Deep Reinforcement Learning for Jointly Resource Allocation and Trajectory Planning in UAV-Assisted Networks

Arwa Mahmoud Jwaifel(✉) and Tien Van Do

Department of Networked Systems and Services, Faculty of Electrical Engineering and Informatics, Budapest University of Technology and Economics (BME), Budapest, Hungary
jwaifel@hit.bme.hu

Abstract. Unmanned aerial vehicles (UAVs) have diverse applications in various fields, including the deployment of drones in 5G mobile networks and upcoming 6G and beyond. In UAV wireless networks, where the UAV is equipped with an eNB or gNB, it is critical to position it optimally to serve the maximum number of users located in high-capacity areas. Furthermore, the high mobility of users leads to greater network dynamics, making it challenging to predict channel link states. This study examines the use of Proximal Policy Optimization (PPO) to optimize the joint UAV position and radio spectrum resource allocation to meet the users' quality-of-service (QoS) requirements.

Keywords: 5G · 6G · Unmanned aerial vehicle (UAV) · resource allocation optimization · deep reinforcement learning (DRL) · Proximal Policy Optimization (PPO) · Deep Reinforcement Learning (DQN)

1 Introduction

The beauty of Unmanned Aerial Vehicles (UAVs), which includes drones, have recently attracted lots of researcher's attention in the industrial fields due to their ability to operate and monitor activities from remote locations; moreover, UAVs are well known for their portability, lightweight, low cost and flying without a pilot. UAV features make it suitable to be integrated into the fifth-generation (5G) and the networks beyond 6G wireless networks, where UAV can be deployed as aerial base stations into what is called the UAV-assisted [1,2]. Such situations include quick service recovery after a natural disaster and offloading base stations or the Next Generation Node B (gNBs) at hotspots in case of failure or malfunction of the ground base station or the gNB. In addition, UAV can be used to enhance network coverage and performance, where the location of the UAV can be controlled and dynamically changed to optimize the network performance according to the users' needs and their mobility model. Such scenarios are represented in Fig. 1.

N. T. Nguyen et al. (Eds.): ICCCI 2023, LNAI 14162, pp. 71–83, 2023.
https://doi.org/10.1007/978-3-031-41456-5_6

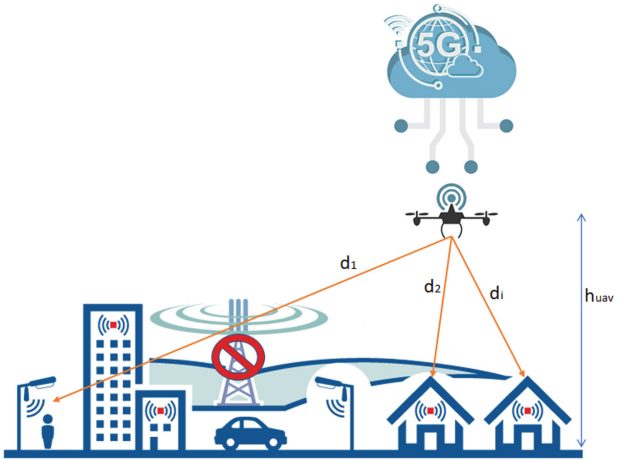

Fig. 1. UAV emergency model.

A UAV-assisted application was investigated in terms of performance analysis, resource allocation, UAV placement and position optimization, channel modeling, and information security as in [3,4] and [5]. UAV-assisted wireless communications have three main types; the first type is called UAV-carried Evolved Node B (eNB) or gNB, where the UAV acts as an aerial base station and is used to extend the network coverage [6,7] and [8]. The second type is called UAV relaying, where the UAVs are used as aerial relays to provide a wireless connection for users that cannot communicate with each other directly [9,10]. Finally, the third type is identified as a UAV-assisted Internet-of-Things (IoT) network, where UAVs assist the IoT network in collecting/disseminating data from/to its nodes or charging its nodes [11] and [12].

However, due to the UAV's limitations, only some applications use UAVs in the existing systems. The fundamental limitation is the battery life of the UAV, which is affected by the high power consumption dissipated in the hovering, horizontal, and vertical movements of the drone. Besides the battery life, the position of the UAV is also a significant concern in implementing real systems.

