLOYMENT AND FREQUENCY BAND ALLOCATION

For the set $\mathcal{M} = \{\mathcal{M}^1, \ldots, \mathcal{M}^{|\mathcal{K}|}\}$, user $m \in \mathcal{M}^k$, $\eta_{k,m} = 1$ must be satisfied. For every cluster of users, a UAV is assigned to cover users in the network. The purpose of the 3D deployment of UAVs is to improve SINR. First, the 3D location of UAVs is optimized, then frequency bands are assigned to every UAV, and finally, by adjusting the altitude, the interference in the network is reduced.

1) 2D Deployment: In this part, the 2D positioning of UAVs is optimized. The channel gain of the users covered by UAV is not considered in the clustering of users. Therefore, the location of UAVs should be optimized, and the QoS of every user should be considered. The parameter $s_{k,m}$ is the 2D distance between the drone k and the user m in $\mathcal{M}^k$, and its value should be as small as possible to maximize the channel gain. Therefore, to minimize the value of the parameter $s_{k,m}$, equation (45) is defined [10]:

$$\min_{\{x_k,y_k\}} \quad \max_{m\in\mathcal{M}^k} \; s_{k,m} \tag{45}$$

$$s.t. \quad x_{min} \le x_k \le x_{max}, \quad \forall k \in \mathcal{K}, \tag{45a}$$

$$y_{min} \le y_k \le y_{max}, \quad \forall k \in \mathcal{K}, \tag{45b}$$

Equation (45) is convex optimization and solves the optimization problem with the CVX optimization toolbox [34]. Then, it is possible to obtain the 2D coordinates of UAVs, namely (x_k, y_k). The parameter r^k_{min} is the minimum service radius of UAV k to cover the users in $\mathcal{M}^k$, and its value is obtained by equation (46) [10]:

$$r^k_{min} = \max_{m\in\mathcal{M}^k} s_{k,m}. \tag{46}$$

2) Frequency Band Allocation: When the 2D location of the drones is optimized, it allocates frequency bands to the drones or clusters in such a way as to prevent inter-cluster interference. Since $P_{k,u} \ge p_0$ must be satisfied, for the SINR of equation $N_{k,u} \ge N_0$ to hold, the upper limit of interference of user u from UAV $k' \in \mathcal{K}_{b_k}$ is obtained by equation (47) [10]:

$$P_{k',u} \le \frac{P_0/N_0 - \sigma^2}{|\mathcal{K}| - 1}, \tag{47}$$

Using equation $\bar{g}_{k',u} = \frac{P_{k',u}}{P_T}$, the channel gain between the drone k' and the user u can be obtained, and the equation (47) is equivalent to equation (48) [10]:

$$\bar{g}_{k',u} \le \frac{P_0/N_0 - \sigma^2}{(|\mathcal{K}| - 1) \times P_T} \triangleq \hat{g}_0. \tag{48}$$

The minimum 2D distance between the drone k' and the user u indicates the minimum interference radius and the channel interference should not exceed $\hat{g}_0$. By substituting the desired altitude h^*_k in equation (48), the interference radius R_I is obtained. If $s_{k',u} \ge R_I, \quad \forall k' \in \mathcal{K}_{b_k}$, the user u can satisfy their SINR constraint [10].

The frequency band allocation algorithm is defined to reduce inter-cluster interference. $\mathcal{K}_{un}$ is a set of UAVs that have not yet been assigned frequency bands and are defined in line 1 of Algorithm 2. Lines 2 and 3, $|\mathcal{F}|$ Indicate the frequency bands assigned to UAVs. In lines 4 to 16, frequency bands are assigned to every UAV or every cluster, respectively. In line 6, it determines the frequency of a new UAV in each run, and K_f is updated so that two UAVs or clusters next to each other do not have the same frequency. For example, the same frequency may be assigned to 3 clusters. So it finds the minimum distance for that frequency. Next, remove the frequency band assigned to that cluster from the frequency domain and assign another frequency to that cluster. In this way, two identical frequency bands are not assigned to two adjacent clusters, and for this reason, inter-cluster interference is avoided as much as possible. $s(k_j, k_{f,near})$ means the 2D distance between the drone k_j and the drone $k_{f,near}$. $n(k_i, k_{f,near})$ means the number of drone k_j's users who experience interference greater than $\hat{g}_0$ of drone $k_{f,near}$.

3) Height Adjustment: In this section, the height of drones is optimized to improve the SINR of users. Users served by other drones in band f are denoted by $\mathcal{M}^c$. $s^{min}_{k,m}$ is the shortest 2D distance between the drone k and users in the set $\mathcal{M}^c$ and is defined as the following equation [10]:

$$s^{min}_{k,m} = \min_{m\in\mathcal{M}^c} s_{k,m}. \tag{49}$$

Therefore, the height of the drone k is obtained from the following 3 methods:

