

Article

# Prioritized User Association for Sum-Rate Maximization in UAV-Assisted Emergency Communication: A Reinforcement Learning Approach

Abdul Basit Siddiqui <sup>1</sup>, Iraj Aqeel <sup>1</sup>, Ahmed Alkhayyat <sup>2</sup>, Umer Javed <sup>1</sup> and Zeeshan Kaleem <sup>1,\*</sup>

<sup>1</sup> Department of Electrical and Computer Engineering, Wah Campus, COMSATS University Islamabad, Wah Cantt, Islamabad 47040, Pakistan; fa19-ree-010@cuiwah.edu.pk (A.B.S.); fa19-ree-009@cuiwah.edu.pk (I.A.); umerjaved@ciitwah.edu.pk (U.J.)

<sup>2</sup> College of Technical Engineering, The Islamic University, Najaf 54001, Iraq; ahmedalkhayyat85@iunajaf.edu.iq

\* Correspondence: zeeshankaleem@gmail.com

**Abstract:** Unmanned air vehicles (UAVs) used as aerial base stations (ABSs) can provide communication services in areas where cellular network is not functional due to a calamity. ABSs provide high coverage and high data rates to the user because of the advantage of a high altitude. ABSs can be static or mobile; they can adjust their position according to real-time location of ground user and maintain a good line-of-sight link with ground users. In this paper, a reinforcement learning framework is proposed to maximize the number of served users by optimizing the ABS 3D location and power. We also design a reward function that prioritize the emergency users to establish a connection with the ABS using Q-learning. Simulation results reveal that the proposed scheme clearly outperforms the baseline schemes.

**Keywords:** aerial base station; reinforcement learning; *k*-means clustering; line of sight; non line of sight



**Citation:** Siddiqui, A.B.; Aqeel, I.; Alkhayyat, A.; Javed, U.; Kaleem, Z. Prioritized User Association for Sum-Rate Maximization in UAV-Assisted Emergency Communication: A Reinforcement Learning Approach. *Drones* **2022**, *6*, 45. <https://doi.org/10.3390/drones6020045>

Academic Editors: Mohammed H. Alsharif and Muhammad Asghar Khan

Received: 3 January 2022

Accepted: 29 January 2022

Published: 15 February 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The third generation partnership project (3GPP) has finalized its 17th release to enable the development and deployment of the fifth-generation (5G) wireless systems. The 5G new radio (NR) implements a flexible physical layer to support mmWave communications and massive antenna systems. Service-oriented 5G network architecture has been presented that has the potential to enable functions according to service requirements. The major target of 5G (and beyond) communications systems is to meet the diverse requirements to enable low-latency, broadband communications, and massive machine-to-machine (M2M) communications [1]. To achieve this, unmanned air vehicles (UAVs)-enabled airborne communications have recently attracted researchers' and industrialists' attention [2].

The demand and popularity of UAVs, also known as aerial base stations (ABSs), has increased in the past few years. ABS can be rapidly deployed with a low cost to enable mobile communication services during disaster situations [2–4]. ABS can assist in establishing line-of-sight (LoS) communications that in turn reduce the effect of shadowing and fading [5]. Moreover, it can be used in the Internet of things (IoT) to assemble information from IoT devices on the ground, that are installed in a specific area where mobile infrastructure is not available. The main advantage of ABSs is that they can be deployed on target locations without having any infrastructure, whereas for terrestrial BSs, infrastructure is required to deploy any new BSs. ABSs can also be used as relays for improving link performance. They provide wireless connection between far users in various military applications to increase throughput of the system. Moreover, the deployment of UAVs can also assist in gathering necessary data from the faulty sensors deployed, to gather data from wireless sensor networks [6].

Moreover, Figure 1 shows that ABS can acts as a relay and can enhance services in any geographical areas. They can be connected with ground BSs for media and wireless connectivity. They can also be connected through satellite for basic connectivity with the core network. ABSs can provide basic services like calls and can also provide advanced features of video calling, data streaming, heavy file downloading and live gaming at good data rates. Sum-rate maximization will be key to provide advanced features to the users, as these features require high data rates and seamless connectivity.

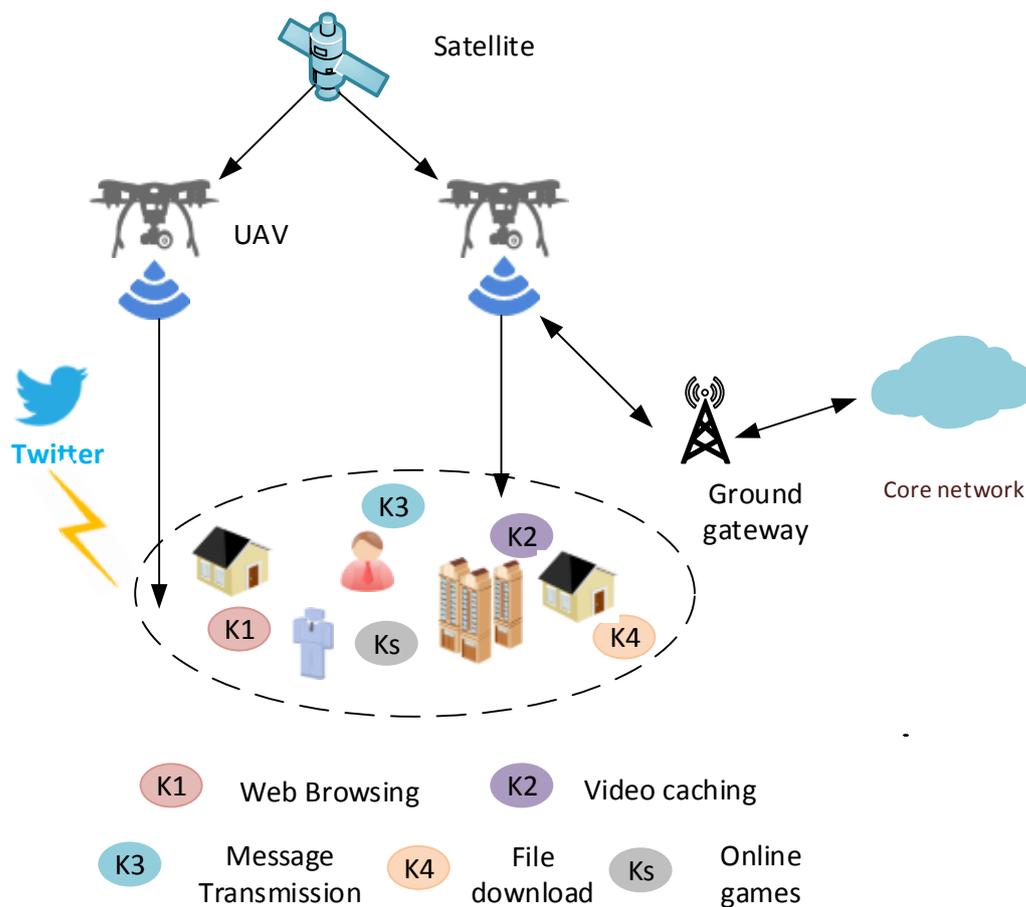


Figure 1. Deployment of multiple UAVs in wireless communications.