One of the significant applications of using the UAV in the communication system is during emergencies (such as floods or earthquakes, ... etc.) while the infrastructure is partially or totally unavailable, and the need to provide mobile service to the users is highly required. In these situations, the UAV can perform this task and provide mobile services to the user equipment (UE's) while granting the required quality-of-service (QoS). The main challenge for using the UAV-assisted network is to find the optimal position of the UAV in the cell area before getting a dead battery. Which is very complicated and challenging to determine, and the traditional optimization methods of artificial intelligence (AI) cannot solve those complicated optimization problems.

In order to address those two concerns, Reinforcement learning (RL) algorithms are applied, especially deep reinforcement learning, which has been proven to outperform the existing traditional algorithm. In this work, we introduced a different deep RL algorithm to solve the UAV-assisted joint position and radio resource allocation optimization problem. The main target is to find the optimal position of the UAV that is dynamically changed concerning the UE's required QoS and consider the UAV battery energy level in each time step, in addition to the required energy to get back to the start point.

Our main contribution in this study is presented as follows. We developed a method that collaboratively optimizes communication resource allocation and position for the UAV based on reinforcement learning, where the position and radio resource allocation joint optimization problem is formulated to obtain the maximum cumulative discounted reward. For the non-convexity nature of the optimization problem, we designed and applied different deep reinforcement learning algorithms for the UAV to solve the joint optimization issue, then we compared these algorithms' performance to solve the proposed problem; these algorithms are Proximal Policy Optimization (PPO) and Deep Reinforcement Learning (DQN).

Section 2 reviews the related literature on optimizing the position and resource allocation in UAV-assisted networks. Also, we review the reinforcement learning application in such optimization problems for UAV-assisted wireless networks. System model and problem formulation are illustrated in Sect. 3, and simulation and results are presented in Sect. 4. The conclusion is discussed in Sect. 5.

2 Related Work

The design of UAV position for improving various communication performance metrics has gained significant attention, as shown in various studies such as in [13], which focused on optimizing the spectrum efficiency and energy efficiency of a UAV-enabled mobile relaying system by adjusting the UAV's flying speed, position, and time allocation. [14] aimed to optimize the global minimum average throughput through optimized UAV trajectories and OFDMA (orthogonal frequency-division multiple access) resource allocation. [15] explored the UAV-enabled wireless communication system with multiple UAVs and aimed to increase the minimum user throughput by optimizing communication scheduling, power allocation, and UAV trajectories. In [16], UAVs served as flying Base Stations (BSs) for vehicular networks, delivering data from vehicular sources to destination nodes. The authors determined the optimal UAV position and radio resource allocation by combining Linear Programming and successive convex approximation methods.

Despite the deployment optimization of UAVs, machine learning (ML) algorithms have been introduced to optimize different QoS network requirements. The reinforcement RL and deep learning (DL) received the foremost researchers' focus in this field. Such researches as in [17], where the authors proposed UAV

autonomous indoor navigation and target detection approach based on a Q-learning algorithm. While in [18], the authors proposed multi-agent reinforcement learning to optimize the resource allocation of the multi-UAV networks, and the algorithm is designed to maximize the systems' long-term reward. The authors of [19] have considered RL algorithms to optimize UAV's position to maximize sensor network data collection under QoS constraints. Moreover, in [20], the researchers adopted deep learning RL based to dynamically allocate radio resources in heterogeneous networks.

Based on the related literature review, a limited number of researchers are solving the UAV position's joint optimization problem and the UE's resource allocation. Motivated by that, we applied the deep RL algorithms to solve this optimization problem.

3 An RL-Based Approach

We considered a multi-rotor UAV with total energy E_{max} that flying at a fixed altitude of h_{max} from a base point denoted by $s_0 = (x_0, y_0)$. The UAV has an onboard gNB that will serve K subscribers within a specific area. At the beginning (τ_i) of time slot i, the gNB decides the assignment of Resource Blocks (RB) for each customer according to specific criteria; in our study, we adopt the customer's QoS requirements, and the channel quality, where the gNB can measure the channel quality of each user's device and allocate the RB's based on a minimum requirement to maintain the network performance. We assume that the gNB receives the CQI values. ($CQI(i) = [CQI_{1,i}, CQI_{2,i}, \ldots, CQI_{k,i}]$) of $k = \{1, ..., K\}$ user equipment (UEs) at time instance τ_i where $i = 0, \ldots$, which is in accordance with the time-slot operation of the gNB, so $\tau_{i+1} - \tau_i = \Delta$. At each time step $\tau_i = a \times i \times \Delta$, the UAV decides to continue flying or get back to the base point while monitoring the battery level. For this problem, we apply Reinforcement learning (RL) for flight control as follows:

- At each time step τ_i, the state $s_i = [(x_i, y_i, h_{max}, E_i), [CQI_{k,i}]] \quad \forall\, k \in [0, K]$ consists of UAV position, which can be denoted by the coordinates (x_i, y_i, h_{max}) and the UAV battery energy level, in addition to the received CQI values, form the UE's $CQI_{k,i} \forall k \in [1, K]$, and the UAV battery level E_i.
- We assume that the altitude of the UAV is fixed in this study, which can lead to the possible actions: backward, forward, left, right, and hovering in the same location and returning to the base point. The action space is $\mathcal{A} == \{L, R, FW, BW, HO, RE\}$.
- The reward function $r_i = \sum_{k=1}^{K} U_{k,i}$ is defined as the total number of served UE's in each time step, where the binary variable $U_k \in \{0, 1\}, \forall k$, which is asserted if the UAV succeeded in serving the k^{th} UE, and allocated the required resources to guarantee the minimum throughput required to provide coverage for the cell in emergencies. Otherwise, U_k is set to 0. In this study, we adopt the max CQI scheduling allocation of the UE's, where the UE's with the highest values of CQI are allocated while there are available resource blocks in the radio frame.

The energy consumption of the UAV consists of mainly two parts: one that is required to provide the onboard gNB with its energy to operate, and the other is the propulsion energy of the UAV so that it can fly around. The UAV will decide to get back to the base point by monitoring its battery energy level (E_i) at each time step τ_i, and compare it with the energy required to fly back to the start point $s_0 = (x_0, y_0)$ from its position point ($E_{i+1,r}$). The UAV battery energy constraints is assumed to be:

$$E_i > E_{i,r} \quad \text{AND} \quad E_{i+1} > E_{i+1,r}. \tag{1}$$

4 Simulation and Analysis

4.1 Models Used in Simulation

User Mobility. User mobility modeled in this research is based on the Gauss-Markov Mobility Model [21]. Where the Mobile nodes (UE's) are located in random locations within the cell area, these nodes will set their speed as for the k^{th} UE the speed is denoted as ($V_{i,k}$) and its direction denoted as ($D_{i,k}$) for each specific step (i). At every step i, the current position of the k^{th} UE coordinates ($x_{k,i}, y_{k,i}$) depends on the previous location ($x_{k,i-1}, y_{k,i-1}$), previous speed $V_{k,i-1}$ and previous direction $D_{k,i-1}$, assuming the directions values can be set to $\in [0, 90, 180, 270]$, to follow the proposed grid world model of the network cell. The k^{th} UE position at the i^{th} step, is expressed as

$$\begin{aligned} X_{k,i} &= X_{k,i-1} + V_{k,i-1} \cos D_{k,i-1}, \\ Y_{k,i} &= Y_{k,i-1} + V_{k,i-1} \sin D_{k,i-1}. \end{aligned} \tag{2}$$

Parameters $V_{k,i-1}$ and $D_{k,i-1}$ are chosen from a random Gaussian distribution with a mean equal to 0 and a standard deviation equal to 1.

RB Scheduling Algorithm. In our study, we adopt the best-CQI scheduling algorithm to allocate RB to the UE, where the gNB Scheduler allocates the RBs to the UE's that reported the highest CQI during Transmission Time Interval (TTI), where the higher CQI value means a better channel condition.

Energy Consumption Model for Multi-rotor UAV. In this study, we considered rotary-wing UAV, the UAV has four brushless motors which are powered by the carried battery, and they rotate at the same constant speed ω_{rotor}. The UAV will fly to a specific position and hover or continue flying to the next position. We follow the forces model in [22] to derive the energy consumption for both UAV motion phases. The propulsion power of the UAV is essential to support the UAV's hovering and moving activities either the vertical movement, where in our study, we assumed the UAV height is constant; thus, we will not consider this movement phase, the other movement type is the horizontal movement from one position to another in the cell grid.

