Resource Allocation for UAV-assisted 5G mMTC Slicing Networks using Deep Reinforcement Learning

Rohit Kumar Gupta (1821cs16@iitp.ac.in)
IIT Patna: Indian Institute of Technology Patna https://orcid.org/0000-0002-1080-2651
Saubhik Kumar
IIT Patna: Indian Institute of Technology Patna
Rajiv Misra
IIT Patna: Indian Institute of Technology Patna

Research Article

Keywords: UAV, 5G Network Slice, Reinforcement Learning, Markov Decision Processes (MDP), Dueling-Deep-Q-Networks (DDQN)

Posted Date: May 4th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1485990/v1

Resource Allocation for UAV-assisted 5G mMTC Slicing Networks using Deep Reinforcement Learning

Rohit Kumar Gupta, Saubhik Kumar and Rajiv Misra

Department of Computer Science and Engineering, Indian Institute of Technology Patna, Patna, 801106, Bihar, India.

Contributing authors: 1821cs16@iitp.ac.in; 1801cs44@iitp.ac.in; rajivm@iitp.ac.in;

Abstract

The Internet of Things (IoT) application scenarios is becoming extensive due to the quick evolution of smart devices with fifth-generation (5G) network slicing technologies and hence IoT becoming significantly important beyond fifth-generation (B5G) networks. However, communication with IoT devices is more sensitive in disasters because the network depends on the main power supply and devices are fragile. In this paper, we consider Unmanned Aerial Vehicles (UAV) as a flying base station (BS) for the emergency communication system with 5G mMTC Network Slicing to improve the quality of service. The UAV-assisted mMTC creates a base station selection method with the aim of maximizing the system energy efficiency. Then, the system model is reduced into the stochastic optimization based problem using Markov Decision Process (MDP) theory. We propose a Dueling-Deep-Q-Networks (DDQN) based approach based on Reinforcement Learning (RL) technique for maximization of energy efficiency to solve the resource allocation problem. We compare the proposed model with DQN and Q-Learning models and found that the proposed DDQN based model performs better for resource allocation in terms of low transmission power and maximum energy efficiency.

Keywords: UAV, 5G Network Slice, Reinforcement Learning, Markov Decision Processes (MDP), Dueling-Deep-Q-Networks (DDQN).

1 Introduction

The rapid innovation and evolution of the fifth generation (5G) networking has been significant impact on standardization bodies academia and industry with extensive use cases supported by the 5G Network Slicing like self driving cars, augmented, mixed and virtual reality (AR/MR/VR), UAVs, etc [1]. The most important technology enablers of Network slicing technique are SDN and NFV. SDN, which decouples control plane with the data plane and provides programmability to network applications. NFV is known to be key technology to benefit industry virtualization by separating hardware network functions from primary hardware appliances [2, 3]. Network slicing is virtualization technique of multi-tenancy type where network functions are abstracted from software and hardware components. Each slice of network provides the specific functionalities covering by the core portions and RAN [4].

With the rapid development of remote mobile technology and innovation, the Unmanned aerial vehicles (UAVs) are becoming more intelligent with machine learning, which further broaden with the extensive UAV use cases. In the past few years, UAVs have been used widely due to high flexibility in traffic monitoring, aerial photography and disaster rescue [5]. The aim of the future generation networking is to provide user coverage capacity and ensure interconnection of everything [6]. In this perspective, since high mobility and versatility UAV features, the integration of communication module in the wireless network aspects UAVs can be occupied as aerial wireless access point for communication platforms in B5G [7]. This enables to create more flexible network in a new way for next generation communication, like UAV-assisted emergency communications. In disaster area using UAVs as access point for communication is effective choice in which continuously users can communicate to others in device-to-device (D2D) multi-cast manner [8]. In this at first, the disaster area can achieve fast response in this way then, UAVs can take advantage of line-of-sight (LoS) through air-to-ground (A2G) and ground-to-ground (G2G) links to access backhaul and access links to communicate and forward information [9]. Furthermore, ultra-dense networks (UDN) scenario are adopted in order to solve massive data computing, congestion and path loss problems and improve LoS with enhanced network capacity. The UDN is defined as massive Internet of Things (mIoT) slice or massive Machine Type Communication (mMTC) slice of 5G networks in which the uses as IoT devices can communicate with UAV-assisted framework.

However, the ground communication systems may suffer from equipment aging, human damage and natural corrosion which, affects the reliability and stability. In case of sudden breakdown situation, the base station (BS) may act as immobile with poor flexibility, which is not suitable for to meet the quality of experience (QoE) requirements of emergency communication systems [10]. The BS of wireless communication is main powered like cluster head (CH) in the D2D network. The emergency communication has been considered as the research hotspot towards high reliability and high flexibility. UAVs are likely to be important key component of 5G and beyond wireless communication networks that can support high bandwidth and facilitate with wireless broadcast. As compared with the fixed infrastructure communications. UAVs have some salient attributes, such as strong LoS connection links, flexible deployment and controlled mobility with additional design degrees [11]. UAVs can be used as wireless aerial base stations to enhance reliability, capacity, coverage and the energy efficiency of wireless communication also, UAVs can be operated as the flying mobile terminal within a wireless mobile network [12]. The future networks also promise massive IoT and ubiquitous coverage other than low latency and extremely high data rates, flying UAV's base stations could be back-up fast to the wireless communication networks, prevent service interruption and enhance the network performance [13]. Furthermore, The study of resource allocation, data transmission, power control and energy efficiency for emergency communication in UAV-assisted 5G which aim to solve resource allocation in disaster area, improve quality of user experience and enhance quality of user communication.

As diverse application advantages such as mobility and flexible deployment, UAVs can carry base stations that can provide temporary communication and combined with existing cellular networks can serve ground users which attracted a lot of attention recently [14]. In this work, trajectory optimization problem which aim to solve transmit power allocation problem by considering the maximum transmit power and minimum user date rate. However, the energy consumption factor and the stability should also be taken in account. These application should also be applicable into 5G communication or next generation 6G networks. The work [15], present UAV based 5G in emergency communication scenario with ability of coverage uplink, downlink, scheduling, mobility, which makes the faster response time and more flexible with compare to traditional networks.

The heterogeneous UDN is known to be promising architecture for coping with massive data traffic in 5G and beyond networks. Resource allocation need to address to meet quality-of-services (QoS and to ensure the service level agreement (SLA) in 5G networks [16, 17]. In order to efficiently manage wireless resource in heterogeneous UDN faces challenges to reduce wireless interference [18]. The mMTC based dense network deploys low-power and short-range type of large number of small cells coverage to macro base stations, and it enables flexible and high throughput services with the 5G networks [19]. The UAV-assisted future networks have accomplished interest in various applications ranging from increase in airborne flight time, robust maneuverability, rapid deployment, payload capabilities, are highly influenced by technologies such as machine learning, artificial intelligence, reinforcement learning, SDN and mobile edge computing (MEC) [20]. The 5G UDN has been known as the key solution for consequential energy consumption and sudden increase in the mobile traffic. Further, popular content caching at the edge in mMTC can resolve the challenges of energy and traffic [21]. However, UAV-assisted 5G mMTC networks brings new challenge for maximization of system energy efficiency (EE) to ensure overall quality of user experience .

To address the above mentioned challenges, we study the optimization problem of UAV-assisted resource allocation in the 5G mMTC slice system. We first design the link selection method which solve base station selection problem. Then, the energy efficiency maximization problem with low transmission power is proposed which aim to optimize the resource allocation strategy. Although the original problem is reduced into MDP concept to make it stochastic optimization based problem and Deep reinforcement learning (DRL) approach is used to solve [22]. The generated massive data can be process efficiently by DRL in mMTC slice as compare to the traditional machine learning techniques and DRL is more consistent to investigation of this study. The main contributions of this paper are as follows.

- We investigate the UAV-assisted resource allocation problem in 5G mMTC Slice Network system. UAVs has been used as flying base station to deal with emergency communications in which flying base station communicate with ground base stations so that users can continuously communicate to others in device-to-device (D2D) multi-cast manner in order to increase quality of experience. As compared with fixed infrastructure communications, UAVs have some salient attributes, such as strong LoS connection links, flexible deployment and controlled mobility with additional design degrees.
- To optimize the space-ground emergency communication system performance, we considered the power allocation and link selection problem for better system performance. We design the base station selection method based on the transmission power and distance of the user with the help of locations of UAV and base station. The user selects best base station for the data transmission based on optimal path which can improve the overall quality of user communication.
- We create scenario for UAV-assisted 5G mMTC slice for the feasibility of this scheme. The problem of maximizing energy efficiency is reduced into MDP. We propose Dueling-DQN (DDQN) based dynamic resource management algorithm based on reinforcement learning to solve MDP problem. We also discuss the Q-Learning and DQN models to compare with proposed DDQN based model for comprehensive study.
- We conduct the extensive experiments to evaluate the system performance of UAV-assisted 5G mMTC network slice scenario for emergency communication with reinforcement learning approaches. The simulation performance determine that the resource allocation based on Dueling-DQN (DDQN) scheme can improve the system energy efficiency than Q-Learning and DQN model. Further, we also investigate the relationship between the users, base stations and system energy efficiency with Q-Learning, DQN and proposed model.