Algorithm 2 Allocation of Frequency Bands

Input: $\mathcal{K}, \mathcal{M}, \mathcal{F}$

Output: $\{\mathcal{K}_f | f \in \mathcal{F}\}$

Initialization: $\mathcal{K}_{un} = \mathcal{K}, \mathcal{K}_f = \emptyset, \forall f \in \mathcal{F}$.
Choose the UAV k_1 nearest to the center of the region.
$\mathcal{K}_1 \leftarrow k_1, \mathcal{K}_{un} = \mathcal{K}_{un} \setminus k_1$.
Choose $|\mathcal{F}|-1$ UAVs nearest to the UAV k_1, for example, UAV $k_2, \ldots, k_{|\mathcal{F}|}$.
$\mathcal{K}_2 \leftarrow k_2, \ldots, \mathcal{K}_F \leftarrow k_{|\mathcal{F}|}$,
$\mathcal{K}_{un} = \mathcal{K}_{un} \setminus \{k_2, \ldots, k_{|\mathcal{F}|}\}$.
for $j = |\mathcal{F}| + 1 : |\mathcal{K}|$ **do**
  Choose the UAV $k_i \in \mathcal{K}_{un}$ nearest to UAV k_1, $\mathcal{K}_{near} = \emptyset$.
  for $f = 1 : j - 1$ **do**
    Choose the UAV $k_{f,near} \in \mathcal{K}_f$ nearest to UAV k_i, $\mathcal{K}_{near} \leftarrow k_{f,near}$, calculate $s(k_i, k_{f,near})$, $n(k_i, k_{f,near})$.
  end for
  for $j = |\mathcal{F}| + 1 : |\mathcal{M}|$ **do**
    Temp=find $s(k_i, k_{f,near}) == \mathcal{F}(f)$
    if Temp $\sim$=1 **then**
      $s(Temp, 1) = \min_{f \in \mathcal{F}}(s(Temp, 1))$
    end if
  end for
  Choose UAV k_{f_1} and m_{f_2} from $\mathcal{K}_{near}$, where
  UAV k_{f_1} satisfies $s(k_i, k_{f_1}) = \max_{f \in \mathcal{F}} s(k, k_{f,near})$,
  UAV k_{f_2} satisfies $n(k_i, k_{f_2}) = \min_{f \in \mathcal{F}} n(k, k_{f,near})$.
  If the UAV k_{f_2} is not unique, select the UAV with higher $s(k_i, k_{f_2})$.
  if $n(k_i, k_{f_1}) = 0$ **then**
    $\mathcal{K}_{f_1} \leftarrow k_i$
  else
    $\mathcal{K}_{f_2} \leftarrow k_i$
  end if
  $\mathcal{K}_{un} = \mathcal{K}_{un} \setminus k_j, k_1 = k_j$.
end for
return $\{\mathcal{M}_f | f \in \mathcal{F}\}$

1) If $s_{k,m}^{min} > R_I$
The altitude h_k is obtained by the equation (50):

$$h_k = \begin{cases} h_{min}, & r_{min}^k \tan\theta_{k,u}^* < h_{min}, \\ r_{min}^k \tan\theta_{k,u}^*, & h_{min} < r_{min}^k \tan\theta_{k,u}^* < h_{max}, \\ h_{max}, & r_{min}^k \tan\theta_{k,u}^* > h_{max}, \end{cases} \tag{50}$$

2) If $r_{min}^k < s_{k,m}^{min} < R_I$
The interference radius R_I is obtained by the equation (51):

$$R_I = s_{k,m}^{min} - \epsilon, \tag{51}$$

ϵ represents a small positive number.
To satisfy the interference power equation of the user from the UAV, the altitude h_k is obtained by the equation (52):

$$\bar{g}(R_I, h_k) = \hat{g}_0, \tag{52}$$

To ensure that users are covered by drones, the height h_k is obtained by the equation (53):

$$\bar{g}(R_I, h_k) = \hat{g}_0. \tag{53}$$

3) If $s_{k,m}^{min} < r_{min}^k$
If $s_{k,m} = s_{k,m}^{min}$, by adjusting the radius of the drone, the user of m will not be covered by the drone. After user m is removed from $\mathcal{M}^c$, $\mathcal{M}'^c = \mathcal{M}^c \setminus m$, $m \in \mathcal{M}^c$ is obtained, where every user m satisfies $s_{k,m} > r_{min}^k$. Then, $s_{k,m}^{'min}$ is obtained as the minimum 2D distance between the drone k and the users in $\mathcal{M}'^c$, which is $r_{min}^k < s_{k,m}^{'min} < R_I$. So, the parameters R_I, h_k, and r_k are obtained by equations (51), (52), and (53), respectively.

TABLE 1. Simulation parameters.

Parameter	Physical Meaning	Value
x_{min}, y_{min}	The lower limit of the test area	0 m
h_{min}	Minimum height of UAVs	100 m
h_{max}	Maximum height of UAVs	500 m
$\mathcal{S}_{max}$	Serviceability	8
L_{max}	Maximum learning iteration	45
T_M	Maximum iteration time	800
P_S	Population size	500
τ	Attenuation loss parameter	0.01 [31]
ξ_0	Path loss parameter	7×10^{-5} [31]
$\bar{g}_0$	Channel gain threshold	-100 dB
N_0	SINR threshold	2
σ^2	Gaussian white noise signal power	-110 dBm [35]
P_T	Transmitted signal power	30 dBW
ϵ	Small positive number	1

IV. SIMULATION RESULTS

In this section, the simulation results are presented and then analyzed. The communication system consists of fixed-ground users and drones, with 200 users randomly distributed in a complex urban environment. The environment parameters are $\alpha = 11.95$, $\beta = 0.14$, and the modeling parameter of path loss is $\nu = 2$ [35]. This article was simulated in a MATLAB environment. Algorithms are compared to extract the best algorithm based on the number of drones for user coverage, drone coverage rate, processing time, received signal power before and after optimization, and interference signal power before and after optimization of the communication system. Table 1 shows other simulation parameters that were used to simulate the article.

In section (III-A), KKT conditions are used to obtain R_{max} and $\theta_{k,u}^*$. The optimal R_{max} and $\theta_{k,u}^*$ can be obtained according to three theorems 1, 2, and 3. Therefore, according to the conditions of KKT, the conditions (24a) to (24i) must be satisfied. According to Figure 4, in theorem 1, all the conditions of KKT are satisfied, and the solution of all parameters is obtained from this proposition.

Therefore, $R_{max} = 578$ meters, $\theta_{k,u}^* = 0.69$, and $h_k^* = 477$ meters have been obtained.

```
================================================================
CaseA       = 1 ----> **** KKT condition is satisfied ****
teta_star_opt  = 0.68561
r_ser_opt      = 577.6062
h_opt          = 472.4757
Lambda_1       = 202162183.5169
Lambda_2       = 0
Lambda_3       = 0

================================================================
CaseA     = 2 ----> KKT condition is not satisfied
teta_star = 1
r_ser     = 321.0463
h_opt     = 500
Lambda_1  = 57851568.7003
Lambda_2  = -3.0166e-05
Lambda_3  = 0

================================================================
CaseA     = 3 ----> KKT condition is not satisfied
teta_star = 1
r_ser     = 64.2093
h_opt     = 100
Lambda_1  = 462812.5496
Lambda_2  = 0
Lambda_3  = 3.0166e-05

================================================================
The KKT conditions are not satisfied for cases 2 and 3, so they are removed,
but the KKT conditions are satisfied for case 1, so all the parameters are selected from
case 1
================================================================
r_ser_opt  = 578
teta_star  = 0.69
h_opt      = 477
>>
```

FIGURE 4. Review of KKT conditions.