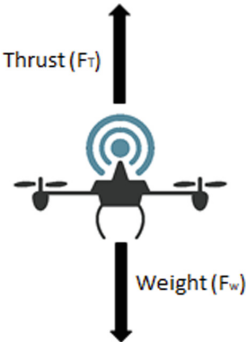

Fig. 2. UAV hovering state forces.

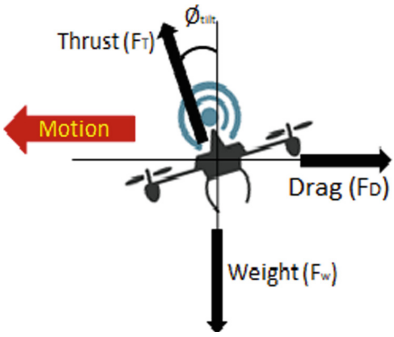

Fig. 3. UAV forward state forces.

Hovering is one of the motion activities of the drone, where the thrust of the rotor is used to equilibrate the gravity effect completely; Fig. 2 represents the hovering phase forces. Thrust is denoted by:

$$F_T = \frac{1}{2}\rho N_{rotor} A_{uav} V_{uav} \omega_{rotor}, \tag{3}$$

where ρ is the air density and equals to $(1.225\,\mathrm{kg/m^2})$, the rotor propeller area is A_{uav} and is equal to $A_{uav} = \pi r_{uav}^2$ where r_{uav} is the propeller radius. Finally, the number of UAV rotors is represented by the variable N_{rotor}. The V_{UAV} is the resultant velocity of the drone, and the hovering phase is equal to the motor speed, which is denoted by ω_{rotor} and can also be defined as the induced velocity of the rotor blades.

In the hovering phase, the thrust of the drone motors must equal the gravitational force $(m_{tot} \times g)$, where the value of $V_{uav} = \sqrt{2m_{tot} \times g/(\rho A_{uav} N_{rotor})}$. Accordingly, in time step duration Δ where the power is equal to $P_{hov} = F_T V_{uav}$, with $V = 0$, the energy that the battery must supply is only that to defy the weight force, and considering the UAV motor efficiency η_{mot} and the propeller efficiency η_{pro} is defined as

$$E_{hov} = \sqrt{\frac{2(m_{tot} \times g)^3}{\rho A_{uav} N_{rotor}}} \times \frac{1}{\eta_{mot}\eta_{pro}} \times \Delta. \tag{4}$$

where m_{tot} is the total mass in Kg and equals to the sum of UAV mass (m_{uav}), the payload (the carried gNB) (m_{pld}) and the battery (m_b), i.e. $m_{tot} = m_{uav} + m_b + m_{pld}$. The earth gravitational force g and equals to $(9.81 \approx 10\,\mathrm{m/s^s})$. Finally, η_{mot} is the efficiency of the UAV motor.

The UAV horizontal movement is considered the most challenging drone motion to estimate; where according to Newton's 1^{st} low where the drone required to generate motors thrust force (F_T) that is equal and opposite to the total sum of forces consists of drag force (F_D) due to the drone speed and

the weight force ($F_W = m_{tot} \times g$) due to the total weight of the drone and it is cargo (battery and carried gNB). All horizontal movement forces are shown in Fig. 3. The vertical forces under the equilibrium condition are mathematically represented by

$$F_T \times \cos \phi_{tilt} = m_{tot} \times g. \tag{5}$$

Applying Newton's 1^{st} law to find the UAV velocity required to maintain the required conditions, the forces are denoted by

$$F_D = F_W \times \tan \phi_{tilt} = \frac{1}{2} C_D \rho A_{uav}^{eff} V_{UAV}^2, \tag{6}$$

where C_D represents the drag coefficient, and A_{uav}^{eff} represents the vertical projected area of the UAV and can be evaluated as $A_{uav}^{eff} = A_{uav}^{side} \sin(90 - \phi_{tilt}) + A_{uav}^{top} \sin \phi_{tilt}$, where A_{uav}^{side} and A_{uav}^{top} represents the side and top surface of the UAV, which can be approximated as $A_{uav}^{eff} = A_{uav}^{top} \sin \phi_{tilt}$. To evaluate the UAV energy consumed in the horizontal movement of the drone with constant speed, and using Eqs. 5 and 6, the power formula denoted by $P_{hor} = F_T V_{uav}$ is presented as

