Received:

DOI:

RESEARCH ARTICLE

WILEY

Resource allocation of fog wireless access network based on deep reinforcement learning

Jingru Tan[1] | Wenbo Guan[2]

[1]School of Communication and Information Engineering, Xi'an University of Posts and Telecommunication, Xi'an, China

[2]School of Microelectronics, Xidian University, Xi'an, China

Correspondence
Jingru Tan, School of Communication and Information Engineering, Xi'an University of Posts and Telecommunication
Xi'an, China
Email: 747100768@qq.com

Abstract

Aiming at the problem of huge energy consumption in the Fog Wireless Access Networks (F-RANs), the resource allocation scheme of the F-RAN architecture under the cooperation of renewable energy is studied in this paper. Firstly, the transmission model and Energy Harvesting (EH) model are established, the solar energy harvester is installed on each Fog Access Point (F-AP), and each F-AP is connected to the smart grid. Secondly, the optimization problem is established according to the constraints of Signal to Noise Ratio (SNR), available bandwidth and energy harvesting, so as to maximize the average throughput of F-RAN architecture with hybrid energy sources. Finally, the dynamic power allocation scheme in the network is studied by using Q-learning and Deep Q Network (DQN) respectively. Simulation results show that the proposed two algorithms can improve the average throughput of the whole network compared with other traditional algorithms.

KEYWORDS

fog wireless access networks (F-RANs), renewable energy, resource allocation, deep reinforcement learning

1 | INTRODUCTION

In recent years, with the rapid development of wireless communication networks, people's demand for data traffic has increased rapidly. Cloud Radio Access Networks (C-RANs) can cover the data information of all User Equipments (UEs) in the cell, but how to effectively solve the lack of fronthaul link capacity and reduce the transmission delay of information are the primary problems for the networks [1-7]. Under this background, the F-RAN with cache nodes has appeared. There are four communication modes in the F-RAN: C-RAN mode, F-AP mode, Device to Device (D2D) mode and High Power Node (HPN) mode [8-11]. Compared with the C-RAN , F-RAN is equipped with file caching equipment in the F-APs and caches files with high popularity in advance. When a user requests a file, F-AP can quickly send the file to the UEs without requesting a file from the centralized baseband unit pool (BBU), which effectively relieves the pressure on the fronthaul link and reduces the request delay of high-speed users [12-21]. However, while F-RAN can effectively meet the Quality of Service (QoS) of UEs, it also brings a lot of energy consumption. Relevant data show that the current cellular network consumes about 6×1010kwh power every year in the world, and the energy consumption of base stations accounts for about 80% of the total power consumption of the network [22]. The energy consumption of the network affects its performance, so F-RANs need to improve energy efficiency. The key point to solve this problem is the cooperation between green communication and EH[23-28], so as to

 © 2021 John Wiley & Sons, Ltd
https://doi.org

improve economic benefits and reduce environmental pollution while transmitting information at high speed.

Until now, many recent researches on power control, energy pricing and energy efficiency of C-RANs with hybrid energy sources have attracted great interest from academia and industry [29-31], while there is fewer work on power distribution of F-RANs with hybrid energy sources. In [32], F-AP is divided into EH-FAP and Grid-FAP. To maximize the energy efficiency, the beamforming vector of the transmitter and the user-centric clustering scheme are jointly optimized under fronthaul capacity and transmission power constraints. On the premise of considering edge computing, reference [33] proposed an energy-saving F-RAN by studying user association and beamforming. Taking the maximization of energy efficiency as the optimization goal, the heuristic algorithm is used to solve the optimization problem, which can significantly improve the system energy efficiency. In [34], FAP mode and D2D mode are used as transmission modes, and system throughput and energy harvesting models are established. Under energy constraints, an optimization problem with the optimization goal of maximizing throughput is established, and a joint mode selection and power allocation algorithm are proposed. By setting up the transmit power threshold, the optimal transmit power is obtained when each D2D user pair selects FAP mode and D2D mode, and then the maximum throughput under energy coordination is obtained.

However, the research shows that the transmission power is one of the largest factors of energy consumption in F-APs [35-37], while the previous works [32-33] only took the transmission power as a constraint, not as the focus of optimization. Therefore, how to reasonably allocate the transmit power in the F-RANs of energy cooperation to meet the traffic needs of UEs need to be further studied. In addition, with the gradual increase in the number of UEs and F-APs, the traditional heuristic algorithm used in the optimization in [32-34] is not feasible due to the "dimension disaster". Reinforcement Learning (RL), as a new optimization theory emerging in recent years, has advantages over traditional optimization methods in exploring environment, machine learning, and intelligent decision-making.

Based on the above two points, this paper considers how to adopt an effective power allocation scheme in F-RANs from the perspective of green communication.This paper takes power allocation as the optimization variable, and uses Q-learning algorithm in RL and DQN algorithm in Deep Reinforcement Learning (DRL) to maximize the average throughput. Simulation results show that the power allocation schemes under the two algorithms can maximize the average throughput compared with the power allocation schemes under the traditional algorithms. The contributions of this paper are as follows: (1) The resource allocation scheme in F-RAN under renewable energy cooperation is considered. An optimization problem is proposed to maximize the average throughput while meeting the constraints of SNR, available bandwidth and energy harvesting. At the same time, the optimization model also considers the randomness of channel conditions and energy arrival. (2) In order to solve the complex convex optimization problem, taking Channel State Information (CSI), energy arrival information, and Energy Queue State Information (EQSI) as the system state, power allocation as the system action, and average throughput as the return function, two solutions based on Q-Learning and DQN are studied respectively. (3) The simulation results show that the two algorithms can improve the average throughput compared with the traditional algorithms.