The current section is introduction of this paper and the remaining part of this paper is organized as follows. In the section 2, we discuss literature work which is related work to our system. The system model is described in the session 3 which, describes channel model, network model and computational model. The section 4 formulate the problem and transform into MDP technique which is solved by reinforcement learning and DQN techniques. Then we present the performance evaluation and analysis in the section 5. Finally, we conclude this paper in section 6 with the future scope.

2 Related Work

In the literature, continuous and sincere efforts have been made towards network performance improvement in response to resource allocation strategies such as game theory and deep-Q networks. In the 5G networks, UAV assistance and resource allocation have great significance that continuously got more attention in 5G development. The UAV's application in 5G networks includes energy transmission and communication buffering with the aim to improve system flexibility. For instance, the authors in [23] investigates multi-optimization for resource management and massive access in mMTC slice scenario by UDN environment throughput maximization problem of machine learning based user clustering method. The paper [24] studied CPU-cycle and stochastic task allocation for MEC system to aim for energy consumption minimization with upper bounded task queue length and proposes stochastic based optimization algorithm.

The paper [25] proposes MEC-UAV cooperation on task offloading with aim to minimizing energy consumption by using reinforcement learning approach. The authors in [26] aim to extend 5G network slicing using UAV with MEC by proposing reinforcement learning based power optimization to maximize objective functions. In paper [27] studied cooperative MEC wireless-powered UAV supported system, optimizing energy consumption and trajectory of UAV-assisted system. In [28] authors investigate the scenario of UAV-alone and UAV with fixed base stations in which aim to load balance maximization and minimization of number of UAVs. The authors in [29] studied base station bandwidth allocation and UAV access selection problems with game framework in IoT communication UAV assisted network. In [30], the joint power splitting and precoding vectors optimization technique is proposed which aim to secure UAV-assisted NOMA networks. The paper [31] proposes a multi UAV-assisted NOMA scheme for better spectrum and energy efficiency of uplink cellular communication system. The work [32] jointly optimize bandwidth allocation, power transmission and UAV route to maximization of energy efficiency to meet difference quality-of-experience (QoE) requirements. The paper [33] proposes to utilize relay node and computing node to improve use latency in UAV-assisted MEC networks.

In addition, considering intelligence as an main characteristic for the future wireless communication, many work have been investigated recently. The UAV-enabled networks in which flying object is uses as access point of given flight period, seek to maximise the common throughput across the ground users [34–36]. In the paper [37], to address the massive connected devices UAV has used as air base station which aim to increase the performance by improving placement and power allocation. In this paper [38], two UAVs are used, one moves to interact with various users on the ground, while the other jams eavesdroppers to safeguard the desired users' communications. While the work based on UAV-enabled network system model in which a UAV is used to connect and communication with ground node while some jammers are present that aim to optimize the trajectory of UAV and maximize energy efficiency [39]. The paper [40] studies UAV's line-of-sight in 3D space for air-to-ground communication and proposes UAV-assisted data collecting strategy for gathering information which aim to reduce the total time by optimising the UAV's altitude, velocity, trajectory and data linkages with ground users. The deployed UAVs are used to collect sensor data in the work [41], in which authors aim to optimize sensor node in 3D trajectory.

The cellular based network system with UAV-mounted base station with numerous terrestrial based base stations is considered in the article [42], with each base station servicing multiple customers. Further, designing the 3D placement for the aerial base station and the transmission power allotment for all nodes in the uplink and downlink in a probabilistic channel based environment have been done. The problem of drone BS placement with resource allocation in a certain hotspot area is investigated in paper [43] to ensure QoS for users within the hotspot area. The paper [44] investigates UAV-assisted networks in which multiple UAV has been considered as the base stations. In this scenario UAV-BS connects with the ground BS and ground BS forward the data between user and core networks and proposes the framework to maximise user overall throughput while maintaining fairness among users within the flight-time of base stations.

These are some of the work that are close to our work, the paper [46], authors aim to power optimization by maximizing social group utility with user downlink and QoS being satisfied. The UAV-supported communication can provide the better coverage in the remote areas with capability to manage high traffics. In case of damage of the terrestrial network UAV can be used for the emergency communication. The authors in [48] discuss the UAV-assisted based emergency communication model in heterogeneous IoT and distributed NOMA scheme. The IoT has problem of spectrum resource due to widespread and power consumption due to battery powered. In [47], authors aim to maximize the energy efficency of UAV-assisted UDN networks. Through the flexible deployment of the system UAV can be used as the flying base station when any disaster occurs. The authors in paper [49], study on mMTC based UAV-assisted optimization for energy efficiency.

3 Preliminaries and System Model

3.1 System and Channel models

We have considered a natural disaster scenario where ground base station is not fully functional and UAV have been used as flying base station as emergency communication. As shown in the Fig. 1, the disaster can occurs suddenly within the wireless network-zone in which the torn-down scope fall out which is described as dotted circular region. The heterogeneous IoT has been considered in this region, where all the devices and users work with batteries and use to communicate in the network. Moreover, we also assume that some of the base stations and cluster heads are down due to no battery back ups as it is main powered. In this case, an UAV-assisted BS is deployed for message transmission between the disaster region devices to the workable outside-disaster BSs. In order to deliver messages from the base station to disaster region, the UAV's of outside the disaster region should relay on UAV's above. The terminal UAV's could communicate with the devices and cellular users within the coverage area. Further, the users within the disaster region can communicate with the other disturb region and outside users.

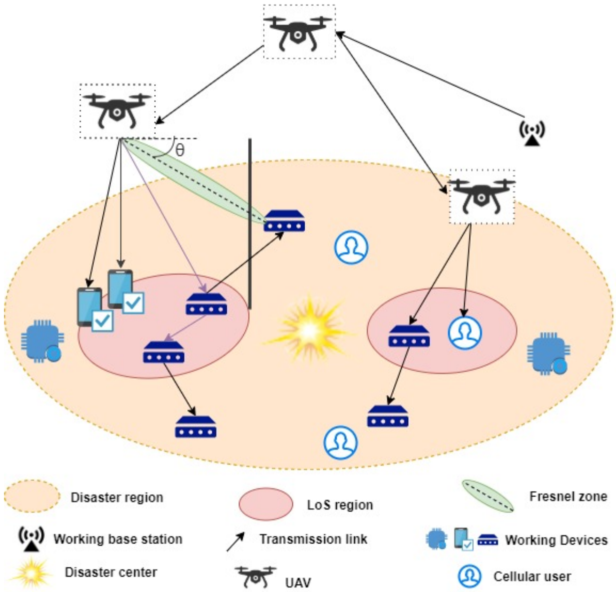

Fig. 1 An overview of UAV-assisted 5G based emergency communications

The links of LoS region are preferred that can be easily obtained within the UAV coverage. The LoS link depends on θ angle that is the elevation of the UAV's to reach signals [27]. If the value of θ will be small then the probability for UAV trajectory to transmitting signals with LoS links will be low. Thus, we define and say θ^{TH} is the threshold of LoS-based elevation angle. If suppose the angle of signal is less than θ^{TH} then, no valid link is defined from UAV to users on the ground. Hence, the coverage of UAV region with valid LoS links can be find if the height and the location of UAV are known as illustrated in Fig. 1 enclosed by the solid circles .

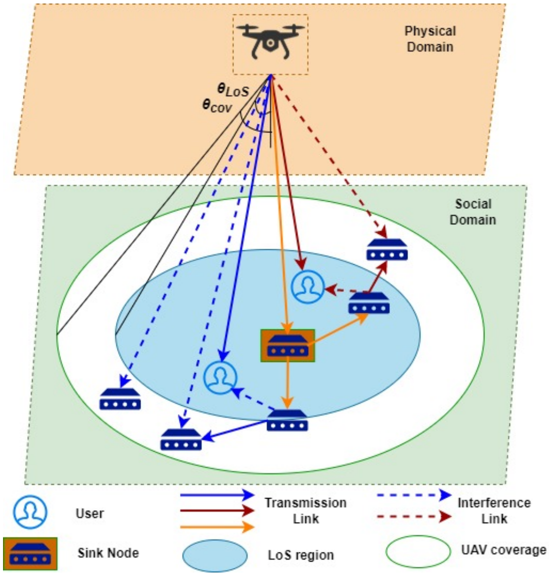

Fig. 2 The coexisting links with UAV-assisted communication model

In the case of emergency communication, many messages need to be urgently delivered and exclusive spectrum allocation for devices and users is inefficient. Hence, we accommodate to use the same spectrum between the links of D2D outside of LoS region and between the links of UAV and the user or devices. Fig. 2 shows the Model of the coexisting links surrounded by single UAV in disaster region. θ_{LoS} and θ_{cov} are denoted as the maximum angle to the LoS region and UAV coverage region from perpendicular line in order to denote the UAV coverage and line-of-sight (LoS) region for the user. It is likely to be that θ_{LoS} is the complementary angle of θ^{TH}.