$$E_{hor} = \sqrt{\frac{2(m_{tot} \times g)^3}{C_D \rho A_{uav}^{eff} N_{rotor}} \times \frac{\sin \phi_{tilt}}{\cos^3 \phi_{tilt}}} \times \frac{1}{\eta_{mot} \eta_{pro}} \times \Delta, \tag{7}$$

where η_{mot} and η_{pro} are the efficiency of the motor and the propeller, respectively. The UAV properties and parameters values used in the simulation are represented in Table 1, in addition to the UAV battery specifications, which represent the battery model installed with DJI Matrice 600 Pro drone models [23]. At a given trajectory (x_i, y_i, h_{max}), the remaining energy of the UAV can be expressed as

$$E_i = E_{max} - \sum_{i=0}^{i} E_i. \tag{8}$$

Energy Model of gNB. Path loss is modeled as the probability model that consists mainly of two components, i.e., LoS and NLoS. LoS connection probability between the receiver and transmitter is an essential factor and can be formulated as [24]

$$p_{\text{LoS},k}(i) = \frac{1}{1 + a_{LOS} \cdot \exp\left(-b_{LOS}\left(\phi_k(i) - a_{LOS}\right)\right)}, \tag{9}$$

where a_{LOS} and b_{LOS} are environmental constants, and $\phi_k(i)$ is the elevation angle in degree, and it depends on the UAV height as well as the distance between the UAV and user k, the elevation angle can be evaluated from $\phi_k = -\frac{180}{\pi} \sin^{-1}(\frac{h(i)}{d_k(i)})$. Furthermore, $h(i)$ is the UAV height, and $d_k(i)$ is the distance between the UAV and the k^{th} UE and defined as

$$d_k(i) = \sqrt{h^2(i) + (x(i) - x_k(i))^2 + (y(i) - y_k(i))^2}. \tag{10}$$

Table 1. UAV energy model parameters simulation values.

UAV and motor parameters		
Notations	Physical definition	Simulation value
m_{uav}	UAV Weight (6×TB48S batteries)	10 kg
ρ	Air density in kg/m^3	1.225
C_D	Drag coefficient	0.044
r_{uav}	Propeller radius in meter [m]	0.1905
A_{uav}^{top}	UAV top area [m^2]	0.3
V_{uav}	Max UAV speed	18 m/s
N_{rotor}	Number of rotors	6
ϕ_{tilt}	Tilt Angle values	25°
$N_{battery}$	Number of batteries	6
η_{mot}	Motor efficiency	0.8
η_{pro}	Propeller system efficiency	0.8
UAV battery model parameters		
Parameter		Simulation value
Battery model		TB48S
Battery type		LiPo 6S
Weight		680 g
Capacity (Q)		5700 mAh
Voltage		22.8 V
Energy		129.96 Wh
eNB model parameters [26]		
Parameter		Simulation value
LTE Mode		TDD
Frequency Bands		400 Mhz: (400–430) Mhz
		600 Mhz: (566–626) Mhz, (606–678) Mhz
		1.4 Ghz: (1447–1467) Mhz
		1.8 Ghz: (1785–1805) Mhz
Channel Bandwidth		5/10/15/20 MHz
Max Output Power		15 W
Power Supply		48V DC or 220 V AC
Power Consumption		150 W
MIMO		2 × 2
Dimensions		330 * 260 * 110 mm
Weight		5.5 kg
Users		200

The probability of having NLoS communication between the UAV and k^{th} UE is denoted by:

$$p_{\mathrm{NLoS},k}(i) = 1 - p_{\mathrm{LoS},k}(i). \tag{11}$$

Hence, the mean path loss model (in dB) we adopt the following equation from [24]

$$L_{\mathrm{k}}\left(h, d_k, i\right)(\mathrm{dB}) = L_{\mathrm{LoS},k}(i) \times p_{\mathrm{LoS},k}(i) + L_{\mathrm{NLoS},k}(i) \times p_{\mathrm{NLoS},k}(i), \tag{12}$$

where, $L_{\text{LoS},k}(i)$ and $L_{\text{NLoS},k}(i)$ are the path loss for LoS and NLoS communication links and denoted by

$$L_{\text{LoS},k}(i) = 10 \times \alpha_{pl} \log\left(\frac{4\pi f_c d_k(i)}{c}\right) + \delta_{\text{LoS}}, \tag{13}$$

$$L_{\text{NLoS},k}(i) = 10 \times \alpha_{pl} \log\left(\frac{4\pi f_c d_k(i)}{c}\right) + \delta_{\text{NLoS}}, \tag{14}$$

where α_{pl} is the path loss exponent and its environment-dependent variable, both of the δ_{LoS} and δ_{NLoS} are the mean losses due to LoS and NLoS communication links, $c = 3 \times 10^8$ the speed of light and f_c is the network operating frequency.