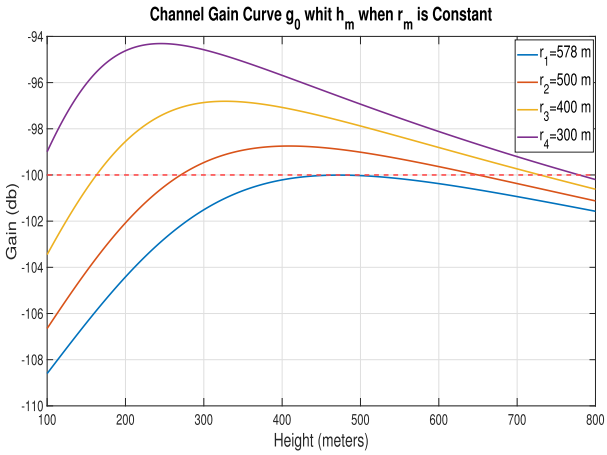


FIGURE 5. Channel gain curve $\hat{g}_0$ with h_k when r_k is constant.

Figure 5 shows the channel gain curve as a function of altitude when the radius is constant.

Figure 6 shows the optimal altitude according to the coverage radius of each UAV. Here, the goal is to have the lowest path loss according to the height of the UAVs, that is, to obtain the height of the UAVs in a way that has the lowest path loss.

It can be seen in Figure 6 that when the height of the UAV increases for a certain radius, the path loss first decreases and then increases.

In this article, the ABC-based placement (OAP) algorithm [10], the ACO-based placement (OACP) algorithm, the GA-based placement (OGP) algorithm, and the HS-based placement (OHP) algorithm are selected, and by comparing them, the best suitable algorithm in the communication system is introduced.

The OAP algorithm uses the ABC meta-heuristic algorithm for clustering users in the communication system, which is

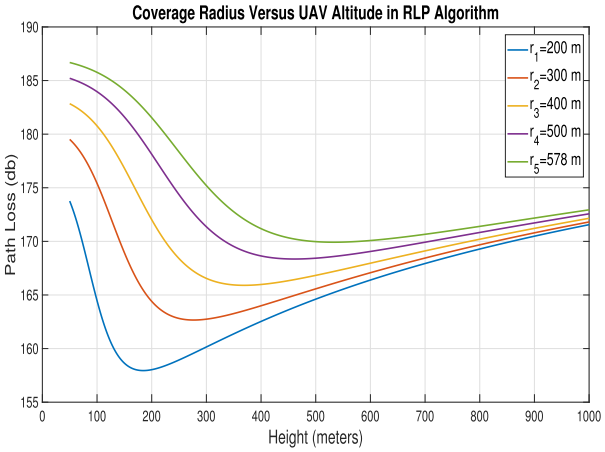


FIGURE 6. Coverage radius versus UAV altitude.

an ordered algorithm for UAV deployment. In this algorithm, it is $T_M = 800$ and $P_S = 500$.

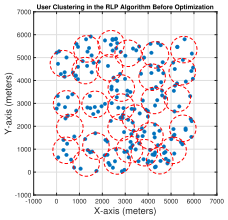
The OGP algorithm works similarly to the OAP algorithm to find the user k_0, but it uses the genetic algorithm (GA) algorithm to cluster the users in the communication system. k_0 is the user of the boundary area and has the largest distance from the center of the area that is not covered by the UAV. In this algorithm, it is $T_M = 800$ and $P_S = 500$.

The OACP algorithm works similarly to the OAP algorithm to find the user k_0, but it uses the ant colony optimization (ACO) algorithm to cluster the users in the communication system. The OACP algorithm uses the traveling salesman problem (TSP) in the ACO algorithm. It specifies a path with boundary points, and by moving these points, the shape of the path can be changed. Therefore, the path can be determined using these points. According to the center point obtained in each iteration, the largest path can be drawn around this center point. The maximum radius of the distance from the center point to the edge points should not be more than $2R_{max}$, because later it can make its circle tangent to its center point and get a new center. In this algorithm, it is $T_M = 800$ and $P_S = 500$.

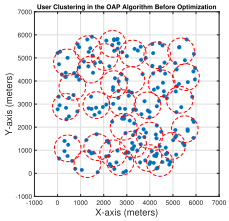
The OHP algorithm works similarly to the OAP algorithm to find the user k_0, but it uses the harmony search (HS) algorithm to cluster the users in the communication system. In this algorithm, the maximum iteration times, size of the memory harmony, and number of new harmonies of the HS algorithm are 800, 500, and 200, respectively. 200 users are randomly distributed in a region of 6×6 km. The serviceability of every UAV is 8, as well as the number of bands used in the communication system of 8 frequency bands.

Figure 7 shows the clustering of users before the optimization by each of the algorithms. In user clustering before optimization in the communication system, the service radius or coverage radius of the UAVs is the same for clustering all users in different groups and is equal to $R_{max} = 578$ meters.

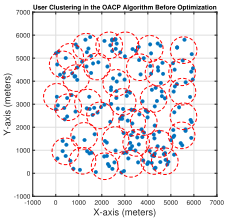
Figure 8 shows the clustering of users after the optimization by each of the algorithms. In this section, in addition to



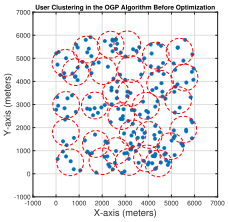
(a) clustering of users in the RLP algorithm before optimization.