Furthermore, it is condonable for assuming that at least one device transmitter is there that has located within the UAV's LoS coverage region. Therefore, this transmitter is not able to allocate the spectrum to the devices and users. Since the air-to-ground (A2G) channel is more suitable for the existing device in the LoS region of UAV, this device could be used as the sink node of the IoT. By using of the relay in sink node, the messages from the base station can be disseminated to the devices. As a result, suppose the unit bandwidth in one subchannel is set as B_0, while total number of subchannels available to these devices and users for sharing is assumed to be Z^T, then the corresponding power gain of the links can be computed using equations

(1–4). We used the notations for channel model as given in work [27]. $G^{A2C}_{u,z}$ and $G^{D2D}_{v,z}$ are the power gains for the transmitting links from the UAV to the user u and between the device pair v on subchannel z, respectively. The power gains from the UAV to the receiving device and from the transmitting device to the cellular user on subchannel are $g^{A2D}_{v,z}$ and $g^{D2C}_{u,v,z}$.

$$G^{A2C}_{u,z} = \rho_{(LoS)} \cdot d^{(-\alpha)}_{u,A} \tag{1}$$

$$G^{D2D}_{v,z} = d^{(-\beta)}_{v} \cdot \|H_z\|^2 \tag{2}$$

$$g^{A2D}_{v,z} = \rho_{(NLoS)} \cdot d^{(-\alpha)}_{v,A} \tag{3}$$

$$g^{D2C}_{u,v,z} = d^{(-\beta)}_{u,v} \cdot \|H_z\|^2 \tag{4}$$

The distance between UAV to user is denoted by $d_{u,A}$, and the user and UAV distance is marked by d_v between the device pair. Further, the distance between the UAV and the receiving station must be considered. The distance between the receiving device and UAV is $d_{v,A}$, and $d_{u,v}$ is distance of transmitting device u and the user v. To define the gain in LoS transmission of messages from the UAV to cellular users and the NLoS gain of the UAV's interference to the ground station, on the same spectrum as the device receiver, $\rho_{(LoS)}$ and $\rho_{(NLoS)}$ are used. Additional attenuation factors are denoted for the sake of simplicity, $\rho_{(LoS)}$ is normalised to 0 dB. α and β stand for route loss factors in A2G and G2G are two different types of channels. Based on Rayleigh distribution the multipath-caused frequency channel gain is H_z on a tiny scale for G2G transmissions on subchannel z.

3.2 Network Model

We consider the 5G mMTC system with UAV-assisted in which there are set of base stations followed by UAV. The UAV is considered as the flying base station within the system. A centre access point (CAP) is positioned at centre of the system as depicted in Fig. 3 which is used to manage base stations and UAV. The set for UAV is represented to A and there is only one element a in A, such that $A = \{a\}$. The set of base station is defined as B such that $B = \{1, 2, .., b, .., B\}$. The set for users in the mMTC network is represented to U, such that $U = \{1, 2, .., u, .., U\}$. In the system, the users have been assigned random to the base stations and UAV. Further, we represent $F = A \cup B$ as total feasible set. Consider F_u as the set of feasible for the user u, which is represent as $F_u = \{f_a, f_b\}$, where f_a defines that the the user u select UAV a for data transmission and f_b defines that the user u select base station b for data transmission. Further, the notation of symbols and detailed description is provided in the Table 1.

Table 1 Notations and description

Notation	Definition
A, B, U	set of UAVs, base stations, users, respectively
F, F_u	a feasible policy set of all user, user u, respectively
c_u	users connection status
P_a, P_b, P_c	location of UAV, base station, user, respectively
$H^t_{u,a}$, $H^t_{u,b}$	channel gain from UAV to user and from BS to user
r_{bs}	base station initial channel gain
r_{uav}	channel gain of UAV within 1 meter
α	path loss index from user to base station
β	path loss index from user to UAV
E_{user}	maximum power received by user
$E^t_{u,a}, E^t_{u,b}$	assigned power to user by UAV and base station
$\mathcal{B}_a, \mathcal{B}_b$	bandwidth of UAV and base station
$\omega(u)$	differentiating connection coefficient for users
σ_2	additive-white-gaussian noise (AWGN)
$G^{A2C}_{u,z}, G^{D2D}_{v,z}$	power gains of UAV to the user and between device pair
$g^{A2D}_{v,z}, g^{D2C}_{u,v,z}$	interference power gains of UAV to device and device to cellular
$\rho_{(LoS)}$, $\rho_{(NLoS)}$	attenuation factors of line-of-sight and outside

We consider the system model in which center access point can capture a variety of environmental data within the system. The center access point then chooses the best base station for the given user based on data-link selection method and manages the base station's resource allocation policy. In case of UAV-assisted, the user directly get connected with the UAV and UAV is connected to the nearest center access point so that user can connect with outside and within region in order to data transmission as shown in Fig. 3.

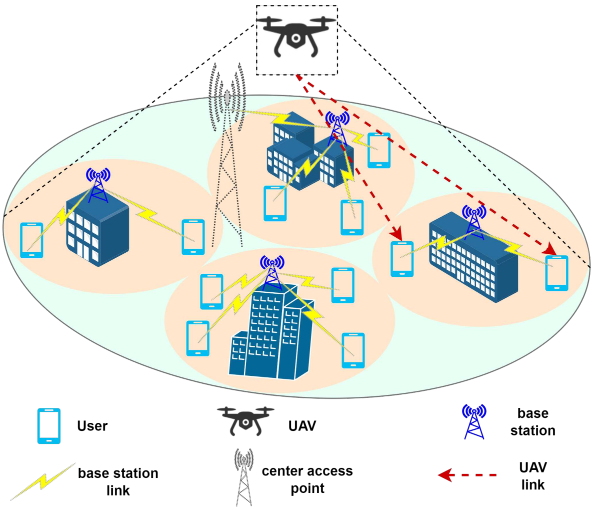

Fig. 3 The UAV-assisted 5G mMTC slice model

3.3 Computational Model

In order to represent communication connection status in better way, we denote a function c_u, where $c_u \in F_u$ which express the user u connection with base station b or UAV u. We assume that the base station location is fixed because of the UAV is staying in the air. Furthermore, we define the UAV location as $P_a = [x_a, y_a, z_a]$, where x_a, y_a and z_a is coordinate of UAV a . The base station location is defined as $P_b = [x_b, y_b, z_b]$, where x_b, y_b and z_b is coordinate of base station a. The users can be randomly distributed all over the system and the location of user u is defined as $P_u = [x_u, y_u, z_u]$, where x_u, y_u and z_u is the coordinate of user u. Channel gain is defined as the tight upper bound on the rate at which information can be transmitted with high reliably over a communication channel. The channel gain has mainly the relation between the devices and base station. The channel gain at time slot t from base station b to the user u is defined as

$$H_{u,b}^t = r_{bs} \| P_b^t - P_u^t \|^{-\alpha} \tag{5}$$

where $\|P_b^t - P_u^t\|$ is the distance between the base station b and user u at the time slot t, α represents the path loss index from user to base station and r_{bs} is the rayleigh fading factor i.e. initial channel gain. Similarly, the channel gain at time slot t from UAV a to the user u is defined as

$$H_{u,a}^t = r_{uav} \| P_a^t - P_u^t \|^{-\beta} \tag{6}$$

where $\|P_a^t - P_u^t\|$ is the distance between the UAV a and user u at the time slot t, β represents the path loss index from user to UAV and r_{uav} is the channel gain of UAV within 1 m reference distance.

$$\begin{aligned} &\frac{E_a}{E_b} \cdot d_b^u > d_a^u, \; c_u = f_a \\ &\frac{E_a}{E_b} \cdot d_b^u < d_a^u, \; c_u = f_b \end{aligned} \tag{7}$$

where, E_a and E_b are the power transmission of the UAV to user and base station to user, respectively. The distance between UAV a to user u and base station b to user u is defined as d_a^u and d_b^u, respectively. The distance between UAV a to user u and the distance between base station b to user u is described as

$$\begin{aligned} d_a^u &= \|P_a - P_u\| \\ d_b^u &= \|P_b - P_u\| \end{aligned} \tag{8}$$