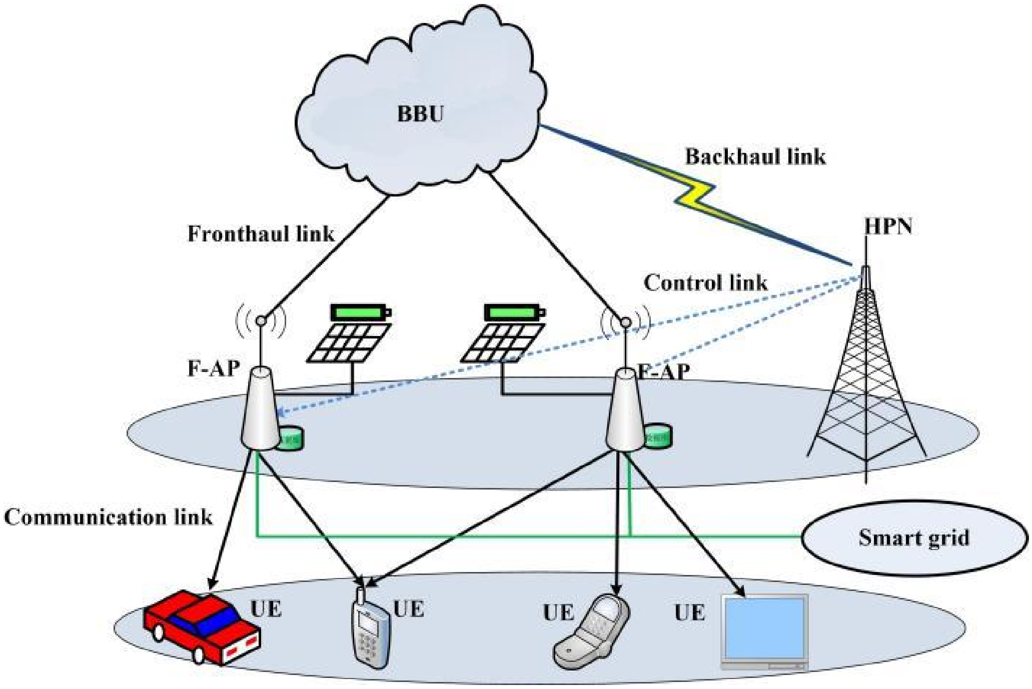
2 | SYSTEM MODEL

Figure 1 shows a downlink F-RAN architecture. In terms of information transmission, F-APs in the network are connected to the BBU pool through fronthaul links with limited capacity. In terms of energy consumption, F-APs include solar panels and energy storage batteries, and are connected to the smart grid. The BBU pool contains all content files that may be requested by UEs. As a fog access node with caching capability, F-APs can cache files with high popularity at the edge of the network.

When the files requested by the UEs are in the cache of F-APs, F-APs can directly transfer the files to the UEs, otherwise F-APs degenerate into Remote Ratio Heads(RRHs), and the system runs in C-RAN mode. Because the distance of the fronthaul link between BBU and F-APs is much greater than the distance of the wireless link between F-APs and UEs, obtaining the cache file from F-APs can reduce the transmission delay [38].

Assuming that the number of F-APs in the network is K , the number of antennas is N and the number of UEs is M , we use $\mathcal{K}=\{1,2,...,K\}$, $\mathcal{N}=\{1,2,...,N\}$ and $\mathcal{M}=\{1,2,...,M\}$ to represent their sets respectively. Considering the downlink transmission scenario of random arrival services in F-RAN, F-RAN works in discrete time slot $t \in \{0,1,2,...\}$.

FIGURE 1 System Model



2.1 | Transmission Model

In cellular mode, the signal received when UEm is connected to F-APk in time slot t can be expressed as:

$$y_{mk}(t)=\sqrt{L_{mk}(t)P_{mk}^{tr}(t)H_{mk}(t)}x_{mk}+\sum_{n\in\mathcal{M},n\neq m}\sqrt{L_{nk}(t)P_{nk}^{tr}(t)H_{nk}(t)}x_{nk}+z_m,\forall m\in\mathcal{M} \quad (1)$$

Where, the first item of the above formula is the useful signal received by the UEm; The second item is the signal sent by F-APk to other users, that is, the interference received by the UEm; The third item $z_m\sim\mathcal{CN}(0,\sigma_m^2)$ is the additive Gaussian white noise received by the UEm and σ_m^2 is defined as a fixed value in this paper. $L_{mk}(t)$ and $H_{mk}(t)$ represent large scale fading and small scale fading from the F-APk to the UEm respectively, $P_{mk}^{tr}(t)$ is the total transmission power of F-APk, x_{mk} is the service data stream corresponding to UEm, and its mean square value can be expressed as $\mathbb{E}\left\{|x_{mk}|^2\right\}=1,\forall m\in\mathcal{M}$.

The global CSI of the network is represented by $\mathbf{H}(t)=\left\{H_{mk}(t),\forall m\in\mathcal{M},k\in\mathcal{K}\right\}$. The assumption on channel model is as follows:

Assumption 1 (Channel Model) [39]: The global CSI is quasi-static in each frame. In particular, in t slot, each element of $H_{mk}(t)$ is an independent identically distributed discrete random variable with mean value of 0 and variance of σ_{mk}^2. Large scale fading remains unchanged in the communication process.

In t slot, the SNR of UEm can be expressed as:

$$\gamma_{mk}(t)=\frac{L_{mk}(t)P_{mk}^{tr}(t)H_{mk}(t)}{\sum_{n\in\mathcal{M},n\neq m}L_{nk}(t)P_{nk}^{tr}(t)H_{nk}(t)+\sigma_m^2} \quad (2)$$

According to Shannon formula, the downlink achievable rate of the UEm is:

$$R_{mk}(t)=B_{mk}\log_2\left\{1+\gamma_{mk}(t)\right\} \quad (3)$$

Where, B_{mk} is the available bandwidth.