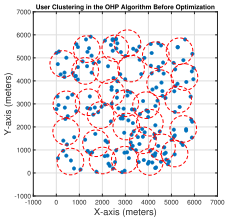
(b) clustering of users in the OAP algorithm before optimization.



(c) clustering of users in the OACP algorithm before optimization.

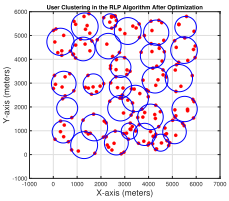


(d) clustering of users in the OGP algorithm before optimization.

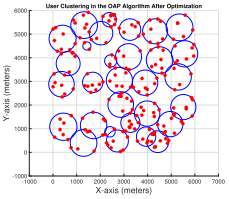


(e) clustering of users in the OHP algorithm before optimization.

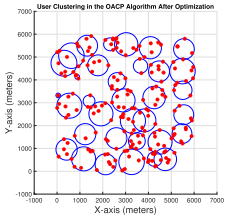
FIGURE 7. Clustering of users in each algorithm before optimization.



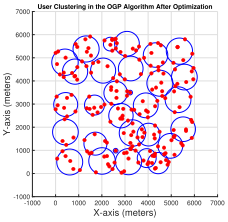
(a) User clustering in RLP algorithm after optimization.



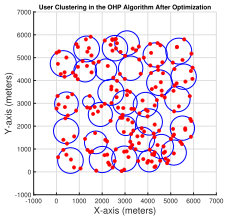
(b) User clustering in OAP algorithm after optimization.



(c) User clustering in OACP algorithm after optimization.



(d) User clustering in OGP algorithm after optimization.



(e) User clustering in OHP algorithm after optimization.

FIGURE 8. Clustering of users in each algorithm after optimization.

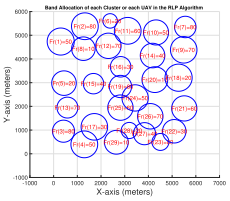
optimizing the 2D position of the drones, the service radius of drones is updated by optimizing with the CVX toolbox.

The frequency band allocation for each UAV in each algorithm is shown in Figure 9. According to the figure, it is not possible to allocate two identical frequency bands in two adjacent clusters, so the interference in the communication system is less.

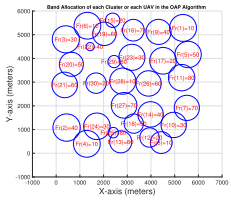
The results of user clustering in the 3D deployment of drones by algorithms are shown in Figure 10. According to Figure 10, the RLP algorithm uses fewer drones to cover or serve all ground users than other algorithms.

Figure 11 presents the number of drones used in different algorithms with different test areas, different serviceability, and different numbers of users. The Monte Carlo method has been used to distribute users, and each point in the figure is a representation of 100 different user distributions. Algorithm RLP has used less number of drones to cover ground users than other algorithms, and also algorithm OACP has used more drones to cover all users than other algorithms.

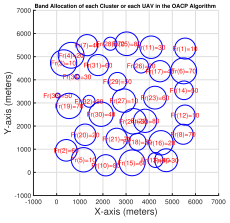
Figure 11(a) compares the number of drones used in the different algorithms based on different test areas. According to the figure, as the area of the test area increases, more



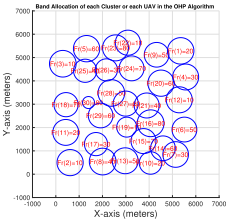
(a) Frequency band allocation of each cluster in the RLP algorithm.



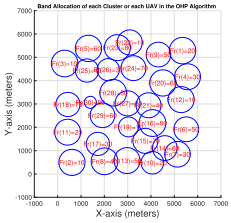
(b) Frequency band allocation of each cluster in the OAP algorithm.



(c) Frequency band allocation of each cluster in the OACP algorithm.

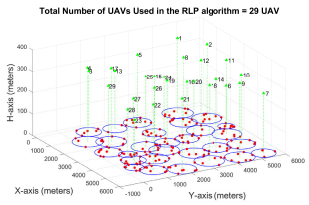


(d) Frequency band allocation of each cluster in the OGP algorithm.

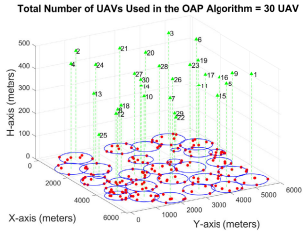


(e) Frequency band allocation of each cluster in the OHP algorithm.

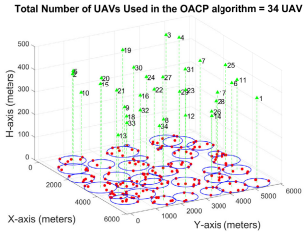
FIGURE 9. Frequency band allocation of each cluster in each algorithm.



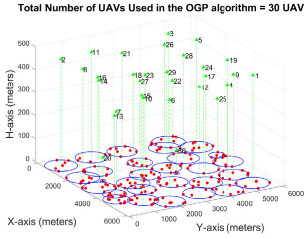
(a) Number of UAVs used in the RLP algorithm: 29



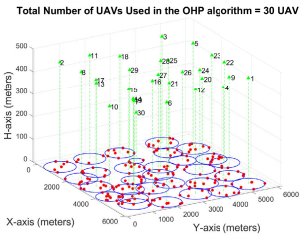
(b) Number of UAVs used in the OAP algorithm: 30