When the product $\frac{E_a}{E_b} \cdot d_b^u$ is greater than the distance from the user u to the UAV a, then the user can access to the base station b and when the product $\frac{E_a}{E_b} \cdot d_b^u$ is less than the distance from the user u to the UAV a, then the user can access to the UAV a. Based upon the connection status of the user, we derive the signal-to-interference and noise ratio (SINR) for the each of the user. The SINR in the downlink from fron UAV a to the user u at time slot t is defined as

$$SINR_{one} = SINR_{u,a}^{t} = \frac{E_{u,a}^{t} \cdot H_{u,a}^{t}}{\sum_{v \in U, v \neq u} E_{u,a}^{t} \cdot H_{u,a}^{t} \cdot c\{c_v = c_u\} + \sigma^2}, \tag{9}$$

where $E_{u,a}^{t}$ denotes the assigned power transmission by the UAV a to the user u and $H_{u,a}^{t}$ is the channel gain from user u to the UAV a, the user connection status is determine by the indicator function $c\{c_v = c_u\}$ and σ^2 is the additive white Gaussian noise (AWGN). The SINR in the downlink from fron base station b to the user u at time slot t is defined as

$$SINR_{two} = SINR_{u,b}^{t} = \frac{E_{u,b}^{t} \cdot H_{u,b}^{t}}{\sum_{v \in U, v \neq u} E_{u,b}^{t} \cdot H_{u,b}^{t} \cdot c\{c_v = c_u\} + \sigma^2} \tag{10}$$

where $E_{u,b}^{t}$ denotes the assigned power transmission by the base station b to the user u and $H_{u,b}^{t}$ is the channel gain from user u to the base station b. In order to calculate the channel capacity or data rate (DR) we use Shannon's channel capacity formula. The data rate in the downlink from user u to the UAV a at a time slot t is defined as

$$DR_{one} = DR_{u,a}^{t} = \frac{E_{u,a}^{t}}{E_{max}^{a}} \cdot \mathcal{B}_a \cdot log_2(1 + SINR_{one}) \tag{11}$$

where E_{max}^{a} is the maximum transmission power of the UAV a and $\mathcal{B}_a$ is the bandwidth of the wireless channel of the UAV a. Similarly, the data rate in the downlink from user u to the UAV a at a time slot t is defined as

$$DR_{two} = DR_{u,b}^{t} = \frac{E_{u,b}^{t}}{E_{max}^{b}} \cdot \mathcal{B}_b \cdot log_2(1 + SINR_{two}) \tag{12}$$

where E_{max}^{b} is the maximum transmission power of the base station b and $\mathcal{B}_b$ is the bandwidth of the wireless channel of the base station b. As a result, the total data transmission rate of the UAV and the base station is defined as

$$DR_{total} = \sum_{u \in U} \omega(u) \cdot DR_{one} + \sum_{u \in U} (1 - \omega(u)) \cdot DR_{two} \tag{13}$$

where $\omega(u)$ is the use for differentiating connection indicator. $\omega(u) = 0$ shows user u is connected to the UAV a. In other way, $\omega(u) = 1$ shows user u is connected to the base station b. The total power consumption E_{total}^{t} contain the power consumption $E_{u,a}^{t}$ to the UAV and power consumption $E_{u,b}^{t}$ of base station. In which the power of UAV and base station can be affected by user link mode. Therefore, at time slot t the total power consumption can be defined as

$$E_{total} = \sum_{u \in U} \omega(u) \cdot E^t_{u,a} + \sum_{u \in U} (1 - \omega(u)) \cdot E^t_{u,b} \tag{14}$$

By using the equations (13) and (14), Energy Efficiency (η) of the system is given as

$$\eta = \frac{DR_{total}}{E_{total}} \tag{15}$$

The resource allocation problem in the UAV-assisted 5G mMTC network slice systems can be given as follows, as mentioned in the study above.

$$\begin{aligned} &\max_{\omega(u)} \ \eta, \\ &\min \sum_{u=1}^{U} E_v \end{aligned} \tag{16}$$

s.t.

$$\begin{aligned} &C1 : \sum_{u \in U, c_u = b} E^t_{u,b} \leq E^b_{max}, \ \forall u \in U, b \in B \\ &C2 : \sum_{u \in U} E^t_{u,a} \leq E^a_{max}, \ \forall u \in U, a \in A \\ &C3 : \omega(u) = \begin{cases} 0, \ c_u = f_a \\ 1, \ c_u = f_b \end{cases}, \ \forall u \in U \\ &C4 : E^t_{u,a} \leq E_{user}, \ E^t_{u,b} \leq E_{user}, \forall a \in A, b \in B \end{aligned}$$

where, C1 exhibit the limit of power transmission from ground base station to the associated user within the coverage region. The limit of transmission power from the UAV to user within the coverage region is represented by C2. In C3, $c_u = 0$ indicates that user u is linked to the ground base station, while $c_u = 1$ indicates that the user u has been connected to the UAV a. We denote the received maximum power capacity by the user as E_{user} in C4, in which the user receive at max E_{user} from the UAV or from the groud base station. The overall energy efficiency of the system should be maximize with the minimum transmission power supplied. The UAV-assisted mMTC system ensures the user connectivity and data transmission to achieve the overall quality of user experience.

Lemma 1: The path loss from UAV to the ground devices can be defined as $g_{u,a} = Pr(LoS) \cdot d_{u,A}^{(-\alpha)} + Pr(NLoS) \cdot \rho_{(NLoS)} \cdot d_{v,A}^{(-\alpha)}$, $u \in U$, where $Pr(LoS)$ and $Pr(NLoS)$ are the probability of line-of-sight and outside connections respectively, $\rho_{(NLoS)}$ is the additional attenuation factor.
Proof:

$$Pr(LoS) = \frac{1}{1 + S \cdot exp(-R[\theta_i - S])} \tag{17}$$

where S and R are dependent parameters which is dependent on environments (such as urban, rural or dense urban), θ_i is the angle of evaluation between user and UAV. Here, θ_i can be calculated as $\theta_i = \frac{180}{\pi} tan^{-1}(\frac{height}{radius_i})$, in this $radius_i = \sqrt{d_{u,A}^2 - height^2}$ [45]. According to above equation (17), the probability of line-of-sight (LoS) region increase when the angle of evaluation increases. So, the probability of Non-line-of-sight (NLoS) region could be represented as $Pr(NLoS) = 1 - Pr(LoS)$. Fundamentally, the UAV altitude defines evaluation angle and signal-propagation distance so, both NLoS and LoS depletion jointly impact the path-loss within UAV to the ground devices [46]. Hence,

$$g_{u,a} = Pr(LoS) \cdot d_{u,A}^{(-\alpha)} + Pr(NLoS) \cdot \rho_{(NLoS)} \cdot d_{v,A}^{(-\alpha)}, \; u \in U. \tag{18}$$

Lemma 2: The total sub-channels from UAV to the ground users and devices can be represented as $Z_{total}^t = \frac{E_{total} \cdot Z^T}{\sum_{l=1}^{U} E_l}$ in which, the total number of available sub-channel is assumed as Z^T.
Proof: According to the channel model assigning sub-channels to users can be simplified into assign bandwidths, which is identical to finding the needed number of sub-channels for each of the user in the system because transmission gain is only directly related to the path loss. As a result, maximizing the least of DR_u^t / E_u^t is the same as satisfying equation (19).

$$\frac{DR_u^t}{E_u^t} = \frac{DR_{u'}^t}{E_{u'}^t}, \quad u \neq u', \quad \forall\, u, u' \tag{19}$$

As a result, if the available power is sufficient,

$$x_{u,z}^t = x_{u',z'}^t = x_{eq}^t, \quad u \neq u', z \neq z' \quad \forall\, u, u', z, z' \tag{20}$$

Based on Shannon's capacity theory, the user's sum rate is defined as follows,

$$DR_{total}^t = \sum_{z=1}^{Z^T} \sum_{u=1}^{U} \left[k_{u,z}^t \cdot \mathcal{B} \cdot log_2(1 + x_{u,z}^t) \right] \tag{21}$$

$$= \mathcal{B} \cdot log_2(1 + x_{eq}^t) \cdot \sum_{z=1}^{Z^T} \sum_{u=1}^{U} k_{u,z}^t$$
$$= Z_{total}^t \cdot \mathcal{B} \cdot log_2(1 + x_{eq}^t) \tag{22}$$

where x_{eq}^t is a SINR benchmark for all users on each sub-channel as a result, (19) becomes,

$$\frac{Z_{total}^t}{E_{total}} = \frac{Z_{total'}^t}{E_{total'}} \tag{23}$$

To enhance data rates for all users equally, a simple and intuitive way is to start the random sub-channel assignment with the entire sub-channel number as the starting point, so the total sub-channels from the UAV to the users and devices can be defined as

$$Z_{total}^t = \frac{E_{total} \cdot Z^T}{\sum_{l=1}^{U} E_l}. \tag{24}$$

Lemma 3: The minimum number of subchannel for device pair can be given as $Z_v^{DD} = \frac{DR}{\mathcal{B} \cdot log_2(1+sinr_v^D)}$,
Proof: According to the device receivers' quality of service (QoS) restrictions

$$sinr_{v,z}^{DD} \geq sinr_v^D \tag{25}$$

The sum data rate of device to device of sub-channels will be grater than or equal to the total data rate at the users to meet the quality of experience,

$$\sum_{z=1}^{Z_T} DR_{v,z}^{DD} \geq DR_v^D \tag{26}$$

Based on Shannon's capacity theory, the user's sum data rate is defined as follows,

$$DR = \sum_{v=1}^{U} \left[x_v^{DD} \cdot \mathcal{B} \cdot log_2(1 + sinr_v^D) \right] \tag{27}$$