The total throughput of the system under T time slots is:

$$C(t)=\sum_{m=1}^{\mathcal{M}}\sum_{k=1}^{\mathcal{K}}R_{mk}(t) \tag{4}$$

2.2 | Energy Harvesting Model

In this paper, solar panels and energy storage devices are installed on F-APs. F-APs can use the energy of smart grid and the green energy harvested by themselves. It is defined that in time slot *t*, the solar energy harvested by F-APs is $\mathbf{X}(t)=\{X_1(t),X_2(t),...,X_k(t)\}$, that is, $\mathbf{X}(t)$ is the energy achieved in time slot. The assumption on renewable energy collection is as follows:

Assumption 2 (Random Renewable Energy Model) [40]: In time slot *t*, random process $X_k(t)\geq 0$ is independent and identically distributed, and its average harvested energy is $\overline{X}_k(t)=\mathbb{E}[X_k(t)]$. In addition, the energy harvesting of each F-AP is independent of each other.

$\mathbf{E}(t)=\{E_1(t),E_2(t),...,E_k(t)\}$ is defined as EQSI, where $E_k(t)$ represents the renewable energy stored by F-AP*k* in time slot *t*. N_k^E is the maximum storage capacity of the energy queue buffer of F-AP*k*. When $N_k^E=E_k$, F-AP*k* will not harvest external solar energy. Therefore, the dynamic change of energy queue of F-AP*k* can be expressed as [40]:

$$E_k(t+1)=\min\left\{\left[E_k(t)-P_{k,e}(t)\right]^+ + X_k(t),N_k^E\right\} \tag{5}$$

Where, $X_k(t)$ is the newly arrived energy in time slot *t*, $P_{k,e}(t)$ is the renewable energy consumed in the current time slot, and $P_{k,e}(t)$ follows:

$$P_{k,e}(t)\leq E_k(t),\forall k \tag{6}$$

The power consumption $P_k(t)$ of F-AP*k* consists of static overhead $P_{k,s}$ (the static overhead is not affected by the transmission rate and is constant) and transmission power $P_k^{tr}(t)$, which can be expressed by the following formula:

$$P_k(t)=P_k^{tr}(t)+P_{k,s}\cdot 1,(P_k^{tr}>0) \tag{7}$$

Due to the instability of solar energy harvesting and the limitation of energy storage battery, renewable energy can only be used as a supplementary energy and can not completely replace the smart grid. Let $P_{k,g}(t)$ represents the smart grid energy consumed by F-AP*k*. Therefore, the total consumption of F-AP*k* can be expressed as:

$$P_k(t)=P_{k,e}(t)+P_{k,g}(t) \tag{8}$$

Therefore, the transmission power of F-AP*k* can be expressed as:

$$P_k^{tr}(t)=\left[P_{k,e}(t)-P_{k,s}\cdot 1,(P_k^{tr}>0)\right]^+ \tag{9}$$

$$P_k^{tr}(t)=\left[P_{k,e}(t)+P_{k,g}(t)-P_{k,s}\cdot 1,(P_k^{tr}>0)\right]^+ \tag{10}$$

Where, (9) indicates that the transmission energy consumption of F-AP*k* comes from the harvested green energy. (10) indicates that the transmission energy consumption of F-AP*k* comes from the harvested green energy and the electric energy of smart grid.

2.3 | Problem Formulation

This paper studies the resource allocation in F-RAN including energy harvesting. The optimization problems are as follows:

$$\begin{aligned}
&\max_{P_{mk}^{tr}} \lim_{T\to\infty} \frac{1}{T}\sum_{t=o}^{T-1} C(t) \\
s.t.\quad & C1: P_{mk}^{tr}(t) \le P_{\max}^{tr}(t), \forall m \in \mathcal{M}, k \in \mathcal{K} \\
& C2: B_{mk} \le B_{\max}, \forall m \in \mathcal{M}, k \in \mathcal{K} \\
& C3: \gamma_{mk} \ge \gamma_{QoS}, \forall m \in \mathcal{M}, k \in \mathcal{K} \\
& C4: 0 \le X_k(t) \le \chi_k(t) \le \chi_{\max}, \forall k \in \mathcal{K}
\end{aligned} \tag{11}$$

The constraint C1 limits the transmission power of F-APk, and $P_{\max}^{tr}$ is the maximum transmission power. The constraint C2 is the available bandwidth limit of F-APk, and $B_{\max}$ is the maximum bandwidth. The constraint C3 is the QoS constraint of the system, and γ_{QoS} is the lowest SNR that satisfies QoS. The constraint C4 represents the energy harvesting constraint of F-APk. Where χ_k is the solar energy in the surrounding environment, and $\chi_{\max}$ is the upper limit of solar energy change over time. In the optimization problem, the optimization goal of the system is to maximize the throughput, which is a strictly convex function. The optimization variable $P_{\max}^{tr}$ is a continuous variable. When solving the optimization problem, the complexity of the problem will increase rapidly with the increasing number of UEs. Therefore, a resource allocation algorithm based on DRL is proposed.

3 | DRL Resource Allocation Algorithm for Energy Harvest and Cooperation

This paper studies the whole transmission process, in which the CSI, currently harvested energy and EQSI of each time slot change dynamically. This requires network manager to make decisions according to the current state in the multi-dimensional system state. Because the system state changes dynamically with time, the traditional algorithms are difficult to solve the above optimization problem, and the machine can take the complex high-dimensional data as the input, so that the network manager can manage the system more efficiently. Therefore, this paper first introduces the Q-learning algorithm in RL for power allocation. Q-learning algorithm is used to convert the optimization goal into the maximum Q value, which stores the Q value of state-action through Q-table. However, because the optimization problem involves more complex continuous states, which will bring high computational complexity to the Q-table storage, consume a lot of memory, and make the traversal of the search state more time-consuming. Therefore, this paper proposes to use DQN algorithm to estimate the Q value of state-action. DQN combines Q-learning and neural network, and uses neural network to replace the Q-table in Q-learning, so it can achieve good performance in continuous large-scale networks[41].