Besides the discussed advantages, ABS’s 3D deployment poses certain challenges to improve coverage, maximize user association, improve energy efficiency, and enhance sum rates. Recently, many schemes have been proposed to improve the coverage, maximizing user association, and maximizing users’ sum rates. For instance, in [7] the authors modeled the user association and ABS placement problem as a mixed-integer nonconvex optimization problem with the goal of maximizing users’ total achievable data rates. They divided the problem into two subproblems and proposed an iterative solution to solve the nonconvex problem. A comparison with the state of the art demonstrated that the developed algorithm was convergent and capable of producing good results. Similarly, in [8], the authors presented a placement algorithm for multiple ABSs with in-band wireless backhaul that jointly optimized resource allocation and user association. To simplify the optimization problem, an equivalent spectrum efficiency was defined given the optimal resource allocation. Simulation results were provided to validate the effectiveness of the proposed method. The solutions discussed above used traditional methods, such as game theory and convex optimization, that require additional accurate network information from the environment.

To solve these challenges, machine/deep-learning-based solutions have recently made remarkable progress over traditional ABS deployment schemes. In machine/deep-learning-based solutions, three major types of schemes exist such as supervised, unsupervised, and reinforcement learning (RL) schemes. Reinforcement learning scheme has proved to be more efficient in the implementation of resource allocation, ABS placement, and user-association as it can interact directly with the environment without requiring prior training. As a result, it has the potential to provide the best possible action to maximize the decided reward. In the literature, there exist several schemes that addressed the ABS placement and user association problem using RL schemes. For example, in [9], a smart user association algorithm named reinforcement learning handoff (RLH) was developed to reduce redundant handoffs in UAV networks, and two methods of UAV mobility control were designed to work in combination with the proposed RLH algorithm to achieve maximum system throughput. The RLH algorithm proved to reduce the number of handoffs by 75%. Hence, motivated by this, we adopted an RL-based solution to solve the challenge of ABS placement and user association to enable emergency communications.

This paper is organized as follows. We summarize the related work in Section 2 of the paper. The detailed system model is presented in Section 3. We formulate the problem in Section 4 and propose the prioritized user association scheme in Section 5. Simulation results are discussed in Section 6 and finally the paper is concluded in Section 7.

## 2. Related Work

In the literature, to maximize ABS coverage, an ABS altitude optimization scheme that considered maximum path loss was proposed in [5]. Similarly, an efficient 3D placement of ABSs was proposed in [10], that aimed at maximizing the coverage by using the minimum transmit power. They decoupled the ABS placement problem into horizontal and vertical dimensions. An optimal placement algorithm was proposed that proved the increase in coverage of users and significant savings in terms of transmit power.

A multiple UAVs deployment along with directional antennas was studied in [11], where the authors proved that ABS placement was a function of altitude and antenna gain. Moreover, they also proved that the ABS altitude must be properly adjusted in accordance with the beam width of the antenna to maximize coverage. Results revealed that optimum altitude and location were determined on the basis of the number of UAVs and the beam width of directional antennas. To deal with the problem of traffic congestion, a UAV-aided cellular communication network was presented in [12]. Here, UAVs used reinforcement learning methods to select relay policy for mobile users, when base station was heavily congested. Results showed the reduction in bit error rate and transmission energy.

The authors in [13] proposed a coverage enhancement solution for a single UAV, to provide wireless coverage for indoor users during a disaster situation. Their objective was to minimize the total transmit power with maximum path loss using a gradient descent algorithm. A method to optimize the height of a UAV to maximize coverage and minimize outage probability was presented in [14], where they adopted a decode and forward-relaying method.

In [15], the authors jointly optimized user association, trajectory, and power for each user to improve data offloading by observing QoS requirements. They decomposed the problem into subproblems and then iteratively solved each problem. The results showed a significant performance gain compared to existing schemes.

The authors in [16] presented mobile-edge-computing-assisted UAV communications to enable emergency communications. They addressed the key challenges of user association and resource allocation for energy efficient UAV deployment at edge. To solve this challenge, they adopted RL-assisted resource allocation and user association, which in turn proved beneficial compared to existing solutions.

To improve the throughput in overloaded and outage situations, an RL-based ABS 3D deployment approach was proposed in [17], where UAVs found their optimal location to increase system performance gain. Similarly, 3D ABS deployment was presented in [18] to

maximize users' coverage by finding ABS optimal altitude and location using a bisection search algorithm.

Throughput maximization by adjusting the position of UAV in software-defined disaster areas was presented in [19]. The results showed that a 26% throughput improvement was achieved by optimizing the ABS altitude. In [20], the joint optimization of ABS 3D placement and path loss compensation factor was presented to obtain the ABS maximal coverage. The results improved the coverage by consuming less energy.

The deployment of UAVs equipped with intelligent reflecting surface (IRS) was presented in [21] to maximize the sum rate by optimizing the base stations' power allocation, phase shift of the IRS, and horizontal position of UAV using deep reinforcement Learning. The ABS efficient placement as a relay node was studied in [22] with the objective to maximize throughput. They firstly adopted a particle swarm optimization algorithm to identify the optimum ABS location, and then three different approaches were adopted to maximize throughput that involved equal power allocation, water filling, and modified water filling. The results showed that the water filling method gave better results as compared to the other two methods.