Hence, the number of subchannels assigned can be reduced. The minimum number of required subchannels can be calculated using (25), (26) and (27) as

$$Z_v^{DD} = \frac{DR_v^D}{\mathcal{B} \cdot log_2(1 + sinr_v^D)}. \tag{28}$$

4 Problem Formulation

The problem of allocating resources with UAVs goes beyond conventional optimization techniques due to the huge number of users in 5G mMTC systems. Considering the discontinuous nature of data the traditional resource allocation algorithms cannot quickly find an appropriate solution to allocate spectrum and power. In order to solve resource allocation problem in our system, first we reduce our problem to Markov-Decision Process (MDP) technique. The MDP problem can be easily solve by the reinforcement learning algorithms. We propose a Dueling-DQN based algorithm for resource allocation for the maximization of overall system energy efficiency. The DDQN technique is the type of reinforcement learning technique which is based on improved version to Q-Learning algorithm. We also consider the Q-Learning and DQN models to compare with the DDQN model. The DDQN technique of reinforcement learning does not rely on prior knowledge and it can successfully address the data explosion problem generated by a vast state and action space. It can also solve the resource allocation problem to maximize energy efficiency. Hence, we use the reinforcement learning based DDQN model in UAV-assisted 5G mMTC system for resource allocation.

4.1 MDP Model

Considering a system model of UAV-assisted 5G mMTC, for the large action space the reward as feedback should generate instantaneously. The characteristics of MDP is more suitable for the state, reward and action states. The markov property is stated as "The future is independent of past for the given current state". In this model the decision maker which is agent has surrounded by the environment. The environment provides a reward on each of the next state based on the agent's action. Here, we define the state space, action space and reward function below.

- **State space:** The space state indicates agent position in the environment at a given time stamp t per user. Further, the updated space state t+1 is given to the agent for further action. We have allocated different data rates to each users since the data-transmission rate is variable. The state space for the given time t is given as:
  $s_t = \left\{DR(1), DR(2), .., DR(U), E_{total}\right\}^t$

- **Action space:** To maximize data transfer rates for the users the base station and UAV power allocation should in control. The action space are the learners in the reinforcement learning and the agent interact with the environment through the actions. The action state is given as $a_t = \left\{E(1), X(1), E(2), X(2), .., E(N), X(N)\right\}^t$ and each action corresponds to a state. where,
  $e_u^t = E_{min}(\frac{E_{max}}{E_{min}})^{m/(|D|-1)}$

- **Reward function:** In our system the reward function has calculated based upon the total rate of data transmission. The reward function is given as: $r_t = \sum_{u=1}^{U} \omega(u) \cdot DR_{u,a}^{t} + \sum_{u=1}^{U} (1 - \omega(u)) \cdot DR_{n,b}^{t}$.

Our aim is to solve the optimization problem of the resource allocation in which our task is to maximize the energy efficiency with the given lower data transmission rate.

4.2 Q-Learning

The reinforcement learning contains state, action and reward in which the agent interact to the environment with the action state. The environment sends the next reward and space states to the agent. The Q-Learning is based upon off policy reinforcement learning in which agent takes best action based on the current state. We can say that in order to receive maximum reward agent choose best action. RL can be categorised into a variety of categories depending on the problem like model based (the environment can be understood by the actor), model-free on-policy (same actors for interacting as well as learning), model-free off-policy (different actors for learning and interacting). The dynamic resource allocation algorithm based on Q-Learning is given in algorithm 1 in which δ is the learning rate, $\gamma =$ is the discount factor, E_{min} is minimum transmission power, H_u, H_a are channel gain and σ^2 is environment noise.

Algorithm 1 Dynamic Resource allocation algorithm based on Q-Learning

Require: δ, γ, E_{min}, σ^2, H_u, H_a
Ensure: Optimal resource allocation strategy

State value function and Q-learning parameter initialization;
Reward R initialization;
for *episode* $\leftarrow$ 1 to N **do**
  for `each step of` *episodes* **do**
    Initial system state selection randomly from the state space S;
    An action a_t is chosen from set of possible allocations actions from the current action space A;
    Compute action a_t for observation of reward r_t and next state s_{t+1};
    The parameter updation of $Q(s_t, a_t)$;
    When the expected state is reached, terminate.
  end for
end for
The optimal resource allocation and strategy for power distribution is obtained to get maximized of the energy efficiency of the system

The Q table value decide the next reward to take best action by the agent. In order to energy efficiency maximization the action state is given as
$a_t = \arg\max Q_t(s_t, a_t), a_t \in \pi$
The Bellman equation is used to continually update the value of Q given by the following equation,
$Q_{t+1}(s_{t+1}, a_{t+1}) = Q_t(s_t, a_t) + \delta(R + \gamma \max_{a'_t \epsilon A} Q_t(s'_t, a'_t) - Q_t(s_t, a_t))$.
Here, δ denotes the learning rate and γ denotes the discount factor. Q value is $Q_t(s_t, a_t)$ for the time slot t.

4.3 DQN Model

The Deep-Q Network model is combination of deep learning and reinforcement learning models to take more advantage of the decision making.

Algorithm 2 Dyananmic Resource allocation algorithm based on DQN

Require: δ, γ, E_{min}, σ^2, H_u, H_a
Ensure: Optimal resource allocation strategy

DQN parameters, experience buffer R and state value function $Q(s, a)$ initialization;
The Parameter of CNN θ is initialised with different weights;
for *episode* $\leftarrow 1$ to N **do**
    initialize system state;
    for `each time from` $[1, T]$ **do**
        An action a_t is chosen ranomly by epsilon greedy approach
        Compute state-action pair and observe the reward r_t and next state s_{t+1};
        Save the transitions (s_t, a_t, r_t, s_{t+1}) in R;
        if size of R has reached its maximum capacity **then**
            Select a batch of transitions (s_i, a_i, r_i, s_{i+1}) from R at random
        end if
        if next state $s_{t+1} == s_M$ **then**
            change state value function and reward is observed
        else
            $L(\omega) = \mathbb{E}[(\mathrm{y}_t - Q(s_t, a_t;\ \theta))^2]$
            update the parameter of CNN θ after each c iteration
            update the policy equation to be written
        end if
    end for
end for
The optimal power allocation strategy π^* is obtained

RL and deep learning is combined by DRL, maximizing the advantage of RL which includes decision-making with the perceiving benifits of deep learning.

As a result, agents are able to notice more complicated environmental circumstances and develop more sophisticated action plans. We employ framework DQN which does not depend on the DRL framework in order to solve the MDP problem which does not require any prior knowledge.

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[r_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t)]$$

4.4 DDQN Model

The DDQN is a type of Q-Learning that contains two streams to separate estimator in which both streams shares common convolutional module. Dueling DQN is an extension of DQN. It uses two CNNs in order to get better policy evaluation in the presence of many similar-valued actions. The first network is the primary network which is used for selecting an action. The second network, the target network is used for generating a Q-value for the action. During training, the target-Q values are used to compute the loss function for every action. Target network is basically a copy of primary network and its weights remain fixed. The weights of target network are updated after some constant number of iterations, which is initialised as hyperparameter. During training of the model, the benefit of experience replay is used and the tuple (s_t, a_t, r_t, s_{t+1}) is stored in it.

Algorithm 3 Dynamic Resource allocation algorithm based on DDQN

Require: δ, γ, E_{min}, σ^2, H_u, H_a
Ensure: Optimal resource allocation strategy

1: Initialise parameters of DDQN, state value and experience buffer R
2: **for** *episode* $\leftarrow$ 1 to N **do**
3:     initialsie system state;
4:     **for** `each time from [1,T]` **do**
5:         an action a_t is chosen ranomly by epsilon greedy approach
6:         execute state-action pair and and the reward r_t and next state s_{t+1} is observed
7:         (s_t, a_t, r_t, s_{t+1}) is stored in experience buffer M ;
8:         Get samples from the experience buffer M.
9:         Calculate the target Q value using $\mathrm{Q}(\mathrm{s}_t, a_t) = V(s_t;\ \theta, \beta) + [A(s_t, a_t;\ \theta, \alpha) - max_{a'} A(s_t, a_t;\ \theta, \alpha)]$
10:         Calculate the loss value by the loss function $L(\theta) = \mathbb{E}[(r_t + \theta \cdot max_a Q(s_{t+1}, a_{t+1};\ \theta) - Q(s_t, a_t;\ \theta))^2]$
11:         Update e and parameters with Adam Optimizer.
12:     **end for**
13: **end for**
14: The optimal power allocation strategy π^* is obtained

---

5 Performance Evaluation and Analysis

5.1 Simulation Environment

For simulation experiments and analysis, we used tyrone DIT400TR-48RL workstation with 128 GB RAM which configure NVIDIA Quadro RTX 5000 GPU card on intel-C621 chipset. We perform extensive experiments in order to give proper argument of our system model. We employ conda environment and packages on Linux to run python to assess the performance of the proposed model of UAV-assisted 5G mMTC resource allocation approach. Convolutional Neural Network (CNN) is used in the model which comprises of two hidden layers with 256 and 64 neurons in each. The memory bank's value is determined by its memory capacity. At any one time, the batch size refers to the amount of samples gathered from the memory bank. Memory and batch size have an impact on DDQN's accuracy and training rate. We created simulation settings that are compliant with 5G specifications and recommendations as per 3GPP standards. The system begins with a single UAV and a large number of base stations. The system radius has been adjusted at 500 metres in our scenario. We have shown some important simulation parameter used in the Table 2.