3.1 | Solving Resource Allocation Problem Based on Q-Learning

In F-RAN, HPN acts as an agent to interact with the environment, and selects a reasonable power allocation for the information transmission between F-APs and UEs. The three elements of RL are defined as follows:

State space s_t : It is represented by the CSI, the energy currently harvested by the F-APs and the energy currently stored by the F-APs. The states of the agent in each time slot are as follows:

$$\begin{aligned}
s_t &= \{\mathbf{H}(t), \mathbf{X}(t), \mathbf{E}(t)\} \\
&= \begin{Bmatrix} H_1(t), H_2(t), ..., H_k(t), \\ X_1(t), X_2(t), ..., X_k(t), \\ E_1(t), E_2(t), ..., E_k(t) \end{Bmatrix}
\end{aligned} \tag{12}$$

Action space a_t : It is expressed by the transmission power of F-APs. And the constraint C2 need to be satisfied by a_t .

$$a_t = \{P_1^{tr}(t), P_2^{tr}(t), ..., P_k^{tr}(t)\} \tag{13}$$

Reward function r_t : It is expressed by the total throughput in the network at each time.

$$r_t = C(t) \tag{14}$$

Q-learning is a classical RL algorithm. The agent performs action a_t according to the current state space s_t, then obtains the cumulative return $Q(s_t, a_t)$, and stores its calculation in the Q-table. The correction formula of Q value is [41]:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha\left[r_{t+1} + \lambda \max_a Q(s_{t+1}, a') - Q(s_t, a_t)\right] \tag{15}$$

Where, s_t and a_t represent the current state and action respectively; s_{t+1} and a_{t+1} respectively represent the state and action at the next time; a' is all possible action values in state s_{t+1}; α represents the learning factor, and λ represents the discount factor, and the values of both satisfy [0,1]. Specially, when $\lambda \to 0$, it means that the agent focuses more on the current return value; When $\lambda \to 1$, it represents the future return value and is also the focus of the agent.

In Q-learning, the strategy that the agent selects action a_t according to state s_t is called learning strategy. In this paper, ε-greedy strategy is used as the learning strategy of Q-learning, and its expression is [42]:

$$\pi^*(s_t) = \begin{cases} \arg\max Q(s_t, a), \delta < \varepsilon \\ a_{rand} \qquad ,\delta \geq \varepsilon \end{cases} \tag{16}$$

Where, $0 \leq \varepsilon \leq 1$ is the greedy value. When $\varepsilon \to 1$, it means that the learning strategy focuses more on action selection according to the maximum Q value. δ is a random number between 0~1. $\arg\max Q(s_t, a)$ indicates the action value when the Q value is the maximum. a_{rand} is the action value randomly selected by the agent.

By iterating on $Q(s_t, a_t)$, we can obtain the optimal power allocation. Algorithm 1 is the dynamic power allocation algorithm based on Q-Learning.

Algorithm 1 Dynamic power allocation algorithm based on Q-Learning

Input：Give learning factor α and return matrix, and initialize $Q(s_t, a_t) = 0$

1：For episode =1, 2, ... M do

2： Initialize current state s_t

3： For t =1, 2, ... T do

4： Among all the actions that can be selected in the current state s_t, select the action a_t with the largest return according to the ε-greedy strategy to obtain r_t

5： Take action a_t， observe the next state s_{t+1}

6： Obtain $Q(s_t, a_t)$ according to formula (15)

7： Update states and actions $s_t = s_{t+1}$， $a_t = a_{t+1}$ until all $Q(s_t, a_t)$ converge

8： End for

9：End for

Output：According to the optimal allocation policy $\pi^*(s_t) = \arg\max_a Q(s_t, a)$, the throughput corresponding to the policy is obtained

3.2 | Solving Resource Allocation Problem Based on DQN

Both DQN algorithm and Q-learning algorithm are based on value iteration, but DQN algorithm combines Q-learning with neural network, takes system state and action as the input of neural network, and then generates Q value. This eliminates the need for a large Q-table to record Q values.Therefore, based on Q-learning, DQN is used to solve the optimization problem.

First, the DQN algorithm uses the parameter θ of the deep neural network to approximate the Q function to the optimal Q value, which can be expressed as [43]:

$$Q(s_t, a_t, \theta) \approx Q^*(s_t, a_t) \tag{17}$$

DQN uses two neural networks: Target Network (TN) and Estimated Network (EN). TN uses the previous parameter θ^* to predict the realistic value of Q. $Q^*(s_t, a_t, \theta^*)$ represents the output of TN, which is used to solve the target Q value. EN uses the latest parameter θ to predict the estimated value of Q. $Q(s_t, a_t, \theta)$ represents the output of current EN, which is used to evaluate the value function of state-action in the current system. In each training period of DQN, the agent will use the ε-greedy strategy to select the next action, then obtain the immediate reward r_t and the state s_{t+1} of the next time slot, and finally store the quadruple information (s_t, a_t, r_t, s_{t+1}) in memory D. EN obtains samples from the memory in each time slot to update parameter θ and make $\theta^* = \theta$.

The current target value can be expressed as:

$$y_j = \begin{cases} r_j & \text{, for terminal } s_{j+1} \\ r_j + \lambda \max Q(s_{j+1}, a_{j+1}, \theta^*), & \text{otherwise} \end{cases} \tag{18}$$

Because DQN estimates the Q value through neural network, there is a certain difference between this value and the Q value obtained by Bellman equation. In order to minimize this difference, a loss function is introduced:

$$Loss(\theta) = E\left[y_j - Q(s_j, a_j; \theta)\right]^2 \tag{19}$$

Algorithm 2 is the dynamic power allocation algorithm based on DQN.

Algorithm 2 Dynamic power allocation algorithm based on DQN

Input：Initialize replay memory D to capacity N.
Initialize θ^* and θ.