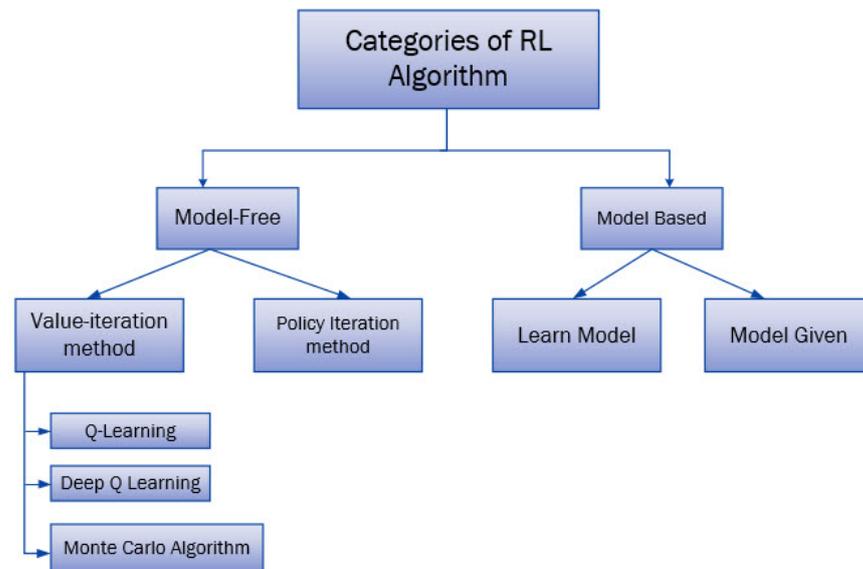
The deployment of a UAV by optimizing its trajectory to maximize the mean opinion score using a deep Q-learning method was presented in [23]. Similarly, the data rate maximization of an ABS-assisted downlink cellular system using RL was presented in [24]. Here, the Q-learning technique was used to optimize the ABS location, where simulation results revealed that RL performed better than a  $k$ -means algorithm to find the optimal ABS positions. Similarly, in [25], a Q-learning approach was proposed to solve the resource allocation challenge by considering the user's fairness and several other quality-of-service (QoS) constraints. We also summarize some of the key papers with their contributions in Table 1.

As the deployment of ABS is a challenging task considering that it should provide a proper coverage to the ground user, maximize user association, maximize sum rate as well as optimize power consumption. To optimize the ABS deployment, the 2D location as well as the altitude of the ABS to cover a maximum number of ground users are considered important factors during deployment. Another challenging task is the deployment of multiple ABSs compared to a single ABS, because of increased cochannel interference from using the same frequencies to improve the spectral efficiency. It is difficult to deploy an ABS in such a way that interference between the ABS to the ground BS is minimum. The position of the ABS should be flexible to cover more users with less interference.

Recently many conventional schemes have been proposed to improve the coverage, maximize user association, and maximize sum rates. However, recently, machine learning has made remarkable progress and RL is considered to be most effective in tackling ABS-assisted communication challenges. RL is the machine learning technique that interacts with the system and optimize its performance by finding the best possible action to maximize the given reward. RL gets the raw data and uses a trial-and-error method to detect the errors at the output, and then feeds it back to the system to enhance the system efficiency. Moreover, the RL is usually categorized into two main categories: model-free and model-based approaches as shown in Figure 2.

**Table 1.** Summary of related papers.

Paper	Problem Statement	Technique/Scheme Used	Improvement Observed
[5]	An optimal ABS platform to maximize coverage by finding the optimum altitude	Analytical approach	Maximum coverage was achieved at optimal low altitude
[10]	A 3D placement of UAV-BS by decoupling vertical from horizontal dimensions to maximize the coverage of users by minimizing the transmit power	Optimal placement algorithm	Savings in transmit power and maximized coverage were achieved
[11]	Deployment of multiple UAVs having directional antennas and optimization of the altitude of UAVs to maximize the coverage area and lifetime of UAVs	Circle-packing theory	The optimum altitude can be obtained on the basis of the number of UAVs and beam width of directional antennas
[12]	UAV-aided cellular communication network against jamming	Reinforcement learning	A minimized bit error rate and energy saving for the cellular network
[13]	Single UAV to provide wireless coverage for indoor users when cellular network goes down	Gradient descent algorithm	A minimum transmit power with maximum path loss was obtained
[14]	Optimizing the height of a UAV to maximize coverage and minimizing outage probability	Decode and forward-relaying method	Maximum coverage with minimum outage was obtained by finding the optimum height of a UAV
[17]	A 3D deployment of UAV to improve throughput in overloaded and outage situations	Reinforcement learning	A maximum performance gain in terms of throughput was achieved
[18]	A 3D deployment of UAV to maximize revenue of the network	Bisection search algorithm	The maximized revenue of the network was achieved
[19]	Throughput maximization by adjusting the position of a UAV in software-defined disaster areas	Centralized algorithm	The throughput was improved by 26%
[26]	Channel model of backhaul and delay-aware was taken into account to minimize delay by finding the optimum height of a UBS	Backhaul and delay-aware positioning of UBS (BaDPU) algorithm	It was observed that the delay was less for low arrival rates and increased for high arrival rates
[27]	Optimal UAV placement to maximize the sum rate by using a minimum transmit power	Genetic algorithm	The optimal placement of UAV was achieved with minimum transmit power and minimum path loss
[21]	Deployment of a UAV equipped with intelligent reflecting surface (IRS) to maximize the sum rate by optimizing the power allocation of a base station (BS), phase shift of the intelligent reflecting surface (IRS), and horizontal position of the UAV	Deep reinforcement learning	An enhanced sum rate was obtained
[22]	Efficient placement of a UAV-BS serving as a relay node to maximize throughput	Equal power allocation method, water filling method and modified water filling method	Results showed that water filling method gives better results as compared to other two methods
[23]	Deployment of a UAV by optimizing its trajectory to maximize the mean opinion score (MOS)	Deep Q-learning	The maximized mean opinion score (MOS) was achieved



**Figure 2.** RL categories; model-free and model-based approaches.

Hence, RL was adopted to solve the problems of maximizing user association and sum rate. RL is the best option to optimize location of the ABS according to the environment and users' density. In RL, the agent (ABS) will optimize its location according to the requirements.

### 2.1. Contributions

Most of the presented solutions in the existing literature mostly considered a conventional cellular network scenario where the challenges are different compared to an emergency communication scenario. Thus, their presented solutions are not valid for emergency communication scenarios without proper modification. Motivated by this, in this paper, we propose a prioritized user association solution that adopts a Q-learning algorithm and prioritize the emergency users by using the proposed reward function, that in turn significantly enhances the sum rate and reduces the outage. Moreover, we also compare the impact of varying the number of deployed ABSs on the system sum rate and outage probability.

### 2.2. Reproducible Research

The simulation results can be reproduced by updating part of the code available at: <https://github.com/ZeeshanKaleem/Unmanned-Air-Vehicles-UAV-Simulator-for-Placement-and-Power-Allocation> accessed on 31 January 2022.