Table 2 The setup of hyper-parameter for Dueling DQN, DQN and Q-network training

Parameter	Value
r_{uav}	-60 dB
r_{bs}	-90 dB
α	2.5
β	2.5
UAV bandwidth	32 MHz
BS bandwidth	23 MHz
batch size	128
memory size	4096
epochs	10000
UAV maximum power	32 dBm
BS maximum power	23 dBm
number of users	15
number of base stations	4
learning rate (δ)	0.01
discount factor (γ)	0.99

5.2 Simulation results and parameter analysis

In this section, we discuss the simulation results and performance analysis by investigating the consumption, throughput of power and the energy efficiency of the systems. We have considered emergency communication scenario in which the communication link has broken due to disaster. An UAV works as a base station in this region to establish the communication connection with the outside since the ground base stations are mostly power driven and might disturb in the disaster. The UAV-assisted 5G UDN network has been considered for the resource allocation which aim to maximize the energy efficiency. We simulated the proposed Dueling DQN model with 10000 epochs. The learning rate and discount factor are 0.01 and 0.99, respectively. We perform the experiments with 15 users, 4 ground base stations and 1 UAV.

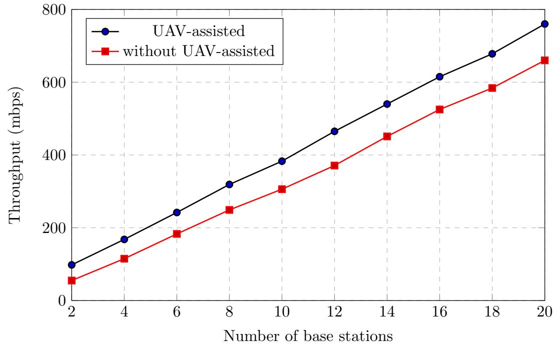

Fig. 4 The system throughput vs number of base stations

The Fig. 4 shows the system throughput with the different number of base stations as (2 , 4 , 6 , . . . , 20) in which UAV-assisted network provides a better throughput with compare to the without UAV-assisted network. With the increasing in number of base stations the UAV-assisted provides better overall throughput than the normal environment.

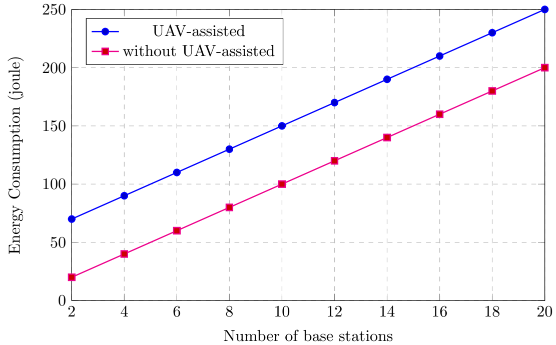

Fig. 5 The system energy consumption vs number of base stations

The Fig. 5 shows the energy consumption with the different number of base stations as (2 , 4 , 6 , . . . , 20) in which UAV-assisted network consume more energy than without UAV-assisted network. The plot is linear because of we provides uniform type of UAV and base station power in the system model. The operational expenses and throughput are two significant performance indicators, as excessive consumption raises operating costs and insufficient throughput negatively impacts the user's experience. The energy efficiency vs different users for different Q-learning algorithms as shown in Fig. 6. We have considered the number of users as (10, 15, 20, 25, 30) with DDQN model and shown comparison with DQN, Q-Learning and Random algorithms. The performance with 10 users is better than the other because of low dense network environment. With the increasing the number of users the overall energy efficiency and the DDQN model performs better in each of the case.

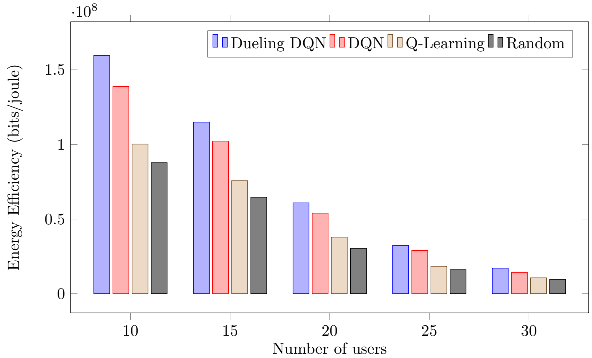

Fig. 6 The system energy efficiency for different users

The different learning rate of DDQN and DQN model for energy efficiency is shown in the Fig. 7. We consider the learning rate as (0.001, 0.01, 0.1) in which the learning rate 0.01 performs better while learning rate 0.001 takes more iterations for stable performance. We fix the learning rate 0.01 for the rest of our experiments. We run for 10000 episodes for analysis but the models DDQN and DQN both performs stable about within 1000 episodes.

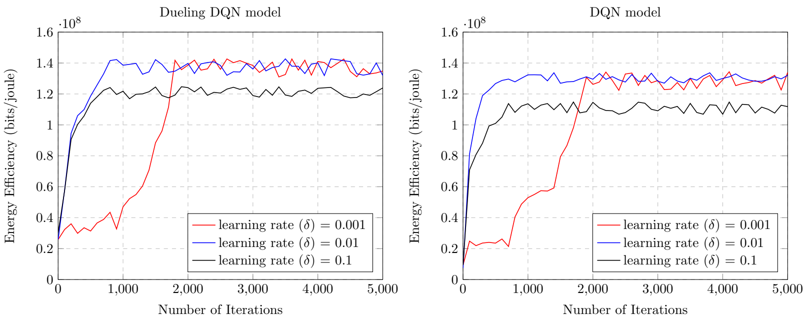

Fig. 7 The different learning rates for DDQN and DQN model

The one of the most essential feature of reinforcement learning algorithm is the learning rate, changing it affects the neural network's weight and changing the depth affects the algorithm's performance. If learning rate tends to 0, most recent feedback function may not be acquired by the agent. It will take a huge time to iterate, and the convergence speed will be slow. If it is close to 1, on the other hand, the speed of convergence will be too fast to get the ideal allocation strategy of resource, resulting in system performance decrease. The performance of learning rate 0.1 is not good as other parameters. Hence, we have considered 0.01 as the learning rate value for our experiments. The overall performance of DDQN model is better than the DQN model. Therefore, we have considered DDQN model for our experiments and the remaining reinforcement learning technique used for comparison.

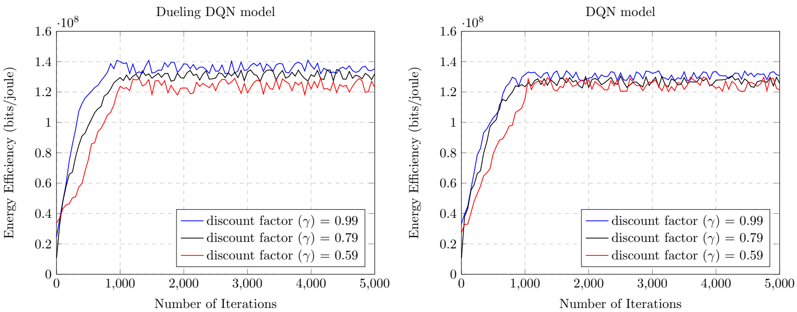

Fig. 8 The different discount factors for DDQN and DQN model

The different discount factor of DDQN and DQN model for energy efficiency is shown in the Fig. 8. We consider the discount factor as (0.99, 0.79, 0.59) in which the discount factor 0.99 performs better while the discount factor 0.79 and 0.59 provides less performance. Hence, we have considered 0.99 as the discount factor for our experiments. Also, the overall performance of DDQN model is better than the DQN model. Therefore, we have considered DDQN model for our experiments and the remaining reinforcement learning technique used for comparison.

5.3 Performance comparisons

We study the UAV-assisted framework in the 5G mMTC where a lot of connected IoT devices as user which is connected to the ground base stations. The ground base stations can be compromised in the disaster due to main power supplied. We considered the UAV-assisted base station which can connected to the outside region for emergency communication. The UAV can have the line-of-sight (LoS) in which it could be connected to the users. In the Fig. 9 we have shown the UAV throughput vs different θLoS range for the given devices and UAV power and hover-height. The UAV provides better throughput when the hover-height is less and UAV, and devices power is more. Here, by increasing the θLoS degree range the throughput performance of two network is increases, this is very interesting phenomenon because widely distributed users are close to the NLoS region [27]. Furthermore, by increasing the distance between user and UAV decreases the link gain, the larger θLoS decreases the interference to the user.