1：For episode =1, 2, ... M do
2： Initialize current state s_t
3： For t =1, 2, ... T do
4： Among all the actions that can be selected in the current state s_t, select the action a_t with the largest return according to the ε-greedy strategy to obtain r_t
5： Take action a_t， observe the next state s_{t+1}
6： Store transition (s_t, a_t, r_t, s_{t+1}) in D
7： IF D is full, do:
8：   Sample a part of data (s_j, a_j, r_j, s_{j+1}) uniformly from D
9： Let the output of dqn be formula (18)
10： Update parameter θ according to equation (19) until convergence
11：   End for
12：End for

Output：According to the optimal allocation policy $\pi^*(s_t) = \arg\max_a Q(s_t, a; \theta)$, the throughput corresponding to the policy is obtained

4 | Simulation Results and Analysis

In this section, the experimental simulation of resource allocation under hybrid energy cooperative in F-RAN is carried out, and the proposed two algorithms are compared with the traditional greedy algorithm and random allocation algorithm under different states change. The simulation results show that the proposed two algorithms can better improve the average throughput of the system. This section indicates the advantages of the proposed Q-learning algorithm and DQN algorithm.

4.1 | Parameters Setting

In order to verify the superiority of the proposed resource allocation algorithm, this section uses MATLAB for numerical simulation. The simulation scenario is shown in Figure 2. Set the coverage of HPN to a circular area with a diameter of 1000m, and place 4 F-APs and 8 UEs in this area. The solar energy harvesting on F-AP is simulated by using system advisory model. Other simulation parameters are shown in Table 1.

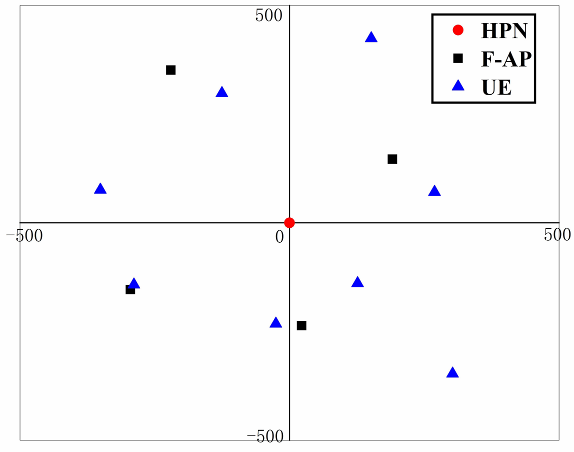


FIGURE 2 System Model

TABLE 1 Parameters setting

Parameters	Value
The number of F-AP	4
The transmission loss of F-AP	140.7+36.7log(km)
The maximum transmit power of F-AP	38dBm
Noise power spectral density	-174dBm/Hz
The maximum bandwidth	50MHz
The maximum capacity of memory	40000
Total training time	500
The number of neurons	400
Learning rate	0.005

4.2 | Simulation Results and Analysis

Figure 3 analyzes the average throughput changes of the proposed DQN algorithm and Q-learning algorithm under different channel gains, and compares the simulation results with the traditional greedy algorithm and random allocation algorithm. It can be seen from that the average throughput under the four algorithms increase with the increase of channel gain. This is because when the channel conditions are better, the UEs are allocated more power, which increases the average throughput. At the same time, the average throughput of the two proposed algorithms is higher than that of the two conventional algorithms, and the average throughput of the random allocation algorithm is the lowest. This is because agents can intelligently allocate energy by learning the system state (transfer information of channel conditions), while the allocation strategies of greedy algorithm and random allocation algorithm are fixed and do not consider the system state and other information. Therefore, their average throughput is lower than DQN algorithm and Q learning algorithm.

FIGURE 3 Variation of average throughput with channel gain

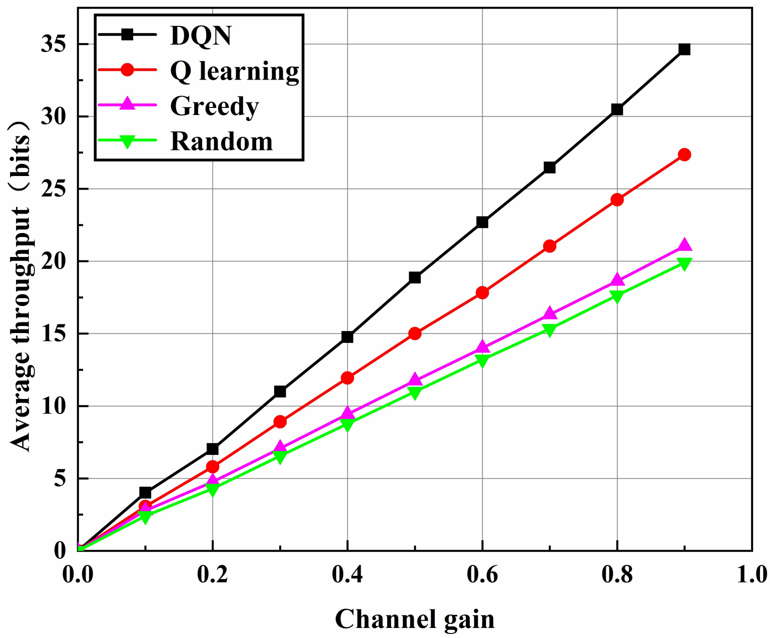


Figure 4 shows the relationship between learning rate and average throughput, where the channel gain is 0.7. It can be seen from that the average throughput of both algorithms increases with the decrease of learning rate. This is because the smaller the learning rate, the more inclined the agent will learn new actions to get rewards, so that the proportion of learned actions becomes smaller. In other words, the reduction of learning rate will promote the agent to explore new actions, which makes it easier to find a better power allocation scheme. At the same time, the excessive learning rate reduces the learning speed of the agent and prolongs the training time. Therefore, the appropriate learning rate should be selected for power allocation.