## 3. System Model

In this paper, we considered  $M - 1$  multiple ABSs accompanied by a single antenna and one ground base station (GBS) ( $m = 0$ ) in a downlink heterogeneous network model, with the total base stations represented by the set  $\mathcal{M}$ . We randomly distributed the total  $U$  of users inside the coverage area where the total users are represented by the set  $\mathcal{U}$ . The ABSs were deployed with the GBS where both used the same frequency, which may cause intercell interference. Resources were allocated orthogonally within the cell to avoid intracell interference. The coordinates for the users are represented as  $(x_u, y_u)$  and the 3D coordinates of the ABSs are represented as  $(x_m, y_m, h_m)$  for the base stations, as shown in Figure 3. However, to model the emergency communication scenario, we considered that the GBS did not exist and the remaining users were connected with the ABSs based on their priority. The key notations and symbols utilized in the paper are summarized in Table 2.

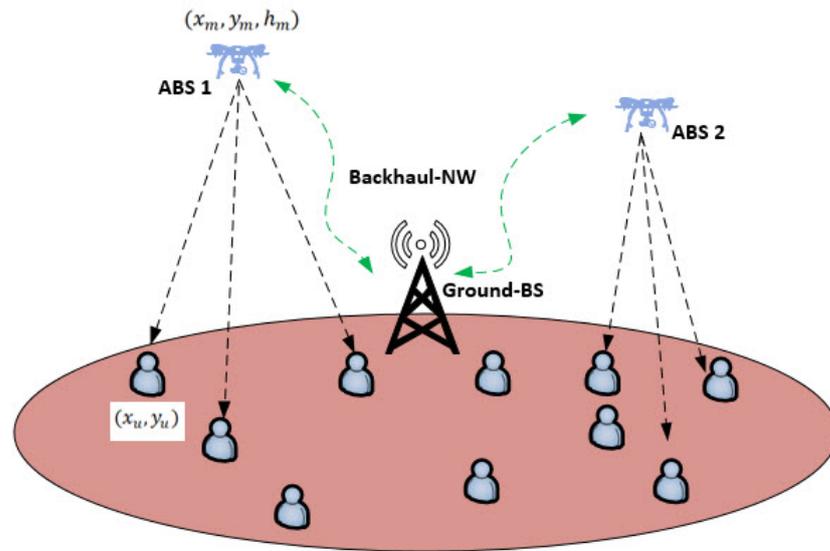


Figure 3. System model.

Table 2. Notations and description.

Notations	Description
ABS	Aerial base station
GBS	Ground base station
N	Number of users connected with ABS
$(x_m, y_m, h_m)$	ABS coordinates
$(x_u, y_u)$	Users coordinates
$P_m$	ABS transmit power
$\theta_u$	Elevation angle
$a, b$	Environmental parameters
$f_c$	Carrier frequency
$d$	Distance between ABS and ground user
$\eta_{LoS}$	Mean additional loss for LoS
$\eta_{NLoS}$	Mean additional loss for NLoS
$\gamma_{th}$	SINR threshold
$r_t, s_t, a_t$	Reward, state, action at time $t$
$\hat{s}_t, \hat{a}_t$	Next state and action
$\beta$	Discount factor
$\alpha$	Learning rate

The channel model from ABS to ground user communication can be line-of-sight (LoS) or non-line-of-sight (NLoS). LoS communication depends on factors such as building density, users' location, and placement of the ABS as well as the angle between the ABS and the user. The LoS channel is modeled as Rician fading whereas the NLoS is modeled as Rayleigh fading. Therefore, the channel distribution of the channel fading,  $l$  is given by:

$$f_l(l) = \begin{cases} f_{LoS}(l) & \text{for LoS case} \\ f_{NLoS}(l) & \text{for NLoS case} \end{cases}$$

where  $f_{NLoS}(l)$  and  $f_{LoS}(l)$  follow a noncentral Chi-squared distribution and exponential distribution, respectively, and are given by [28]

$$\begin{aligned}
f_{\text{LoS}}(l) &= \frac{1 + K(\theta_u)}{\overline{H}_{\text{LoS}}} \exp\left(-K(\theta_u) - \frac{1 + K(\theta_u)}{\overline{H}_{\text{LoS}}} l\right) \\
&\quad \times I_0\left(2\sqrt{\frac{K(\theta_u)(1 + K(\theta_u))}{\overline{H}_{\text{LoS}}} l}\right) \\
&= \frac{1}{2} \exp\left(-K(\theta_u) - \frac{l}{2}\right) I_0\left(\sqrt{2K(\theta_u)l}\right) \\
f_{\text{NLoS}}(l) &= \frac{1}{\overline{H}_{\text{NLoS}}} \exp\left(-\frac{l}{\overline{H}_{\text{NLoS}}}\right) = \exp(-l).
\end{aligned}$$

where  $I_0(\cdot)$  is the first kind with zero order Bessel function, and  $\overline{H}_{\text{LoS}} = 2 + 2K(\theta_u)$  and  $\overline{H}_{\text{NLoS}} = 1$  are the mean of the LoS and NLoS fading gains, respectively.

The probability of LoS is defined as:

$$P_{\text{LoS}} = \frac{1}{1 + e^{-b(\theta_u - a)}} \quad (1)$$

Here,  $a$  and  $b$  are environmental constant,  $\theta_u$  is the elevation angle that depends on the ABS height and also on the distance between users and ABSs. From the above equation, we can notice that by increasing  $\theta_u$ , the LoS probability increases, where:

$$\theta_u = \frac{-180}{\pi} \tan^{-1}\left(\frac{h_m}{r}\right) \quad (2)$$

The horizontal distance  $d_{u,m}$  between the  $m$ th ABS and user  $u$  is calculated as

$$d_{u,m} = \sqrt{(x_u - x_m)^2 + (y_u - y_m)^2} \quad (3)$$

where  $(x_u, y_u)$ ,  $(x_m, y_m)$  are the locations of a user and the ABS, respectively. By using the probability of LoS, the probability of NLoS communications can be calculated as

$$P_{\text{NLoS}} = 1 - P_{\text{LoS}} \quad (4)$$

The LoS path loss for a connected user is given as

$$L_{\text{LoS}} = 20 \log \frac{4\pi f_c d_{u,m}}{c} + \eta_{\text{LoS}} \quad (5)$$

Furthermore, due to buildings and tree, the LoS could disturb the communication, and the loss for a reflected signal with an NLoS component is calculated as:

$$L_{\text{NLoS}} = 20 \log \frac{4\pi f_c d_{u,m}}{c} + \eta_{\text{NLoS}} \quad (6)$$

Here,  $f_c$  is the carrier frequency,  $c$  is the speed of light,  $\eta_{\text{LoS}}$  means an additional loss for the LoS, and  $\eta_{\text{NLoS}}$  is the mean additional loss for the NLoS. As a result, the probabilistic mean path loss is defined as:

$$L = L_{\text{LoS}} \times P_{\text{LoS}} + L_{\text{NLoS}} \times P_{\text{NLoS}} \quad (7)$$

Let us assume  $A = \eta_{\text{LoS}} - \eta_{\text{NLoS}}$ ,  $B = 20 \log \frac{4\pi f_c}{c} + \eta_{\text{NLoS}}$ , and  $h_m^2 + d_{u,m}^2 = \left(\frac{d_{u,m}}{\cos(\theta_u)}\right)^2$ .