(c) Number of UAVs used in the OACP algorithm: 34



(d) Number of UAVs used in the OGP algorithm: 31



(e) Number of UAVs used in the OHP algorithm: 30

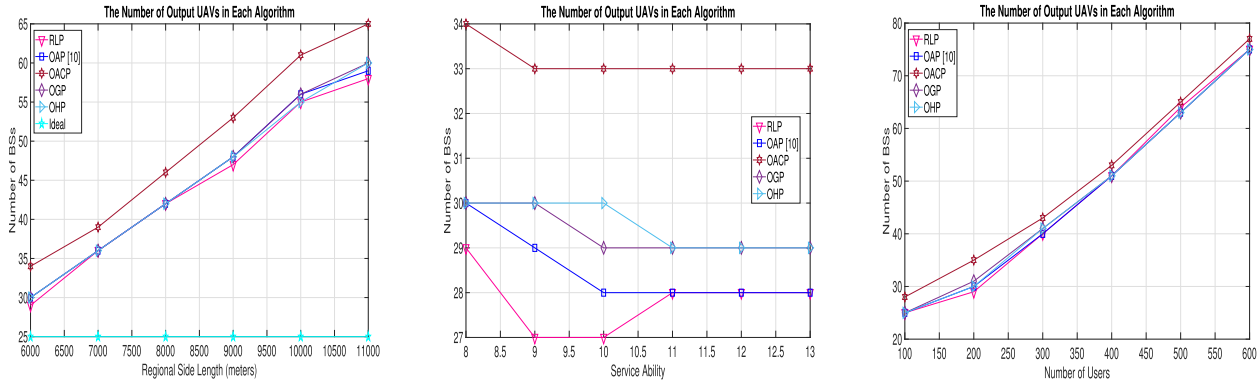
FIGURE 10. 3D deployment of UAVs in each algorithm.

drones are needed to cover all users. The ideal solution is to assume that the Service radius or coverage radius of drones is infinite, and the number of drones is defined as $\frac{|\mathcal{U}|}{\mathcal{S}_{max}}$. According to the figure, when the area of the test area is small, the RLP algorithm is closer to the ideal solution than other algorithms. However, when the area of the test area increases, in terms of solving the clustering problem, the distance between the RLP algorithm and the ideal solution increases, and the RLP algorithm uses more UAVs to cover all users in the communication system.

Figure 11(b) compares the number of drones used in the different algorithms based on different serviceability. As the serviceability of drones increases, fewer drones are needed to cover all users. According to the figure, with the increase in the serviceability of the RLP algorithm compared to other algorithms, it has used a smaller number of drones to cover all users.

Figure 11(c) compares the number of drones used in the different algorithms based on different users. The figure indicates that as the number of users increases, more drones are required to cover all users.

According to the simulation results in Figure 11, regardless of whether different test areas, different serviceability, or different numbers of users, the RLP algorithm has used

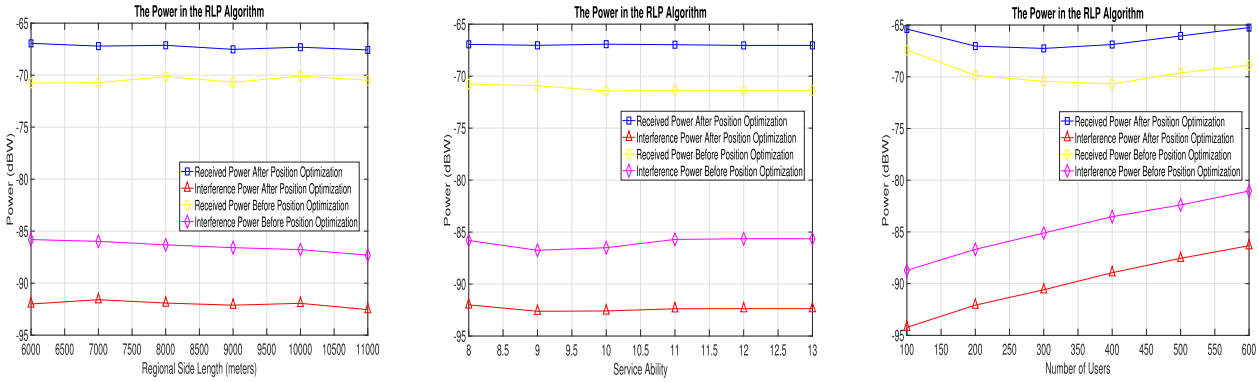


(a) Comparison of the number of drones used based on the different test areas.

(b) Comparison of the number of drones used based on the different serviceability.

(c) Comparison of the number of drones used based on the different users.

FIGURE 11. Comparison based on the number of drones used in the communication system under different conditions.



(a) Comparison of power based on the different test areas.

(b) Comparison of power based on the different serviceability.

(c) Comparison of power based on the different users.

FIGURE 12. Comparison based on the received signal power and the interference signal power before and after optimization in the RLP algorithm.

a smaller number of drones to cover or serve all users. Therefore, the RLP algorithm has performed better in terms of the minimum drone used in the communication system.

Figure 12 presents a comparison based on received signal power and interference signal power before optimization and after optimization in the RLP algorithm for $P_T = 30$ dBW, with different test areas, different serviceability, and different number of users.

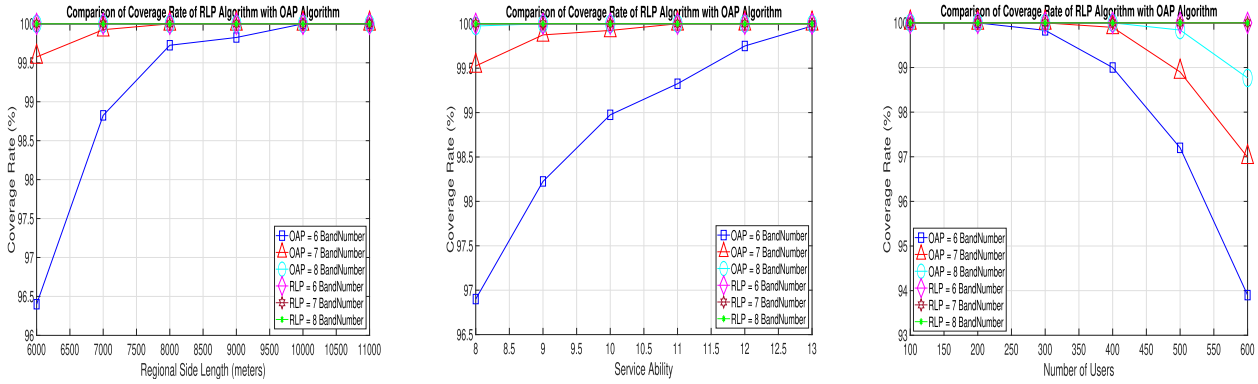
Figure 12(a) presents the comparison of the received signal power and the interference signal power in the RLP algorithm based on the different test areas, before and after optimizing the 3D locations of the drones and assigning frequency bands. According to the figure, with the increase of the area of the test area after optimization, the received signal power has increased, and the interference signal power has decreased. Hence, the figure presents that the RLP algorithm has had an acceptable result after optimization.