With $\gamma_k(i)$ represents the Signal-to-Noise Ratio (SNR) of the k^{th} UE at the i^{th} step, while assuming $P_{r,k}(i)$ is the received signal power at the k^{th} UE, the SNR is defined as

$$\gamma_k(i) = \frac{P_{r,k}(i)}{\sigma^2}. \tag{15}$$

The SNR can be rewritten in terms of the path loss and transmitted UAV power as

$$\gamma_k(i) = \frac{P_k(i) \times L_k(i)}{\sigma^2}, \tag{16}$$

where $P_k(i)$ is the transmitted power from the UAV to the k^{th} UE at the i^{th} step.

The 5G NR maximum data rate of the k^{th} UE can be evaluated in Mbps using the formula defined in [25], and expressed as:

$$R_k(i) = 10^{-6} \cdot \sum_{j=1}^{J} \left(\Omega j, k \cdot M_{j,k} \cdot \zeta_{j,k} \cdot C_{R,\max} \cdot \frac{N_{j,k}^{RB}(i) \cdot 12}{T_s^{\mu}} \cdot (1 - OH_{j,k})\right), \tag{17}$$

where J represents the number of aggregated component carriers, $(\Omega_{j,k})$ is the maximum number of layers, and $(M_{j,k})$ is the modulation order. In contrast, $(\zeta_{j,k})$ is a scaling factor that has values of (1, 0.8, 0.75, and 0.4). The code rate is denoted by $(C_{R,\max})$, and is can have the values in Tables 5.1.3.1-1, 5.1.3.1-2 and 5.1.3.1-3 in 3gpp.38.214 with a maximum value of (948/1024). The numerology μ can have the values of $[0, 1, 2, 3, 4]$ which responds to the subcarrier spacing (SCS) of 15 kHz, 30 kHz, 60 kHz, 120 kHz and 240 kHz. The variable T_s^{μ} represents the average OFDM symbol duration for certain μ and can be evaluated as $(T_s^{\mu} = \frac{10^3}{14 \times 2^{\mu}})$. The $N_{j,k}^{RB}(i)$ is the number of allocated RBs to the k^{th} UE at the i^{th} step. Finally, $OH_{j,k}$ denotes the overhead and can have the values of (0.14, 0.18, 0.08, and 0.10).

Moreover, the data rate can have another formula to be evaluated according to [25], as

$$R_k(i) = 10^{-3} \cdot \sum_{j=1}^{J} TBS_{j,k}(i) \times 2^{\mu}, \tag{18}$$

where $TBS_{j,k}$ is the total maximum number of DL-SCH transport block bits received within a 1ms TTI for the k^{th} UE and j^{th} carrier.

4.2 Simulation Results

In our case study, we considered one UAV that flies at a maximum altitude of $h_{max} = 200\,\text{m}$ over grid area size (1500×1500), The simulation parameters listed in Table 1 and the network setting listed in Table 2. In each episode, there are two scenarios for the UE's mobility. One scenario is considering 20 number of UE which are generated and distributed randomly in the cell area while assuming random walk mobility model to be the mobility model for the UE within the cell, moreover, the second scenario considered placing four UE's and fix their positions at the corners of the cell.

Table 2. Parameters for simulation.

Parameter	Value
Bandwidth	10 MHz
Transmitted power	23 dBm
Frequency	2 GHz
Noise power	−174 dBm
Path loss threshold	−220 dBm
MIMO	2×2

The deep RL algorithms PPO and DQN models were constructed and trained on a random proposed environment where the UAV carries the eNB and flies around the cell to provide mobile services to the maximum number of UE's. The battery capacity of the UAV was initialized with a value E_{max}. The first scenario where the mobility of the UE's is considered, we illustrate the comparison results for applying PPO and DQN RL algorithms then tuned them with different learning rates (lr) values: $[0.01, 0.001, 0.0001]$.