FIGURE 4 Variation of average throughput with learning rate

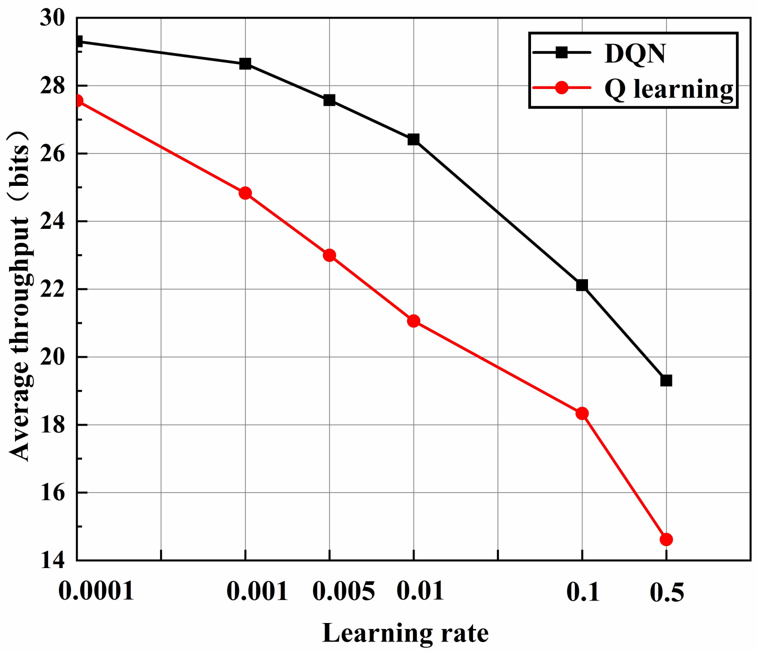


Figure 5 shows the average throughput of the four algorithms under different $X_{\max}$, where the channel gain is 0.7. It can be seen that the average throughput of the four algorithms increases with the increase of solar energy harvesting. This is because F-APs can obtain more energy from the environment for information transmission, which increases the average throughput. At the same time, under any $X_{\max}$, the average throughput obtained by DQN algorithm is the largest, and the average throughput of greedy algorithm and random allocation algorithm is the lowest. This is also because the agent can intelligently distribute the renewable energy obtained in the system according to the system state. It can also be seen that with the increase of X_{max} , the average throughput under the random allocation scheme tends to be stable. This is because the F-APs can transmit limited information capacity under this scheme and cannot use the excess energy to communicate with UEs. This results in the energy redundancy harvested by F-APs, so that the average throughput can not continue to improve.

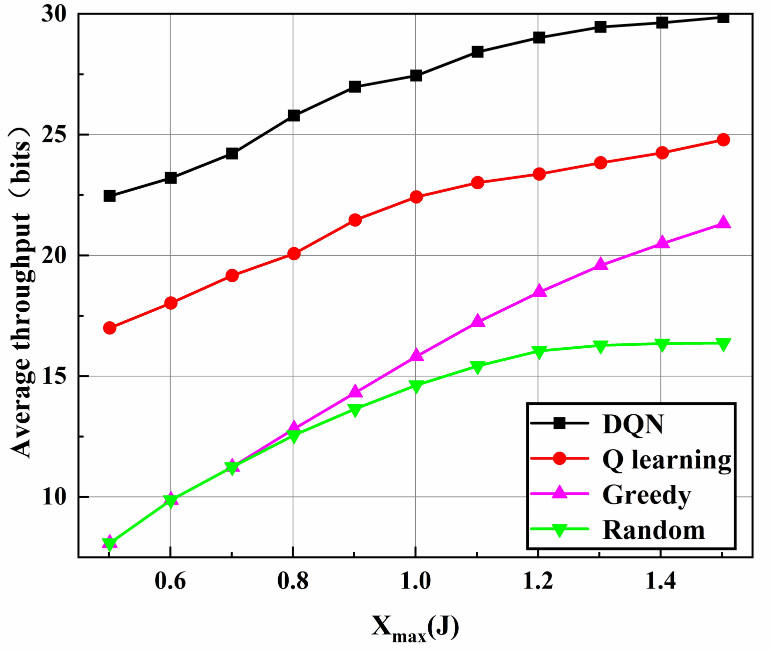


FIGURE 5 Variation of average throughput under different X_{max}

Figure 6 shows the average throughput of the four algorithms under different E_{max} , where the channel gain is 0.7. It can be seen that the average throughput of the four algorithms increases with the increase of battery storage capacity. This is because the energy stored in F-APs increases, so that the system can have more energy for power distribution. Therefore, the average throughput increases. In addition, with the increase of E_{max} , the average throughput of the four algorithms tends to be stable. This is because at this time, the system can use all renewable energy for information transmission and no longer use smart grid energy. The stored energy overflows, so the average throughput no longer grows.