Figure 12(b) presents the comparison of the received signal power and the interference signal power in the RLP algorithm based on the different serviceability, before and after optimizing the 3D locations of the drones and assigning frequency bands. According to the figure, with the increase in serviceability after optimization, the received power has increased, and the interference power has decreased. Therefore, the figure presents that the RLP algorithm has had an acceptable result after optimization.

Figure 12(c) presents the comparison of the received signal power and the interference signal power in the RLP algorithm based on the different users, before and after optimizing the 3D locations of the drones and assigning frequency bands. In this part, users are randomly distributed in a region of 10×10 km.

According to Figure 12, with the increase in the users after optimization, the received power has increased, and the interference power has decreased. Therefore, the figure presents that the RLP algorithm has had an acceptable result after optimization.

Figure 13 presents the comparison of the coverage rate of the RLP algorithm with the OAP algorithm with different test areas, different serviceability, and different number of users.



(a) Comparison of the coverage rates based on the different test areas.

(b) Comparison of the coverage rates based on the different serviceability.

(c) Comparison of the coverage rates based on the different users.

FIGURE 13. Comparison of the UAV coverage rate in the RLP algorithm with the OAP algorithm in the communication system under different conditions.

In this figure, the coverage rate of each algorithm with the number of different frequency bands 6, 7, and 8 has been obtained, and then the coverage rate of 2 algorithms with the same number of frequency bands has been compared.

Figure 13(a) presents the comparison of the coverage rate of 2 algorithms based on the area of the different test areas. In this part, 400 users are randomly distributed in the region. According to the figure, with the increase in the area of the test areas, the density of users is less. Then, the interference of users in the communication system is reduced. Finally, the coverage rate of users is increased. According to the figure, the coverage rate of the RLP algorithm with the number of frequency bands 6, 7, and 8 is 100%. In the OAP algorithm, when the number of frequency bands is 8, the coverage rate is 100% in all test areas (cyan curve). When 7 frequency bands are used in the OAP algorithm, for example, with the distribution of 400 users in a 6 $\times$ 6 km area, the coverage rate of drones is 99.6%. But when 400 users are distributed in the 11 $\times$ 11 km area, the coverage rate has reached 100% (red curve). When 6 frequency bands are used in the OAP algorithm, for example, with the distribution of 400 users in a 6 $\times$ 6 km area, the coverage rate of drones is 96.4%. But when 400 users are distributed in the 11 $\times$ 11 km area, the coverage rate has reached 100% (blue curve).

Figure 13(b) presents the comparison of the coverage rate of 2 algorithms based on the different serviceability. In this part, 400 users are randomly distributed in a region of 6 $\times$ 6 km. According to the figure, with the increase in serviceability, fewer UAVs are used to cover ground users, and inter-cluster interference is reduced, thus increasing the user coverage rate. According to the figure, the coverage rate of the RLP algorithm with the number of frequency bands 6, 7, and 8 is 100%. In the OAP algorithm, when the number of frequency bands is 8, the coverage rate is 100% with any number of UAVs' serviceability (cyan curve). When 7 frequency bands are used in the OAP algorithm, for example, by distributing 400 users in an area of 6 $\times$ 6 km when the serviceability of each UAV is 8, the coverage rate of UAVs is 99.5%. But when 400 users are distributed in an area of 6 $\times$ 6 km when the service capability of each UAV is 11, the coverage rate has reached 100% (red curve). When 6 frequency bands are used in the OAP algorithm, for example, by distributing 400 users in an area of 6 $\times$ 6 km when the serviceability of each UAV is 8, the coverage rate of UAVs is 96.9%. But when 400 users are distributed in an area of 6 $\times$ 6 km when the service capability of each UAV is 11, the coverage rate has reached 100% (blue curve).

Figure 13(c) presents the comparison of the coverage rate of 2 algorithms based on the different users. In this part, 400 users are randomly distributed in a region of 8 $\times$ 8 km. According to the figure, with the increase in users, the inter-cluster interference is increasing because it is more possible for users to be present in these inter-cluster areas. Therefore, the coverage rate decreases, but by increasing the frequency bands, this problem can be solved and the coverage rate can be increased in the communication system. According to the figure, the coverage rate of the RLP algorithm with the number of frequency bands 6, 7, and 8 is 100%. In the OAP algorithm, for example, with the distribution of 100 users in an area of 8 $\times$ 8 km, the serviceability of each drone is 8, and with the number of frequency bands 8, the coverage rate is 100%. However with the distribution of 600 users in an area of 8 $\times$ 8 km, and the serviceability of each drone is 8, and the coverage rate has reached 98.8% (cyan curve).

According to the simulation results in Figure 13, regardless of whether different test areas, different serviceability, or the different number of users, it can be seen that the RLP algorithm performed better than OAP. The coverage rate of the RLP Algorithm is always 100%.