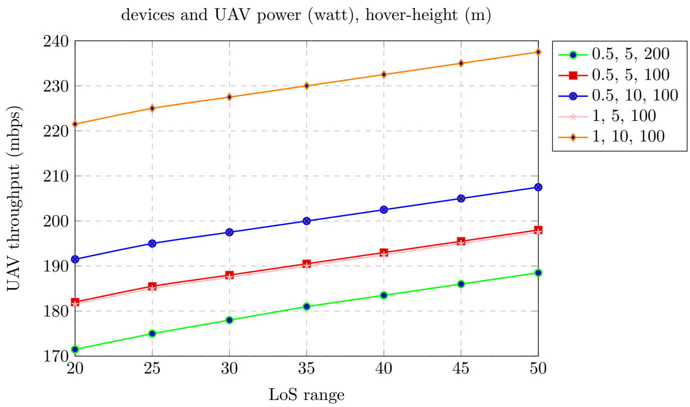

Fig. 9 Maximum line-of-sight θLoS range (degree)

The transmission power of the UAV-assisted base station has major role which can effect the overall wireless communication. The Fig. 10 shows the system's performance with power transmission and with some mentioned different users. It is observed that with the increment in power transmission there is a decrements in energy efficiency of the system. We have used the transmission power as 3 dBm with the 15 users for our rest of the experiments.

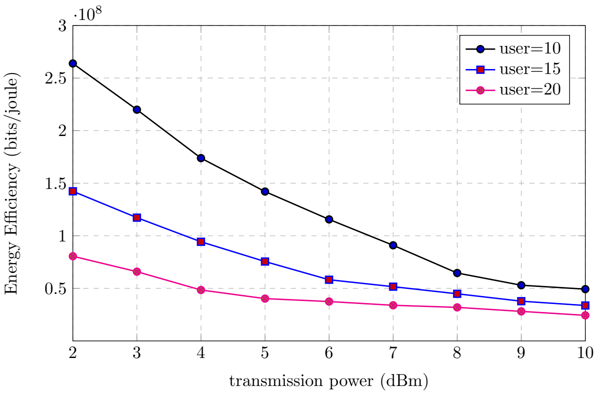

Fig. 10 The system energy efficiency for different power transmission rate

We simulated our system models with Dueling DQN, DQN, Q-Learning, Random distribution algorithms for better performance comparison. In the Fig. 11

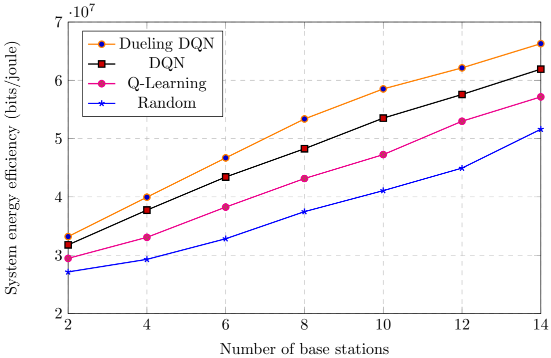

Fig. 11 The system energy efficiency for different number of base stations

we shows the energy efficiency with different number of users. The peak distribution indicates that the base station sends data at the highest transmission power possible to the user. In which Dueling DQN provides better energy efficiency as compared to the DQN and Q-Learning model. The term random refers to the fact that the base station provides transmission power to the user at random.

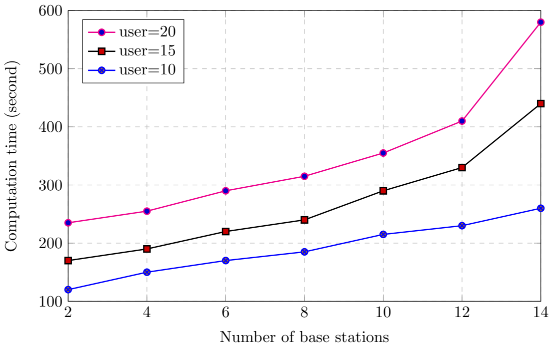

Fig. 12 The computation time for different number of base stations

The computation time is an important parameter by which we can think towards the scalability. We perform the computation analysis with the DDQN algorithm as shown in the Fig. 12. The number of base stations we considered as (2, 4, 6, 8, 10, 12, 14). The number of user can be dynamic which can be change with interval of time, we have simulated the computation time with numbers of users as 10, 15 and 20. The computation time is more for user 20 and for each user groups, with the increasing the number of base stations the computation time increases.

6 Conclusion

In this paper, we study UAV-assisted 5G mMTC slice for emergency communication in disaster scenario. We investigate resource allocation maximization problem in our environment of UAV-assisted wireless network. We consider UAV as a flying base station for the emergency communication system with 5G mMTC Network Slicing to overcome communication limits induced by natural disasters. We formulate our problem to improve overall energy efficiency and we separate the problem of resource allocation into two modules as user link selection strategy and power control method. Then, we reduce the problem into stochastic optimization problem using Markov Decision Process (MDP) theory. We proposed Dueling Deep Q-Network (DDQN) based algorithm based on reinforcement learning for dynamic resource allocation. We perform extensive experiments with proposed model, Q-Learning and DQN in order to present better analysis. We found that the overall performance of DDQN model is better. In the future, we will look at various deep reinforcement learning techniques and other UAVs as the auxiliary equipment, such as UAVs that act as handover and UAVs that act as the wireless charging equipments.

References

[1] Navarro-Ortiz, J., Romero-Diaz, P., Sendra, S., Ameigeiras, P., Ramos-Munoz, J. J., Lopez-Soler, J. M. (2020). A survey on 5G usage scenarios and traffic models. IEEE Communications Surveys Tutorials, 22(2), 905-929.

[2] Rafique, W., Qi, L., Yaqoob, I., Imran, M., Rasool, R. U., Dou, W. (2020). Complementing IoT services through software defined networking and edge computing: A comprehensive survey. IEEE Communications Surveys Tutorials, 22(3), 1761-1804.

[3] Gupta, R. K., Sahoo, B. (2018). Security Issues in Software-Defined Networks. IUP Journal of Information Technology, 14(2).

[4] Bonati, L., Polese, M., D'Oro, S., Basagni, S., Melodia, T. (2020). Open, programmable, and virtualized 5G networks: State-of-the-art and the road ahead. Computer Networks, 182, 107516.

[5] Liu, X., Chen, M., Liu, Y., Chen, Y., Cui, S., Hanzo, L. (2020). Artificial intelligence aided next-generation networks relying on UAVs. IEEE Wireless Communications, 28(1), 120-127.

[6] Chen, S., Liang, Y. C., Sun, S., Kang, S., Cheng, W., Peng, M. (2020). Vision, requirements, and technology trend of 6G: How to tackle the challenges of system coverage, capacity, user data-rate and movement speed. IEEE Wireless Communications, 27(2), 218-228.

[7] Li, B., Fei, Z., Zhang, Y. (2018). UAV communications for 5G and beyond: Recent advances and future trends. IEEE Internet of Things Journal, 6(2), 2241-2263.

[8] Wang, B., Sun, Y., Sun, Z., Nguyen, L. D., Duong, T. Q. (2020). UAV-assisted emergency communications in social IoT: A dynamic hypergraph coloring approach. IEEE Internet of Things Journal, 7(8), 7663-7677.

[9] Khawaja, W., Guvenc, I., Matolak, D. W., Fiebig, U. C., Schneckenburger, N. (2019). A survey of air-to-ground propagation channel modeling for unmanned aerial vehicles. IEEE Communications Surveys Tutorials, 21(3), 2361-2391.

[10] Yan, K., Ma, L., Zhang, Y. (2020, December). Research on the Application of 5G Technology in UAV Data Link. In 2020 IEEE 9th Joint International Information Technology and Artificial Intelligence Conference (ITAIC) (Vol. 9, pp. 1115-1118). IEEE.

[11] Li, B., Fei, Z., Zhang, Y. (2018). UAV communications for 5G and beyond: Recent advances and future trends. IEEE Internet of Things Journal, 6(2), 2241-2263.

[12] Mozaffari, M., Saad, W., Bennis, M., Nam, Y. H., Debbah, M. (2019). A tutorial on UAVs for wireless networks: Applications, challenges, and open problems. IEEE communications surveys tutorials, 21(3), 2334-2360.

[13] Xilouris, G. K., Batistatos, M. C., Athanasiadou, G. E., Tsoulos, G., Pervaiz, H. B., Zarakovitis, C. C. (2018, December). UAV-assisted 5G network architecture with slicing and virtualization. In 2018 IEEE Globecom Workshops (GC Wkshps) (pp. 1-7). IEEE.

[14] Li, Y., Zhang, H., Long, K., Jiang, C., Guizani, M. (2021). Joint resource allocation and trajectory optimization with QoS in UAV-based NOMA wireless networks. IEEE Transactions on Wireless Communications.