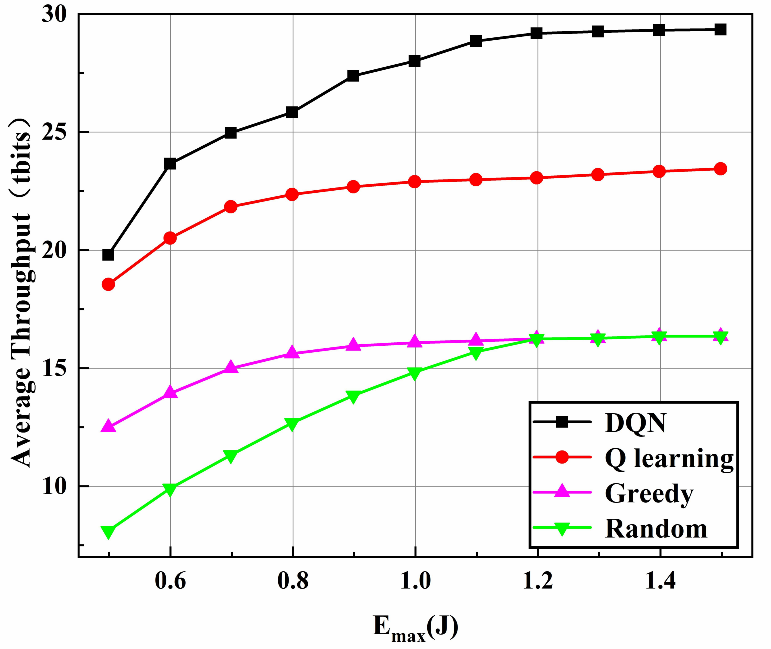


FIGURE 6 Variation of average throughput under different E_{max}

5 | Conclusion

In this paper, innovative research is carried out between F-RAN and energy harvesting cooperation. In the downlink scenario of F-RAN architecture, the dynamic power allocation based on energy harvesting is studied. Firstly, it is proposed to take solar energy as the first energy of F-RAN and the electric energy of smart grid as the second energy to provide electric energy guarantee for the operation of the network, so as to realize a hybrid energy cooperative F-RAN architecture. Secondly, the SNR, available bandwidth and energy harvesting are used as constraints, power allocation is used as optimization variables, and the optimization problem is established with the maximum average throughput as the optimization goal. In addition, due to the defects of the traditional optimization algorithm, the Q-learning algorithm in DL is used to solve the power allocation. At the same time, the state, action and return of DL in the proposed architecture are described. With the increase of the number of users in the network, the system puts forward higher requirements for the intelligence of the algorithm. This paper also uses DQN algorithm in DRL to solve power allocation. Finally, the two proposed algorithms are compared with the traditional greedy algorithm and random allocation algorithm in terms of channel gain, energy harvesting, and energy storage. The simulation results show that the proposed two algorithms can greatly improve the average throughput, and the average throughput under DQN algorithm is the highest.

In the future work , the harvesting rate of solar energy in the system is usually unchanged in a few minutes, but the change of channel state in the system is usually in milliseconds. Therefore, energy dispatch and resource allocation under two time scales are worth studying. In addition, in the F-RAN architecture, D2D users as energy limited devices also need to be studied on energy harvesting.

CONFLICT OF INTEREST

The author declares no potential conflict of interest.

ORCID

Jingru Tan https://orcid.org/0000-0003-0836-2420

REFERENCES

[1] CHECKO A, CHRISTIANSEN H L, YAN Y, et al. Cloud RAN for Mobile Networks—A Technology Overview. Communications Surveys & Tutorials. 2015, 17(1): 405-426.

[2] M. Jaber, M. A. Imran, R. Tafazolli and A. Tukmanov. 5G Backhaul Challenges and Emerging Research Directions: A Survey. IEEE Access. 2016, 4: 1743-1766 .

[3] M. Peng, C. Wang, V. Lau and H. V. Poor. Fronthaul-constrained cloud radio access networks: insights and challenges. IEEE Wireless Communications. 2015, 22(2): 152-160.

[4] M. Peng, Y. Sun, X. Li, Z. Mao and C. Wang. Recent Advances in Cloud Radio Access Networks: System Architectures, Key Techniques, and Open Issues. IEEE Communications Surveys & Tutorials. 2016, 18(3): 2282-2308.

[5] S. -H. Park, O. Simeone, O. Sahin and S. Shamai Shitz. Fronthaul Compression for Cloud Radio Access Networks: Signal processing advances inspired by network information theory. IEEE Signal Processing Magazine. 2014, 31(6): 69-79.

[6] I. A. Alimi, A. L. Teixeira and P. P. Monteiro. Toward an Efficient C-RAN Optical Fronthaul for the Future Networks: A Tutorial on Technologies, Requirements, Challenges, and Solutions. IEEE Communications Surveys & Tutorials. 2018, 20(1): 708-769.

[7] J. Tang, W. P. Tay, T. Q. S. Quek and B. Liang. System Cost Minimization in Cloud RAN With Limited Fronthaul Capacity. IEEE Transactions on Wireless Communications. 2017, 16(5): 3371-3384.

[8] PENG M G, SHI Y, ZHANG K C, et al. Fog Computing based Radio Access Networks: Issues and Challenges. IEEE Network, 2015, 30(4): 46-53.

[9] Tandon and O. Simeone. Harnessing cloud and edge synergies: toward an information theory of fog radio access networks. IEEE Communications Magazine. 2016, 54(8): 44-50.

[10]Y. Ku et al. 5G Radio Access Network Design with the Fog Paradigm: Confluence of Communications and Computing. IEEE Communications Magazine. 2017, 55(4): 46-52.

[11] S. Yan, M. Peng, M. A. Abana and W. Wang. An Evolutionary Game for User Access Mode Selection in Fog Radio Access Networks. IEEE Access. 2017, 5: 2200-2210.

[12] M. Peng and K. Zhang. Recent Advances in Fog Radio Access Networks: Performance Analysis and Radio Resource Allocation. IEEE Access. 2016, 4:5003-5009.

[13] SHIH Y Y, CHUNG W H, PANG A C, et al. Enabling Low-Latency Applications in Fog-Radio Access Networks. IEEE Network, 2017, 31(1): 12-18.

[14] R. Tandon and O. Simeone. Cloud-aided wireless networks with edge caching: Fundamental latency trade-offs in fog Radio Access Networks. 2016 IEEE International Symposium on Information Theory (ISIT); 2019; Barcelona, Spain. pp. 2029-2033.

[15] S. Park, O. Simeone and S. Shamai. Joint cloud and edge processing for latency minimization in Fog Radio Access Network. 2016 IEEE 17th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC),2016; Edinburgh, UK. pp. 1-5.