The processing time of each algorithm to reach the solution is shown in Figure 14. According to the figure, the processing time of the RLP algorithm is less than other algorithms and

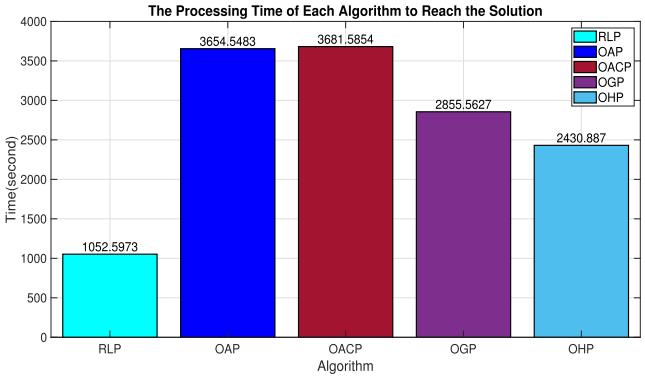


FIGURE 14. The processing time of each algorithm to reach the solution.

the processing time of the OACP algorithm is more than other algorithms.

V. CONCLUSION

In this article, when the users are fixed in the network, a solution is presented for the 3D deployment of multiple drones that are used as BSs to cover downlink users according to the maximum serviceability and maximum coverage radius of the drones. In the first stage, theoretically, the maximum value that the UAV coverage radius can have is obtained using the KKT conditions. In the second stage, the RL-based algorithm was used to cluster users in the communication system to use the least number of drones to cover or serve all the users so that the system is optimal in terms of implementation cost. In the third step, frequency bands were assigned to drones so that two identical frequency bands are not placed next to each other in two clusters and to prevent interference between clusters so that eventually the interference in the communication system reaches its lowest level. In the next step, the 3D locations of the drones are optimized by adjusting the height. According to the simulation results presented in section IV, the RLP algorithm has performed better than other algorithms in terms of the optimal number of drones, the power, the coverage rate of the ground users, and the processing time.

REFERENCES

[1] B. Li, Z. Fei, and Y. Zhang, "UAV communications for 5G and beyond: Recent advances and future trends," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 2241–2263, Apr. 2019.

[2] K. P. Valavanis and G. J. Vachtsevanos, *Handbook of Unmanned Aerial Vehicles*. Dordrecht, The Netherlands: Springer, 2015.

[3] Y. Zeng, R. Zhang, and T. J. Lim, "Wireless communications with unmanned aerial vehicles: Opportunities and challenges," *IEEE Commun. Mag.*, vol. 54, no. 5, pp. 36–42, May 2016.

[4] Z. Xiao, P. Xia, and X.-G. Xia, "Enabling UAV cellular with millimeter-wave communication: Potentials and approaches," *IEEE Commun. Mag.*, vol. 54, no. 5, pp. 66–73, May 2016.

[5] S. Sekander, H. Tabassum, and E. Hossain, "Multi-tier drone architecture for 5G/B5G cellular networks: Challenges, trends, and prospects," *IEEE Commun. Mag.*, vol. 56, no. 3, pp. 96–103, Mar. 2018.

[6] D. Liu, Y. Xu, J. Wang, J. Chen, Q. Wu, A. Anpalagan, K. Xu, and Y. Zhang, "Opportunistic utilization of dynamic multi-UAV in device-to-device communication networks," *IEEE Trans. Cognit. Commun. Netw.*, vol. 6, no. 3, pp. 1069–1083, Sep. 2020.

[7] L. Zhu, J. Zhang, Z. Xiao, X. Cao, X.-G. Xia, and R. Schober, "Millimeter-wave full-duplex UAV relay: Joint positioning, beamforming, and power control," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 9, pp. 2057–2073, Sep. 2020.

[8] Z. Ullah, F. Al-Turjman, and L. Mostarda, "Cognition in UAV-aided 5G and beyond communications: A survey," *IEEE Trans. Cognit. Commun. Netw.*, vol. 6, no. 3, pp. 872–891, Sep. 2020.

[9] L. P. Kaelbling, M. L. Littman, and A. W. Moore, "Reinforcement learning: A survey," *J. Artif. Intell. Res.*, vol. 4, no. 1, pp. 237–285, Jan. 1996.

[10] C. Zhang, L. Zhang, L. Zhu, T. Zhang, Z. Xiao, and X.-G. Xia, "3D deployment of multiple UAV-mounted base stations for UAV communications," *IEEE Trans. Commun.*, vol. 69, no. 4, pp. 2473–2488, Apr. 2021.

[11] Ç. Karahan and B. Canberk, "An intelligent 3D placement methodology for drone networks," in *Proc. IEEE 18th Annu. Consum. Commun. Netw. Conf. (CCNC)*, Jan. 2021, pp. 1–6.

[12] E. Kalantari, H. Yanikomeroglu, and A. Yongacoglu, "On the number and 3D placement of drone base stations in wireless cellular networks," in *Proc. IEEE 84th Veh. Technol. Conf. (VTC-Fall)*, Sep. 2016, pp. 1–6.

[13] X. Liu, Y. Liu, and Y. Chen, "Reinforcement learning in multiple-UAV networks: Deployment and movement design," *IEEE Trans. Veh. Technol.*, vol. 68, no. 8, pp. 8036–8049, Aug. 2019.

[14] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Efficient deployment of multiple unmanned aerial vehicles for optimal wireless coverage," *IEEE Commun. Lett.*, vol. 20, no. 8, pp. 1647–1650, Aug. 2016.

[15] J. Qin, Z. Wei, C. Qiu, and Z. Feng, "Edge-prior placement algorithm for UAV-mounted base stations," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2019, pp. 1–6.

[16] J. Lyu, Y. Zeng, R. Zhang, and T. J. Lim, "Placement optimization of UAV-mounted mobile base stations," *IEEE Commun. Lett.*, vol. 21, no. 3, pp. 604–607, Mar. 2017.

[17] I. Strumberger, N. Bacanin, S. Tomic, M. Beko, and M. Tuba, "Static drone placement by elephant herding optimization algorithm," in *Proc. 25th Telecommun. Forum (TELFOR)*, Nov. 2017, pp. 1–4.

[18] G.-G. Wang, S. Deb, and L. d. S. Coelho, "Elephant herding optimization," in *Proc. 3rd Int. Symp. Comput. Bus. Intell. (ISCBI)*, Dec. 2015, pp. 1–5.