The accumulative rewards for each iteration of the training is illustrated in Fig. 4, in this training results the PPO which was tuned wit learning rate 0.001 has the better performance than the other to solve the optimization problem. The other scenario in which we placed four UE's corners of the cell, Fig. 5 illustrates the accumulative rewards achieved in each training iteration using both model (PPO and DQN) which are in addition tuned with different learning rates (lr) $[0.01, 0.001, 0.0001]$. Comparing the performance of the RL algorithms showed that the PPO agent which tuned with $lr = 0.01$ or $lr = 0.001$ proves a superior performance than the others.

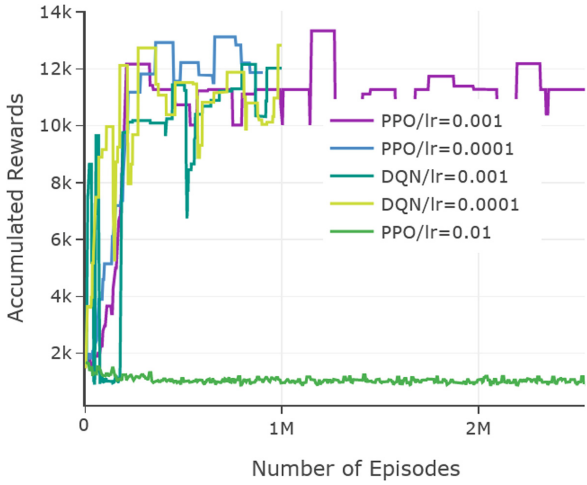

Fig. 4. Max reward per episode - 20 UE with random walk mobility model.

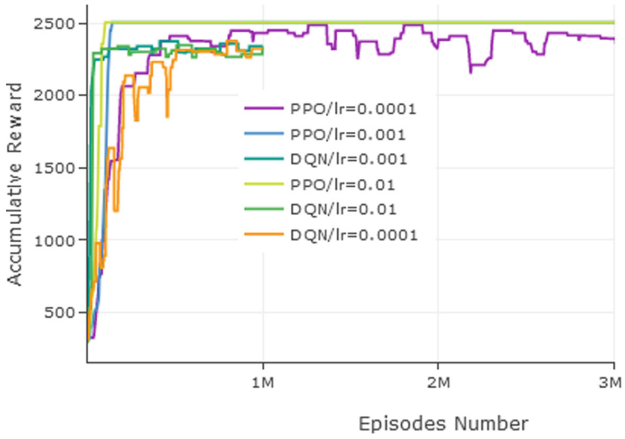

Fig. 5. Max reward per episode - 4 UE places at cell corners.

5 Conclusion

In this paper, we developed a framework for UAV autonomous navigation in urban environments that takes into account trajectory and resource allocation and the battery limitation of the UAV while taking into account the UE mobility within the environment. We deploy the RL PPO-based algorithm, which allows the UAV to navigate in continuous 2D environments using discrete actions, where the model was trained to navigate in a random environment. Then evaluated, the PPO and DQN algorithms while tuning the agents with different learning rate values, and then compared the results accordingly.