By using Equations (5) and (6) in Equation (7) we have:

$$L = \frac{A}{1 + ae^{(-b\frac{180}{\pi}\theta_u - a)}} + 20 \log\left(\frac{d_{u,m}}{\cos(\theta_u)}\right) + B \quad (8)$$

We observe from the above equation that the path loss changes by changing both the height and location of the ABS or by varying either one of them. Hence, the data rate between user  $u$  connected to the ABS is calculated as:

$$C_u = \log\left(1 + \frac{P_m G_u L}{I + \sigma^2}\right) \quad (9)$$

where  $P_m$  is the ABS power,  $I$  is the cochannel interference,  $G_u$  is the small-scale fading channel gain, and  $\sigma^2$  is the noise variance.

We analyzed the outage probability considering the interference and the ABS main links, and it is given by:

$$p_0 = \mathbb{P}[\gamma(\theta_u) < \gamma_{\text{th}}] \quad (10)$$

where  $\gamma_{\text{th}}$  is the target SNR, defined as  $\gamma_{\text{th}} = 2^{\frac{C_{\text{min}}}{B}} - 1$  for the target minimum required rate  $C_{\text{min}}$  and the bandwidth  $B$ .

#### 4. Problem Formulation

In this paper, our objective is to maximize the user association ( $\psi_{u,m}$ ) and the sum rate of the prioritized ABS under the constraints of ABS 3D location ( $x_m, y_m, h_m$ ), power budget, and users' priority ( $\delta_{u,m}$ ).

The 3D ABS optimal placement plays a significant role in maximizing the number of associated users' matrix  $\Psi = (\psi_{u,m})$ , due to the joint optimization of the ABS 3D positions ( $x_m, y_m, h_m$ ). Hence, the proposed optimization framework is mathematically expressed as

$$\max_{x_m, y_m, h_m, p_m, \Psi} \sum_{u=1}^U \sum_{m=1}^{M-1} \delta_{u,m} \psi_{u,m} C_{u,m}, \quad \forall m \in \mathcal{M}, m \neq 0, \quad (11a)$$

$$C_m \geq C_{\text{min}}, \quad \forall m \in \mathcal{M} \quad (11b)$$

$$\sum_m \psi_{u,m} \leq 1, \quad \forall u \in \mathcal{U}, \quad \forall m \in \mathcal{M} \quad (11c)$$

$$\psi_{u,m^*} = 1, \quad m^* = \max_m \mathcal{R}_m, \quad \forall u \in \mathcal{U} \quad (11d)$$

$$|d_{u,m}|^2 \leq R^2, \quad \forall m \in \mathcal{M} \quad (11e)$$

$$h_{\text{min}} \leq h_m \leq h_{\text{max}}, \quad \forall m \in \mathcal{M}, m \neq 0 \quad (11f)$$

$$P_{\text{min}} \leq P_m \leq P_{\text{max}}, \quad \forall m \in \mathcal{M} \quad (11g)$$

$$\delta_{u,m} \in \{0, 1\}, \quad \forall u \in \mathcal{U}, \quad \forall m \in \mathcal{M} \quad (11h)$$

$$\psi_{u,m} \in \{0, 1\}, \quad \forall u \in \mathcal{U}, \quad \forall m \in \mathcal{M} \quad (11i)$$

We aim at maximizing the users' sum rate in Equation (11a) under the constraints of ABS location, associations, and transmit power. Constraint (11b) maintains the minimum sum-rate requirements of the users associated with the GBS. Constraint (11c) imposes that each user is associated with at most one ABS at a time and the constraint (11d) guarantees that a user is associated with the ABS that maximizes the reward function  $\mathcal{R}$ . Constraint (11e) ensures that the user  $u$  lies inside the ABS coverage when located within a distance  $R$  from the ABS center ( $x_m, y_m$ ). Constraints (11f,g) impose that the ABS altitude and transmit powers must be within the feasible region.  $\psi_{u,m}$  and  $\delta_{u,m}$  are the association and priority binary variables, respectively.

#### 5. Proposed Prioritized User Association Algorithm

The problem proposed in Equation (11a) is mathematically challenging because of a nonconvex objective function and nonlinear constraints. Therefore, it is not easy to provide

an optimal solution to this problem. To solve this challenging problem, we adopted a Q-learning-assisted  $\epsilon$ -greedy algorithm that maximized the proposed reward function ( $\mathcal{R}$ ) defined as

$$\mathcal{R} = \underbrace{C_{m,m \neq 0} C_{m,m=0}^2}_{(a)} - \underbrace{(C_{m,m=0} - C_{\min})}_{(b)} - \underbrace{(C_{m,m \neq 0} - C_{\min})^2}_{(c)} \quad (12)$$

In Equation (12), the term (a) implies that the higher the sum rate for users, the higher the reward. Moreover, we can notice that the GBS user sum rate is squared, which in turn prioritizes the emergency users in the proposed scenario. The terms (b), (c) are the deviations of the users' rate from their defined threshold, which in turn are subtracted from the reward.

The key steps involved in the proposal are to apply the  $k$ -means algorithm to form a cluster that places the ABS in the optimum 3D position and then assign the optimum power and associate the users to the ABS and GBS by adopting the Bellman's equation given by:

$$Q(s_t; a_t) = \max_a (E[\mathcal{R}_t + \beta Q(s_t, \hat{a}_t)]) \quad (13)$$

Here,  $E$  is the expectation operator and  $\beta$  the discount factor that lies in the range  $0 \leq \beta \leq 1$ . In Q-learning, the temporal difference (TD) is used to approximate the Q-function. For the Q-learning approach, the easiest one-step approach, adopted here, is to calculate

$$Q(s_t, a_t) \leftarrow (1 - \alpha)Q(s_t, a_t) + \alpha \max_{\hat{a}_t} (\mathcal{R}_t + \beta Q(s_t, \hat{a}_t)) \quad (14)$$

where  $\alpha$  is the learning rate.