[19] X. Luo, J. Xie, L. Xiong, Z. Wang, and C. Tian, "3-D deployment of multiple UAV-mounted mobile base stations for full coverage of IoT ground users with different QoS requirements," *IEEE Commun. Lett.*, vol. 26, no. 12, pp. 3009–3013, Dec. 2022.

[20] F. Pasandideh, F. E. R. Cesen, P. H. M. Pereira, C. E. Rothenberg, and E. P. de Freitas, "An improved particle swarm optimization algorithm for UAV base station placement," *Wireless Pers. Commun.*, vol. 130, no. 2, pp. 1343–1370, Mar. 2023.

[21] Z. Rahimi, R. Ghanbari, A. H. Mohajerzadeh, H. Ahmadi, and M. Sookhak, "3D UAV BS positioning and backhaul management in cellular network via stochastic optimization," in *Proc. IEEE Global Commun. Conf.*, Dec. 2022, pp. 2169–2175.

[22] X. Liu, X. Wang, M. Huang, J. Jia, N. Bartolini, Q. Li, and D. Zhao, "Deployment of UAV-BSs for on-demand full communication coverage," *Ad Hoc Netw.*, vol. 140, Mar. 2023, Art. no. 103047.

[23] R. Ozdag, "Multi-metric optimization with a new metaheuristic approach developed for 3D deployment of multiple drone-BSs," *Peer-Peer Netw. Appl.*, vol. 15, no. 3, pp. 1535–1561, Mar. 2022.

[24] A. Gupta, A. Trivedi, and B. Prasad, "B-GWO based multi-UAV deployment and power allocation in NOMA assisted wireless networks," *Wireless Netw.*, vol. 28, no. 7, pp. 3199–3211, Jul. 2022.

[25] X. Wang, Y. Liu, and S. Chang, "A distributed 3D UAV placement algorithm for integrated ground-air cellular networks," in *Proc. 8th Int. Conf. Big Data Comput. Commun. (BigCom)*, Aug. 2022, pp. 413–421.

[26] H. Guo, Y. Zou, T. Wu, and F. Zhou, "3D location optimization for UAV-aided uplink/downlink transmissions," *IEEE Trans. Veh. Technol.*, vol. 71, no. 4, pp. 4477–4482, Apr. 2022.

[27] A. M. Seid, G. O. Boateng, B. Mareri, G. Sun, and W. Jiang, "Multi-agent DRL for task offloading and resource allocation in multi-UAV enabled IoT edge network," *IEEE Trans. Netw. Service Manage.*, vol. 18, no. 4, pp. 4531–4547, Dec. 2021.

[28] S. Anokye, D. Ayepah-Mensah, A. M. Seid, G. O. Boateng, and G. Sun, "Deep reinforcement learning-based mobility-aware UAV content caching and placement in mobile edge networks," *IEEE Syst. J.*, vol. 16, no. 1, pp. 275–286, Mar. 2022.

[29] H. Zhao, H. Wang, W. Wu, and J. Wei, "Deployment algorithms for UAV airborne networks toward on-demand coverage," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 9, pp. 2015–2031, Sep. 2018.
[30] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proc. 5th Berkeley Symp. Math. Statist. Probab.*, vol. 1, 1967, pp. 281–297.
[31] Y. Zeng, Q. Wu, and R. Zhang, "Accessing from the sky: A tutorial on UAV communications for 5G and beyond," *Proc. IEEE*, vol. 107, no. 12, pp. 2327–2375, Dec. 2019.
[32] 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.
[33] J. T. Linderoth and M. W. P. Savelsbergh, "A computational study of search strategies for mixed integer programming," *INFORMS J. Comput.*, vol. 11, no. 2, pp. 173–187, May 1999.
[34] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[35] Y. Liu, K. Liu, J. Han, L. Zhu, Z. Xiao, and X.-G. Xia, "Resource allocation and 3-D placement for UAV-enabled energy-efficient IoT communications," *IEEE Internet Things J.*, vol. 8, no. 3, pp. 1322–1333, Feb. 2021.

SAHAR BAGHDADY received the bachelor's degree in the field of information engineering and communication technology (ICT), of mobile telecommunications from the Scientific and Applied Faculty of Posts and Telecommunications, in 2015, and the master's degree in the field of electrical engineering of telecommunication systems from Shahrood University of Technology, under the supervision of Dr. Seyed Masoud Mirrezaei, in 2023.

SEYED MASOUD MIRREZAEI received the B.Sc. degree in communication engineering from the K. N. Toosi University of Technology, in September 2004, and the M.Sc. and Ph.D. degrees in electrical engineering from the Amirkabir University of Technology, Tehran, Iran, in 2007 and 2013, respectively. He has been a Ph.D. Student and a Researcher with the Mobile and Wireless Networks Research Laboratory, Amirkabir University of Technology, under the supervision of Prof. Karim Faez. Also, he was a Visiting Research Student with the Signal Design and Analysis Laboratory (SDAL), Queen's University, Kingston, Canada, under the supervision of Prof. Shahram Yousefi, from March 2011 to February 2012. He has been with the Electrical Engineering Department, Shahrood University of Technology, since 2013. His research interests include communications, cloud systems, big data, networks, information theory, signal processing, channel coding, and network coding.

RASHID MIRZAVAND (Senior Member, IEEE) is currently an Assistant Professor with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada, where he leads the Intelligent Wireless Technology Group. He is also an Adjunct Professor with the Department of Mechanical Engineering, University of Alberta, and an Adjunct Fellow with the Faculty of Engineering and IT, University of Technology Sydney, Australia. He was the Co-Founder and the Chief Technology Officer of three companies in the smart sensor (SenZIoT), near-field measurement (Anteligen), and wireless power transfer technologies (WiDyne) with the University of Alberta. His work focuses on developing sensors and wireless sensor networks for smart health, home, and infrastructure applications. He also focuses on developing radio frequency systems, reconfigurable intelligent surfaces and antennas, and measurement systems. He is a Specialty Chief Editor of *Frontiers in the Internet of Things* (IoT Enabling Technologies Section).