[16] G. M. S. Rahman, M. Peng, K. Zhang and S. Chen. Radio Resource Allocation for Achieving Ultra-Low Latency in Fog Radio Access Networks. IEEE Access. 2018, 6:17442-17454.
[17] A. M. Girgis, O. Ercetin, M. Nafie and T. ElBatt. Decentralized coded caching in wireless networks: Trade-off between storage and latency. 2017 IEEE International Symposium on Information Theory (ISIT); 2017; Aachen, Germany. pp. 2443-2447.
[18] T. Chiu, A. Pang, W. Chung and J. Zhang. Latency-Driven Fog Cooperation Approach in Fog Radio Access Networks. IEEE Transactions on Services Computing. 2019, 12(5): 698-711.
[19] M. Mukherjee, Y. Liu, J. Lloret, L. Guo, R. Matam and M. Aazam. Transmission and Latency-Aware Load Balancing for Fog Radio Access Networks. 2018 IEEE Global Communications Conference (GLOBECOM); 2018; Abu Dhabi, United Arab Emirates. pp. 1-6.
[20] T. Dang and M. Peng. Delay-Aware Radio Resource Allocation Optimization for Network Slicing in Fog Radio Access Networks. 2018 IEEE International Conference on Communications Workshops (ICC Workshops); 2018; Kansas City, MO, USA. pp. 1-6.
[21] S. Jia, Y. Ai, Z. Zhao, M. Peng and C. Hu. Hierarchical content caching in fog radio access networks: ergodic rate and transmit latency. China Communications. 2016, 13(12): 1-14.
[22] Yu L, Wu J, Fan P. Energy efficient designs of ultra-dense IoT networks with non-ideal optical front-hauls. IEEE Internet of Things Journal. 2019:1-1.
[23] Y. J. Xu, G. Q. Li, et al. Survey on resource allocation in heterogeneous wireless network. Journal of Chongqing University of Posts and Telecommunications(Natural Science Edition). 2018, 30(3): 289-299.
[24] X. Lu, P. Wang, D. Niyato, D. I. Kim and Z. Han. Wireless Networks With RF Energy Harvesting: A Contemporary Survey. IEEE Communications Surveys & Tutorials. 2018, 17(2): 757-789.
[25] X. Zhou, R. Zhang and C. K. Ho. Wireless Information and Power Transfer: Architecture Design and Rate-Energy Tradeoff. IEEE Transactions on Communications. 2013, 61(11): 4754-4767.
[26] Y. Mao, J. Zhang and K. B. Letaief. Dynamic Computation Offloading for Mobile-Edge Computing With Energy Harvesting Devices. IEEE Journal on Selected Areas in Communications. 2016, 34(12): 3590-3605.
[27] S. Bi, C. K. Ho and R. Zhang. Wireless powered communication: opportunities and challenges. IEEE Communications Magazine. 2015, 53(4): 117-125.
[28] C. K. Ho and R. Zhang. Optimal Energy Allocation for Wireless Communications With Energy Harvesting Constraints. IEEE Transactions on Signal Processing. 2012, 60(9): 4808-4818.
[29] L. Wang and S. Zhou. Flexible Functional Split and Power Control for Energy Harvesting Cloud Radio Access Networks. IEEE Transactions on Wireless Communications. 2020, 19(3): 1535-1548.
[30] LI Y D, YIN C C. Joint Energy Cooperation and Resource Allocation in C-RANs with Hybrid Energy Sources. 2017 IEEE/CIC International Conference on Communications in China (ICCC); 2017;Qingdao, China. pp. 1-6.
[31] Xu B, Zhu P, Li J, et al. Joint Long-term Energy Efficiency Optimization in C-RAN with Hybrid Energy Supply. IEEE Transactions on Vehicular Technology. 2020, PP(99):1-1.
[32] LIU B, PENG M G, Zhou Z. Joint Beamforming and Clustering Optimization of Hybrid-Energy-Powered eRRHs in F-RANs. 2018 IEEE/CIC International Conference on Communications in China (ICCC Workshops). 2018; pp. 319-323.
[33] LIU B, Peng M, Zhou Z. Joint Beamforming and Clustering Optimization of Hybrid-Energy-Powered eRRHs in F-RANs. 2018 IEEE/CIC International Conference on Communications in China (ICCC Workshops); 2018.
[34] D. M. Xu, J. R. Tan, et al. Resource Allocation in F-RAN Based on Renewable Energy Cooperation. Telecommunication Engineering. Accepted.
[35] Z. Hu, C. Hu, Z. Li, Y. Li and G. Wei. Power allocation for video segment based caching strategy in F-RAN architecture. China Communications. 2021, 18(2): 215-227.
[36] R. Rai, H. Zhu and J. Wang. Performance Analysis of NOMA Enabled Fog Radio Access Networks. IEEE Transactions on Communications. 2021, 69(1): 382-397.
[37] C. Liu, Y. Li, Z. Mao and T. Sha. On the Design of OFDM-Based Simultaneous Wireless Information and Power Transfer in Fog-Radio Access Networks. 2018 IEEE/CIC International Conference on Communications in China (ICCC Workshops); 2018; Beijing, China. pp. 272-276.
[38] Z. Zhao et al. On the Design of Computation Offloading in Fog Radio Access Networks. IEEE Transactions on Vehicular Technology. 2019, 68(7): 7136-7149.
[39] J. Li. Research on Theories and Methods of Traffic Queue-based Dynamic Radio Resource Optimization. Beijing University of Posts and Telecommunications. 2018
[40] H. Huang and V. K. N. Lau. Decentralized Delay Optimal Control for Interference Networks With Limited Renewable Energy Storage. IEEE Transactions on Signal Processing. 2012, 60(5): 2552-2561.
[41] B. Q. Huang. Research on the Reinforcement Learning Method and Its Application. Shanghai Jiao Tong University. 2007.
[42] M. Mobasheri, Y. Kim and W. Kim. Toward Developing Fog Decision Making on the Transmission Rate of Various IoT Devices Based on Reinforcement Learning. IEEE Internet of Things Magazine. 2020, 3(1): 38-42.
[43] S. Wang, H. Liu, P. H. Gomes and B. Krishnamachari. Deep Reinforcement Learning for Dynamic Multichannel Access in Wireless Networks. IEEE Transactions on Cognitive Communications and Networking. 2018, 4(2): 257-265.