After defining all the initial parameters, the agent (ABS) performs actions by allocating discrete power between  $P_{\min}$  and  $P_{\max}$  in every state and calculates the corresponding reward, and then updates the Q-table. Once the Q-table is updated, it results in a maximum number of connected users with the highest reward for the given optimized ABS location. Iterations are performed until the maximum reward is achieved. We summarized the proposed scheme in Algorithm 1.

---

**Algorithm 1:** Prioritized User Association for Sum-Rate Maximization Algorithm.

---

- 1 **Input:** Declare key variables and initialize the Q-table ( $Q(s_t, a_t) = 0$ );
  - 2 **Output:** Optimized user association;
  - 3 **for** iterations = 1 : maxIteration **do**
  - 4     Choose action ( $a_t$ ) from a given set of actions (i.e., user association combinations for various ABS heights);
  - 5     Take action and observe the reward using (12), and move to the next state ( $x_{t+1}$ );
  - 6     Update the Q-table using  
 $Q(s_t, a_t) \leftarrow (1 - \alpha)Q(s_t, a_t) + \alpha \max_{\hat{a}_t} (\mathcal{R}_t + \beta Q(s_t, \hat{a}_t))$ ;
  - 7 **end**
  - 8 **Return** Optimum user ABS association that maximizes the sum rate.
- 

## 6. Simulation Results

In this section, we consider the downlink heterogeneous environment where we deployed the ABS to improve the coverage of a destroyed BS. We modeled the short-term fading as a flat fading and the large-scale fading using a probabilistic path loss model given in Equation (7). A UAV network was simulated with  $M$  ABSs, where each ABS supported a minimum of one aerial user. As ABSs were deployed close to each other, they introduced interference between users and their neighbors. To meet the QoS requirements of users, we increased the density of deployed ABSs from  $M = 1$  to  $M = 16$ . We also

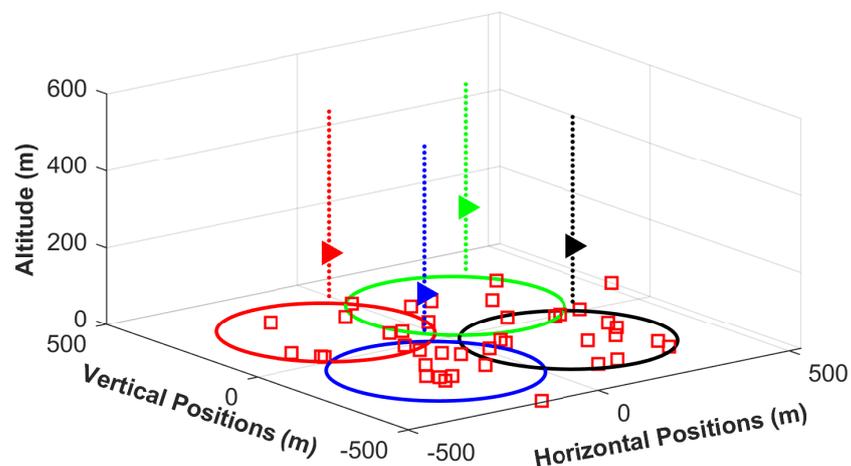
assumed that all the ABSs were operating at a carrier frequency of 900 MHz. The allocated power in the range  $[P_{\min}, P_{\max}]$  was divided into 20 discrete steps with a step size of 1.5 dBm. The minimum sum-rate ( $C_{\min}$ ) requirement for each user was set to 0.5 bps/Hz. We considered a noise power equal to  $-174$  dBm/Hz. The remaining key simulation parameters are summarized in Table 3.

**Table 3.** Simulation parameters.

Parameters	Value
Users ( $U$ )	40
ABS ( $M$ )	16
$P_{\min}$	$-20$ dBm
$P_{\max}$	25 dBm
Step size	1.5
$h_{\min}$	100
$h_{\max}$	600
radius( $R$ )	250 m
$\alpha$	0.5
$\beta$	0.9
maxIteration	50,000

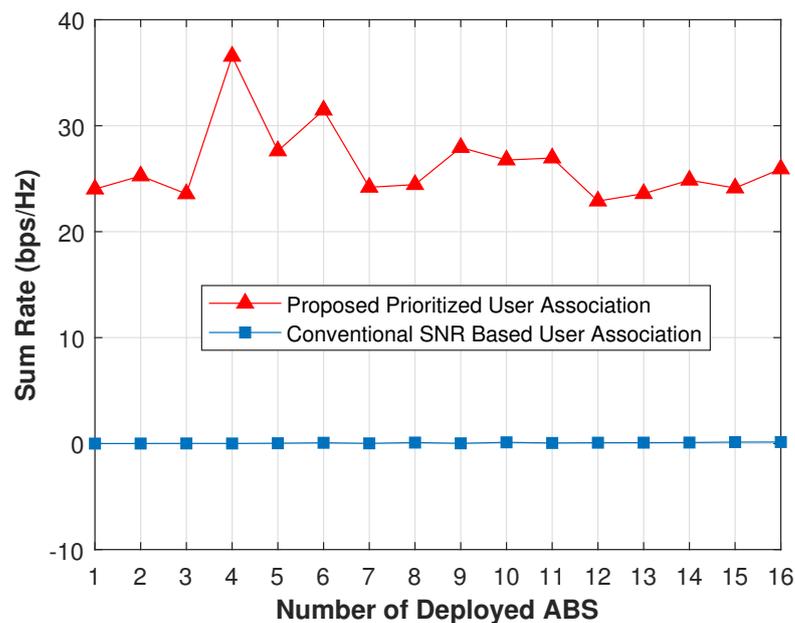
The performance of the proposed scheme was tested to show the effectiveness of the proposed Q-learning-based prioritized user association scheme compared to the conventional signal-to-noise (SNR)-based user association scheme, where users are associated with the ABS receiving the maximum SNR. During simulations, we randomly deployed 40 users in the coverage area following a uniform distribution. We varied the number of deployed ABSs to obtain optimized ABS locations, user association, and transmit power. In this section, we verify the performance of the proposed scheme during an outage on the basis of the sum rate. The simulations were performed using MATLAB.