[15] Volk, M., Sterle, J. (2021). 5G Experimentation for Public Safety: Technologies, Facilities and Use Cases. IEEE Access, 9, 41184-41217.

[16] SHU, Z., TALEB, T., SONG, J. (2021). Resource Allocation Modeling for Fine-Granular Network Slicing in Beyond 5G Systems. IEICE Transactions on Communications.

[17] Gupta, R. K., Misra, R. (2019, December). Machine learning-based slice allocation algorithms in 5G networks. In 2019 International Conference on Advances in Computing, Communication and Control (ICAC3) (pp. 1-4). IEEE.

[18] Zhang, X., Zhang, Z., Yang, L. (2021). Joint User Association and Power Allocation in Heterogeneous Ultra Dense Network via Semi-Supervised Representation Learning. arXiv preprint arXiv:2103.15367.

[19] Guo, H., Liu, J., Zhang, J. (2018). Computation offloading for multi-access mobile edge computing in ultra-dense networks. IEEE Communications Magazine, 56(8), 14-19.

[20] Ullah, Z., Al-Turjman, F., Moatasim, U., Mostarda, L., Gagliardi, R. (2020). UAVs joint optimization problems and machine learning to improve the 5G and Beyond communication. Computer Networks, 107478.

[21] Song, J., Song, Q., Wang, Y., Lin, P. (2021). Energy-Delay Trade-off in Adaptive Cooperative Caching for Energy-Harvesting Ultradense Networks. IEEE Transactions on Computational Social Systems.

[22] Chen, Y., Liu, Z., Zhang, Y., Wu, Y., Chen, X., Zhao, L. (2020). Deep Reinforcement Learning-Based Dynamic Resource Management for Mobile Edge Computing in Industrial Internet of Things. IEEE Transactions on Industrial Informatics, 17(7), 4925-4934.

[23] Shi, Z., Liu, J., Zhang, S., Kato, N. (2021). Multi-Agent Deep Reinforcement Learning for Massive Access in 5G and Beyond Ultra-Dense NOMA System. IEEE Transactions on Wireless Communications.

[24] Chen, Y., Zhang, N., Zhang, Y., Chen, X., Wu, W., Shen, X. S. (2019). TOFFEE: Task offloading and frequency scaling for energy efficiency of mobile devices in mobile edge computing. IEEE Transactions on Cloud Computing.

[25] Faraci, G., Grasso, C., Schembra, G. (2019, April). Reinforcement-learning for management of a 5G network slice extension with UAVs. In IEEE INFOCOM 2019-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS) (pp. 732-737). IEEE.

[26] Faraci, G., Grasso, C., Schembra, G. (2020). Design of a 5G network slice extension with MEC UAVs managed with reinforcement learning. IEEE Journal on Selected Areas in Communications, 38(10), 2356-2371.

[27] Liu, Y., Xiong, K., Ni, Q., Fan, P., Letaief, K. B. (2019). UAV-assisted wireless powered cooperative mobile edge computing: Joint offloading, CPU control, and trajectory optimization. IEEE Internet of Things Journal, 7(4), 2777-2790.

[28] Wang, H., Zhao, H., Wu, W., Xiong, J., Ma, D., Wei, J. (2019). Deployment algorithms of flying base stations: 5G and beyond with UAVs. IEEE Internet of Things Journal, 6(6), 10009-10027.

[29] Yan, S., Peng, M., Cao, X. (2018). A game theory approach for joint access selection and resource allocation in UAV assisted IoT communication networks. IEEE Internet of Things Journal, 6(2), 1663-1674.

[30] Wang, W., Tang, J., Zhao, N., Liu, X., Zhang, X. Y., Chen, Y., Qian, Y. (2020). Joint precoding optimization for secure SWIPT in UAV-aided NOMA networks. IEEE Transactions on Communications, 68(8), 5028-5040.

[31] Wang, J., Liu, M., Sun, J., Gui, G., Gacanin, H., Sari, H., Adachi, F. (2020). Multiple Unmanned-Aerial-Vehicles Deployment and User Pairing for Nonorthogonal Multiple Access Schemes. IEEE Internet of Things Journal, 8(3), 1883-1895.

[32] Zeng, F., Hu, Z., Xiao, Z., Jiang, H., Zhou, S., Liu, W., Liu, D. (2020). Resource allocation and trajectory optimization for QoE provisioning in energy-efficient UAV-enabled wireless networks. IEEE Transactions on Vehicular Technology, 69(7), 7634-7647.

[33] Zhang, L., Ansari, N. (2020). Latency-aware IoT service provisioning in UAV-aided mobile-edge computing networks. IEEE Internet of Things Journal, 7(10), 10573-10580.

[34] Gupta, R. K., Ranjan, A., Moid, M. A., Misra, R. (2020, July). Deep-Learning Based Mobile-Traffic Forecasting for Resource Utilization in 5G Network Slicing. In International Conference on Internet of Things and Connected Technologies (pp. 410-424). Springer, Cham.

[35] Xie, L., Xu, J., Zhang, R. (2018). Throughput maximization for UAV-enabled wireless powered communication networks. IEEE Internet of Things Journal, 6(2), 1690-1703.

[36] Gupta, R. K., Choubey, A., Jain, S., Greeshma, R. R., Misra, R. (2020, July). Machine Learning Based Network Slicing and Resource Allocation for Electric Vehicles (EVs). In International Conference on Internet of Things and Connected Technologies (pp. 333-347). Springer, Cham.

[37] Liu, X., Wang, J., Zhao, N., Chen, Y., Zhang, S., Ding, Z., Yu, F. R. (2019). Placement and power allocation for NOMA-UAV networks. IEEE Wireless Communications Letters, 8(3), 965-968.

[38] Cai, Y., Cui, F., Shi, Q., Zhao, M., Li, G. Y. (2018). Dual-UAV-enabled secure communications: Joint trajectory design and user scheduling. IEEE Journal on Selected Areas in Communications, 36(9), 1972-1985.

[39] Wu, Y., Yang, W., Guan, X., Wu, Q. (2020). Energy-efficient trajectory design for UAV-enabled communication under malicious jamming. IEEE

Wireless Communications Letters, 10(2), 206-210.

[40] Li, J., Zhao, H., Wang, H., Gu, F., Wei, J., Yin, H., Ren, B. (2019). Joint optimization on trajectory, altitude, velocity, and link scheduling for minimum mission time in UAV-aided data collection. IEEE Internet of Things Journal, 7(2), 1464-1475.

[41] You, C., Zhang, R. (2020). Hybrid offline-online design for UAV-enabled data harvesting in probabilistic LoS channels. IEEE Transactions on Wireless Communications, 19(6), 3753-3768.

[42] Ali, M. A., Jamalipour, A. (2020). UAV placement and power allocation in uplink and downlink operations of cellular network. IEEE Transactions on Communications, 68(7), 4383-4393.

[43] Zhang, S., Ansari, N. (2020). 3D drone base station placement and resource allocation with FSO-based backhaul in hotspots. IEEE Transactions on Vehicular Technology, 69(3), 3322-3329.

[44] Qiu, C., Wei, Z., Yuan, X., Feng, Z., Zhang, P. (2020). Multiple UAV-mounted base station placement and user association with joint fronthaul and backhaul optimization. IEEE Transactions on Communications, 68(9), 5864-5877.

[45] Mozaffari, M., Saad, W., Bennis, M., Debbah, M. (2015, December). Drone small cells in the clouds: Design, deployment and performance analysis. In 2015 IEEE global communications conference (GLOBECOM) (pp. 1-6). IEEE.

[46] Xue, Z., Wang, J., Ding, G., Wu, Q., Lin, Y., Tsiftsis, T. A. (2018). Device-to-device communications underlying UAV-supported social networking. IEEE Access, 6, 34488-34502.

[47] Chen, X., Liu, X., Chen, Y., Jiao, L., Min, G. (2021). Deep Q-network based resource allocation for UAV-assisted ultra-dense networks. Computer Networks, 196, 108249.

[48] Liu, M., Yang, J., Gui, G. (2019). DSF-NOMA: UAV-assisted emergency communication technology in a heterogeneous Internet of Things. IEEE Internet of Things Journal, 6(3), 5508-5519.

[49] Yan, Z., Zhang, Z., Meng, Y. (2021, July). Energy Efficiency Optimization for UAV-assisted mMTC Networks with Altitude Differences. In 2021 IEEE/CIC International Conference on Communications in China (ICCC) (pp. 306-311). IEEE.

Statements and Declarations

Funding

The preparation of the manuscript is not funded.

Competing Interests

There are no competing interests declared by the authors.

Author Contributions

The study, implementation and experiments were done by all of the authors. Rohit Kumar Gupta prepared the materials, run the simulations, and conducted the analysis and Saubhik Kumar helped in the simulations under supervision of Dr Rajiv Misra. Rohit Kumar Gupta wrote the first draft of the manuscript, and all contributors provided feedback. The final manuscript was read and approved by all of the authors.

Data Availability

All the used codes/data in this research will be available from the corresponding author on reasonable request.