References

1. Zeng, Y., Zhang, R., Lim, T.J.: Wireless communications with unmanned aerial vehicles: opportunities and challenges. IEEE Commun. Mag. **54**(5), 36–42 (2016)
2. Mozaffari, M., Saad, W., Bennis, M., Nam, Y., Debbah, M.: A tutorial on UAVs for wireless networks: applications, challenges, and open problems. IEEE Commun. Surv. Tutor. **21**(3), 2334–2360 (2019)
3. Wu, Q., Liu, L., Zhang, R.: Fundamental trade-offs in communication and trajectory design for UAV-enabled wireless network. IEEE Wirel. Commun. **26**(1), 36–44 (2019)
4. Wu, Q., Mei, W., Zhang, R.: Safeguarding wireless network with UAVs: a physical layer security perspective. IEEE Wirel. Commun. **26**, 12–18 (2019)
5. Lin, X., et al.: The sky is not the limit: LTE for unmanned aerial vehicles. IEEE Commun. Mag. **56**(4), 204–210 (2018)
6. Zhao, H., Wang, H., Wu, W., Wei, J.: Deployment algorithms for UAV airborne networks toward on-demand coverage. IEEE J. Sel. Areas Commun. **36**(9), 2015–2031 (2018)
7. Sharma, N., Magarini, M., Jayakody, D.N.K., Sharma, V., Li, J.: On-demand ultra-dense cloud drone networks: opportunities, challenges and benefits. IEEE Commun. Mag. **56**(8), 85–91 (2018)
8. Zhang, Q., Mozaffari, M., Saad, W., Bennis, M., Debbah, M.: Machine learning for predictive on-demand deployment of UAVs for wireless communications. In: 2018 IEEE Global Communications Conference (GLOBECOM), pp. 1–6 (2018)
9. Chen, X., Hu, X., Zhu, Q., Zhong, W., Chen, B.: Channel modeling and performance analysis for UAV relay systems. China Commun. **15**(12), 89–97 (2018)
10. Zhang, G., Yan, H., Zeng, Y., Cui, M., Liu, Y.: Trajectory optimization and power allocation for multi-hop UAV relaying communications. IEEE Access **6**, 48566–48576 (2018)
11. Zhan, C., Zeng, Y., Zhang, R.: Energy-efficient data collection in UAV enabled wireless sensor network. IEEE Wirel. Commun. Lett. **7**(3), 328–331 (2018)
12. Xu, J., Zeng, Y., Zhang, R.: UAV-enabled wireless power transfer: trajectory design and energy optimization. IEEE Trans. Wireless Commun. **17**(8), 5092–5106 (2018)
13. Zhang, J., Zeng, Y., Zhang, R.: Spectrum and energy efficiency maximization in UAV-enabled mobile relaying. In: 2017 IEEE International Conference on Communications (ICC), pp. 1–6 (2017)
14. Qingqing, W., Zhang, R.: Common throughput maximization in UAV-enabled OFDMA systems with delay consideration. IEEE Trans. Commun. **66**(12), 6614–6627 (2018)
15. Yu, X., Xiao, L., Yang, D., Qingqing, W., Cuthbert, L.: Throughput maximization in multi-UAV enabled communication systems with difference consideration. IEEE Access **6**, 55291–55301 (2018)
16. Samir, M., Chraiti, M., Assi, C., Ghrayeb, A.: Joint optimization of UAV trajectory and radio resource allocation for drive-thru vehicular networks. In: 2019 IEEE Wireless Communications and Networking Conference (WCNC), pp. 1–6 (2019)
17. Guerra, A., Guidi, F., Dardari, D., Djurić, P.M.: Reinforcement learning for UAV autonomous navigation, mapping and target detection. In: 2020 IEEE/ION Position, Location and Navigation Symposium (PLANS), pp. 1004–1013 (2020)
18. Cui, J., Liu, Y., Nallanathan, A.: Multi-agent reinforcement learning-based resource allocation for UAV networks. IEEE Trans. Wireless Commun. **19**(2), 729–743 (2020)

19. Cui, J., Ding, Z., Deng, Y., Nallanathan, A., Hanzo, L.: Adaptive UAV-trajectory optimization under quality of service constraints: a model-free solution. IEEE Access **8**, 112253–112265 (2020)
20. Tang, F., Zhou, Y., Kato, N.: Deep reinforcement learning for dynamic uplink/downlink resource allocation in high mobility 5G HetNet. IEEE J. Sel. Areas Commun. **38**(12), 2773–2782 (2020)
21. Camp, T., Boleng, J., Davies, V.: A survey of mobility models for ad hoc network research. Wirel. Commun. Mob. Comput. **2**(5), 483–502 (2002)
22. Valavanis, K.P., Vachtsevanos, G.J.: Handbook of Unmanned Aerial Vehicles. Springer, Dordrecht (2014). https://doi.org/10.1007/978-90-481-9707-1
23. DJI matrice 600 prospecs. https://www.dji.com/hr/matrice600-pro/info#specs. Accessed 20 Mar 2023
24. Al-Hourani, A., Kandeepan, S., Lardner, S.: Optimal lap altitude for maximum coverage. IEEE Wirel. Commun. Lett. **3**(6), 569–572 (2014)
25. 3GPP. 5G, NR, User Equipment (UE) radio access capabilities. 3GPP TS, 15.3.0 edition (2018)
26. IWAVE airborne 4G LTE base station. https://www.iwavecomms.com/. Accessed 20 Mar 2023