First of all, we plotted the optimal deployment positions for four ABSs in Figure 4, where clusters were found using the  $k$ -means algorithm. Here, each ring shows the coverage range of the respective ABSs. For example, the red ring shows the coverage range of the red ABS and the users lying inside this ring are connected to that ABS, and similarly, for the green, blue, and black rings, respectively. Users were associated with an ABS on the basis of the maximum reward defined in Equation (12). The coverage range of each ABS was 225 meters, that is, the radius of each ABS. The ABSs were deployed to minimize the coverage overlap among ABSs. The altitude of ABSs was variable as shown with the dotted lines, where each ABS moved between the defined altitude range.



**Figure 4.** Optimized 3D ABS deployment of ABSs and users using the proposed scheme.

Figure 5 shows the sum-rate performance for the proposed prioritized user association scheme compared to a conventional SNR-based association scheme. In the proposed scheme, the Q-learning algorithm was implemented to optimize the power compared to the conventional scheme, which, in turn, considered the real-time environment resulting in significant performance gain. Moreover, we can notice that the sum rate increases as the number of ABSs increases but if it increases above four then the sum rate begins to decrease. The reason is that as the number of ABSs increases, interference among them also increases, which lowers the sum rate. By comparing these two schemes, we achieved a maximum value for the average sum rate of around 36.55 bps/Hz compared to the conventional scheme.

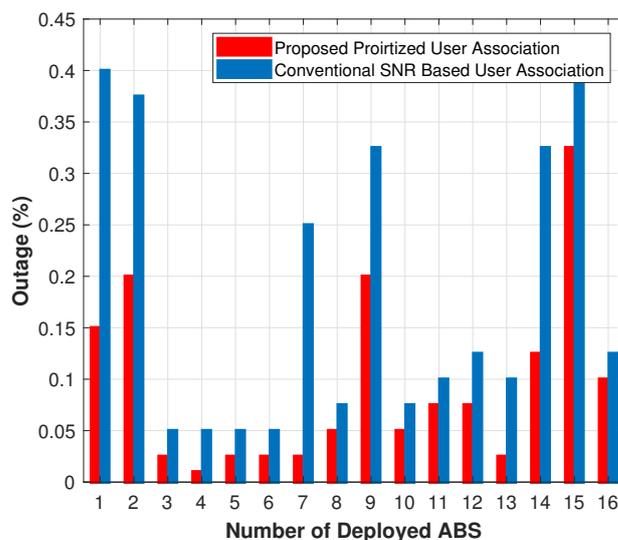


**Figure 5.** Comparison of sum-rate performance under the proposed and the conventional SNR-based user association scheme.

Finally, we compared the user outage performance for both schemes in Figure 6. The simulation results show that the number of associated users is maximum when the number of deployed ABSs are four, under the proposed scheme, as the percentage of outage is far less in that case. By increasing the number of ABSs above four, the interference among users rises, which decreases the received SNR below the threshold. Moreover, by comparing the proposed scheme with the conventional one for a varying number of deployed ABSs, we observe that the number of outage users are less in the proposed scheme. Hence, the proposed scheme performs better in terms of user association than the conventional scheme.

To verify the efficacy of the proposed scheme, we also compare the mean sum-rate performance of the proposed prioritized-based user association scheme with the following schemes:

- Benchmark: we tried every possible combinations of ABS user associations, which resulted in the highest mean sum rate compared to other schemes.
- SINR-based user association: users are associated with the ABS from which they received the maximum SINR.
- Distance-based user association: users are associated with the nearest ABS.
- Random user association: users are associated randomly with any ABS in the vicinity without caring about any requirement.



**Figure 6.** Outage performance comparison.

From the results in Table 4, we can clearly notice that the benchmark algorithm has the best mean sum rate because it tries every deployment positions and associate users with the ABS that maximizes the sum rate. The proposed scheme has the 2nd best performance because it prioritized the users with emergency communication requirements. In distance-based association, performance degrades because users can be associated with the nearest base station, but this does not guarantee good channel conditions. Obviously, in random association, the performance is worst because it disregards any QoS requirements.

**Table 4.** Sum-rate comparison of different algorithms.

Algorithms	Mean Sum Rate (bps/Hz)
Benchmark	30.5
Proposed prioritized user association	24
SINR-based user association	1.5
Distance-based user association	1.3
Random user association	1.1

## 7. Conclusions

Unmanned air vehicle placement and user association is an important challenge to enable emergency communications. In this paper, we handled the problem of optimizing the user sum rate by associating the prioritized users to the deployed emergency aerial base stations. Simulation results verified that the performance under the proposed scheme significantly increased in terms of sum rate and the reduction in the number of users experiencing outages compared to the conventional SNR-based user association scheme. Moreover, we noticed that the sum rate increased as the number of ABSs increased, but after deploying more than four ABSs, the sum rate began to decrease. The reason is that as the number of ABS increases, interference among them also increases, which lowers the sum rate. By comparing these two schemes, we achieved a mean sum rate of around 23 bps/Hz compared to the conventional SINR-based user association scheme. In the future, we plan to further improve the system performance using state-of-the-art deep learning algorithms and also compare the proposal with game- and graph-theory-based approaches.

**Author Contributions:** Conceptualization, Writing—original draft, A.B.S.; Data curation, I.A.; Validation, Funding acquisition, A.A.; Writing—review & editing, U.J.; Conceptualization, Supervision,

Writing—review & editing, Z.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work is funded by Higher Education Commission (HEC) Pakistan under the NRPU 2021 Project with grant no. 15687.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Boccardi, F.; Heath, R.W.; Lozano, A.; Marzetta, T.L.; Popovski, P. Five disruptive technology directions for 5G. *IEEE Commun. Mag.* **2014**, *52*, 74–80. [[CrossRef](#)]
- Zeng, Y.; Zhang, R.; Lim, T.J. Wireless communications with unmanned aerial vehicles: Opportunities and challenges. *IEEE Commun. Mag.* **2016**, *54*, 36–42. [[CrossRef](#)]
- Li, Y.; Zhang, J.; Sun, G.; Lu, D. The sparsity adaptive reconstruction algorithm based on simulated annealing for compressed sensing. *J. Electr. Comput. Eng.* **2019**, *2019*, 1–8. [[CrossRef](#)]
- Shakoor, S.; Kaleem, Z.; Baig, M.I.; Chughtai, O.; Duong, T.Q.; Nguyen, L.D. Role of UAVs in Public Safety Communications: Energy Efficiency Perspective. *IEEE Access* **2019**, *7*, 140665–140679. [[CrossRef](#)]
- Al-Hourani, A.; Kandeepan, S.; Lardner, S. Optimal LAP altitude for maximum coverage. *IEEE Wirel. Commun. Lett.* **2014**, *3*, 569–572. [[CrossRef](#)]
- Ueyama, J.; Freitas, H.; Faical, B.S.; Filho, G.P.; Fini, P.; Pessin, G.; Gomes, P.H.; Villas, L.A. Exploiting the use of unmanned aerial vehicles to provide resilience in wireless sensor networks. *IEEE Commun. Mag.* **2014**, *52*, 81–87. [[CrossRef](#)]
- Xi, X.; Cao, X.; Yang, P.; Chen, J.; Quek, T.; Wu, D. Joint User Association and UAV Location Optimization for UAV-Aided Communications. *IEEE Wirel. Commun. Lett.* **2019**, *8*, 1688–1691. [[CrossRef](#)]
- Qiu, C.; Wei, Z.; Feng, Z.; Zhang, P. Joint Resource Allocation, Placement and User Association of Multiple UAV-Mounted Base Stations With In-Band Wireless Backhaul. *IEEE Wirel. Commun. Lett.* **2019**, *8*, 1575–1578. [[CrossRef](#)]
- Li, Q.; Ding, M.; Ma, C.; Liu, C.; Lin, Z.; Liang, Y.C. A Reinforcement Learning Based User Association Algorithm for UAV Networks. In Proceedings of the 2018 28th International Telecommunication Networks and Applications Conference (ITNAC), Sydney, Australia, 21–23 November 2018; pp. 1–6. [[CrossRef](#)]
- Alzenad, M.; El-Keyi, A.; Lagum, F.; Yanikomeroglu, H. 3-D placement of an unmanned aerial vehicle base station (UAV-BS) for energy-efficient maximal coverage. *IEEE Wirel. Commun. Lett.* **2017**, *6*, 434–437. [[CrossRef](#)]
- Mozaffari, M.; Saad, W.; Bennis, M.; Debbah, M. Efficient deployment of multiple unmanned aerial vehicles for optimal wireless coverage. *IEEE Commun. Lett.* **2016**, *20*, 1647–1650. [[CrossRef](#)]
- Lu, X.; Xiao, L.; Dai, C.; Dai, H. UAV-aided cellular communications with deep reinforcement learning against jamming. *IEEE Wirel. Commun.* **2020**, *27*, 48–53. [[CrossRef](#)]
- Shakhatreh, H.; Khreishah, A.; Ji, B. Providing wireless coverage to high-rise buildings using UAVs. In Proceedings of the 2017 IEEE International Conference on Communications (ICC), Paris, France, 21–25 May 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 1–6.
- Azari, M.M.; Rosas, F.; Chen, K.C.; Pollin, S. Ultra reliable UAV communication using altitude and cooperation diversity. *IEEE Trans. Commun.* **2017**, *66*, 330–344. [[CrossRef](#)]
- Qian, Y.; Wang, F.; Li, J.; Shi, L.; Cai, K.; Shu, F. User association and path planning for UAV-aided mobile edge computing with energy restriction. *IEEE Wirel. Commun. Lett.* **2019**, *8*, 1312–1315. [[CrossRef](#)]
- Wang, L.; Huang, P.; Wang, K.; Zhang, G.; Zhang, L.; Aslam, N.; Yang, K. RL-based user association and resource allocation for multi-UAV enabled MEC. In Proceedings of the 15th International Wireless Communications & Mobile Computing Conference (IWCMC), Tangier, Morocco, 15–19 June 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 741–746.
- Arani, A.H.; Azari, M.M.; Melek, W.; Safavi-Naeini, S. Learning in the sky: Towards efficient 3D placement of UAVs. In Proceedings of the 2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications, London, UK, 31 August–3 September 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 1–7.
- Bor-Yaliniz, R.I.; El-Keyi, A.; Yanikomeroglu, H. Efficient 3-D placement of an aerial base station in next generation cellular networks. In Proceedings of the 2016 IEEE International Conference on Communications (ICC), Kuala Lumpur, Malaysia, 23–27 May 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 1–5.
- ur Rahman, S.; Kim, G.H.; Cho, Y.Z.; Khan, A. Positioning of UAVs for throughput maximization in software-defined disaster area UAV communication networks. *J. Commun. Netw.* **2018**, *20*, 452–463. [[CrossRef](#)]
- Shakoor, S.; Kaleem, Z.; Do, D.T.; Dobre, O.A.; Jamalipour, A. Joint Optimization of UAV 3-D Placement and Path-Loss Factor for Energy-Efficient Maximal Coverage. *IEEE Internet Things J.* **2021**, *8*, 9776–9786. [[CrossRef](#)]
- Jiao, S.; Xie, X.; Ding, Z. Deep Reinforcement Learning Based Optimization for IRS Based UAV-NOMA Downlink Networks. *arXiv* **2021**, arXiv:2106.09616.
- Shakhatreh, H.; Alenezi, A.; Sawalmeh, A.; Almutiry, M.; Malkawi, W. Efficient Placement of an Aerial Relay Drone for Throughput Maximization. *Wirel. Commun. Mob. Comput.* **2021**, *2021*, 1–11. [[CrossRef](#)]
- Lee, W.; Jeon, Y.; Kim, T.; Kim, Y.I. Deep Reinforcement Learning for UAV Trajectory Design Considering Mobile Ground Users. *Sensors* **2021**, *21*, 8239. [[CrossRef](#)] [[PubMed](#)]

24. Gopi, S.P.; Magarini, M. Reinforcement Learning Aided UAV Base Station Location Optimization for Rate Maximization. *Electronics* **2021**, *10*, 2953. [[CrossRef](#)]
25. Amiri, R.; Mehrpouyan, H.; Fridman, L.; Mallik, R.K.; Nallanathan, A.; Matolak, D. A machine learning approach for power allocation in HetNets considering QoS. In Proceedings of the 2018 IEEE International Conference on Communications (ICC), Kansas City, MO, USA, 20–24 May 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1–7.
26. Kumar, S.; Suman, S.; De, S. Backhaul and delay-aware placement of UAV-enabled base station. In Proceedings of the IEEE INFOCOM 2018-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), Honolulu, HI, USA, 15–19 April 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 634–639.
27. Mahmood, A.; Usman, M.Q.; Shahzad, K.; Saddique, N. Evolution of Optimal 3D placement of UAV with Minimum Transmit Power. *J. Wirel. Commun. Mob. Comput.* **2019**, *7*, 13–18. [[CrossRef](#)]
28. Kim, M.; Lee, J. Outage Probability of UAV Communications in the Presence of Interference. In Proceedings of the IEEE Global Communications Conference (GLOBECOM), Abu Dhabi, United Arab Emirates, 9–13 December 2018; pp. 1–6. [[CrossRef](#)]